THE HANDBOOK OF
GROUNDWATER
ENGINEERING

Editor-in-Chief

Jacques W. Delleur

The Handbook of Groundwater Engineering

The Handbook of Groundwater Engineering

Edited by

Jacques Delleur

School of Civil Engineering
Purdue University
West Lafayette, Indiana

Editor

Jacques W. Delleur received his Doctor of Engineering Science degree at Columbia University in 1955, his M.S.C.E. degree at Rensselaer Polytechnic Institute in Troy, New York, in 1950, and his Civil and Mining Engineer degree at the Universidad Nacional de Colombia (National University of Colombia) in 1949. In 1955, he joined Purdue University where he currently is Professor Emeritus of Environmental and Hydraulic Engineering and was Head of the Hydraulic and Systems Engineering Area in the School of Civil Engineering. Dr. Delleur taught intermediate and advanced graduate courses in subsurface hydrology, surface hydrology, statistical hydrology, and hydraulics. He founded the graduate program in Hydrology and Hydraulics in the School of Civil Engineering at Purdue. He is author or co-author of two books on hydrologic time series analysis. He is author or co-author of more than 60 papers in refereed journals, 70 papers in conference proceedings, and 60 technical reports. These cover the areas of subsurface hydrology, hydrologic modeling, stochastic hydrology, urban hydrology, and hydraulics. The most recent research publications related to groundwater co-authored by J.W. Delleur are concerned with the flow and transport of dissolved substances in groundwater and how they are affected by geologic heterogeneity.

Dr. Delleur's research has been supported by the U.S. Department of the Interior, the National Science Foundation, the U.S. Department of Transportation and the U.S. Department of Agriculture. He has served as an advisor to the U.S. Geological Survey, is a member of the international board of advisors of the American Society of Civil Engineers (ASCE) *Journal of Hydrologic Engineering* and is a member of the scientific council of the *Revue des Sciences de l'Eau / Journal of Water Science*. He served as a reviewer for the National Science Foundation, and for the scientific journals *Water Resources Research, Journal of Hydrology, Journal of the American Water Resources Association,* and the *Journal of Hydraulics*. He is a fellow of the Indiana Academy of Sciences, received the 1961 Freeman Fellow Award of the ASCE, in 1983 received an NSF/CNRS US-France Senior Scientist Exchange Award, and in 1992 received the Charles Harold Bechert Award of the Indiana Water Resources Association for significant contribution to the water resources profession in Indiana. While on sabbatical leave, Dr. Delleur did research in hydrology at the French National Hydraulics Laboratory (1968-69 and 1976-77), at the University of Grenoble, France (1961-62 and 1983-84) and at the Vrije Universiteit Brussel (Free University of Brussels), Belgium (1991). He has been a guest lecturer at the Ecole Polytechnique Fédérale de Lausanne (Federal Polytechnic School of Lausanne), Switzerland, at the Free University of Brussels, Belgium, at Imperial College in London, at the University of Tokushima, Japan, at the Indian Institute of Technology in Kanpur, India, at the Mahommadia School of Engineering in Rabat, Morocco, at the Taiwan National University in Taipei, Taiwan, and at the Universidad de los Andes (University of the Andes) in Bogota, Colombia.

Acknowledgments

I am grateful to this excellent group of authors, who are leaders in their respective specialties. Forty-four authors from six countries contributed their knowledge to this handbook. Their names and affiliations appear in the headings of the individual chapters and in the list of Contributors.

The chapters were reviewed in a manner similar to papers in scientific and professional journals. I am also very grateful to the twenty-five reviewers from five countries. They made a special effort to make certain that the information presented is sound. Some reviewers read two chapters. Their names appear in the list of Reviewers.

I wish to acknowledge the support I received from the School of Civil Engineering at Purdue University, in particular the encouragement of Dr. Vincent P. Drnevich, head of the School, and of my immediate colleagues in the Environmental and Hydraulic Engineering Area. I also wish to acknowledge the valuable assistance of the CRC Press staff, in particular Timothy Pletscher, Acquiring Editor, Engineering, and Gail Renard, Suzanne Lassandro, and Dawn Mesa.

<div align="right">

Jacques W. Delleur
Editor

</div>

Contributors

Rafael Angulo-Jaramillo
Laboratoire d'Etude des Transferts en Hydologie et
    Environnement
Institut National Polytechnique de Grenoble (INPG)
Grenoble, France

Ernest R. Blatchley, III
Environmental and Hydraulic Engineering Group
School of Civil Engineering
Purdue University
West Lafayette, Indiana

J. Boonstra
International Institute for Land Reclamation and
    Improvement
Wageningen, The Netherlands

Faycal Bouraoui
Laboratoire d'Etude des Transferts en Hydologie et
    Environnement
Institut National Polytechnique de Grenoble (INPG)
Grenoble, France

Barry R. Christopher
Independent Consultant
Roswell, Georgia

Mark A. Cushey
Department of Civil and Environmental Engineering
University of California
Berkeley, California

John H. Cushman
Departments of Applied Mathematics and Agronomy
Purdue University
West Lafayette, Indiana

Jacques W. Delleur
School of Civil Engineering
Purdue University
West Lafayette, Indiana

Basile Dendrou
ZEi Engineering Inc.
Annandale, Virginia

Stergios Dendrou
ZEi Engineering Inc.
Annandale, Virginia

F. De Smedt
InterUniversity Programme in Water Resources
    Engineering
Free University
Brussels, Belgium

Deborah Elcock
Environment Assessment Division
Argonne National Laboratory
Washington, DC

Bernard A. Engel
Department of Agricultural and Biological Engineering
Purdue University
West Lafayette, Indiana

R. David Espinoza
GeoSyntec Consultants
Columbia, Maryland

Timothy R. Ginn
Department of Civil and Environmental Engineering
University of California
Davis, California

David R. Green
Environmental Assessment Division
Argonne National Laboratory
Washington, DC

Milton E. Harr
School of Civil Engineering
Purdue University
West Lafayette, Indiana

Randell Haverkamp
Laboratoire d'Etude des Transferts en Hydrologie et
    Environnement
Institut National Polytechnique de Grenoble (INPG)
Grenoble, France

Michael F. Houlihan
GeoSyntec Consultants
Columbia, Maryland

Susan S. Hubbard
Department of Civil and Environmental Engineering
University of California
Berkeley, California

David W. Hyndman
Department of Geological Sciences
Michigan State University
East Lansing, Michigan

Vivek Kapoor
School of Civil and Environmental Engineering
Georgia Institute of Technology
Atlanta, Georgia

Peter K. Kitanidis
Environmental Fluid Mechanics and Hydrology
Department of Civil Engineering
Stanford University
Stanford, California

Leonard F. Konikow
National Center
U.S. Geological Survey
Reston, Virginia

P.A. Lapcevic
National Water Research Institute
Groundwater Remediation Project
Burlington, Ontario

Darrell I. Leap
Department of Earth and Atmospheric Sciences
Purdue University
West Lafayette, Indiana

Patrick C. Lucia
GeoSyntec Consultants
Columbia, Maryland

Kumar C.S. Navulur
Department of Agricultural and Biological Engineering
Purdue University
West Lafayette, Indiana

Loring F. Nies
School of Civil Engineering
Purdue University
West Lafayette, Indiana

K.S. Novakowski
National Water Research Institute
Groundwater Remediation Project
Burlington, Ontario

J.-Yves Parlange
Department of Agricultural and Biological Engineering
Cornell University
Ithaca, New York

Nancy L. Ranek
Environmental Assessment Division
Argonne National Laboratory
Washington, DC

Pedro C. Repetto
Woodward-Clyde International
Latin American Operations
Denver, Colorado

Thomas E. Reilly
National Center
U.S. Geological Survey
Reston, Virginia

Yoram Rubin
Department of Civil and Environmental Engineering
University of California
Berkeley, California

Tammo S. Steenhuis
Department of Agricultural and Biological Engineering
Cornell University
Ithaca, New York

E.A. Sudicky
Department of Earth Sciences
University of Waterloo
Waterloo, Ontario

John E. Thompson
Environmental and Hydraulic Engineering Group
School of Civil Engineering
Purdue University
West Lafayette, Indiana

John A. Veil
Environmental Assessment Division
Argonne National Laboratory
Washington, DC

Rony Wallach
Department of Soil and Water Science
The Hebrew University of Jerusalem
Israel

William B. White
Department of Geosciences
Materials Research Laboratory
The Pennsylvania State University
University Park, Pennsylvania

Amy Wilson
Department of Civil and Environmental Engineering
University of California
Berkeley, California

Christian Zammit
Laboratoire d'Etude des Transferts en Hydologie et
    Environnement
Institut National Polytechnique de Grenoble (INPG)
Grenoble, France

Jorge G. Zornberg
GeoSyntec Consultants
Columbia, Maryland

Preface

A substantial amount of new knowledge in the field of groundwater engineering has been accumulated over the last decade. In particular, many new developments have taken place in the analysis of contaminant transport both in the saturated and in the unsaturated zones. New techniques for the remediation of contaminated groundwater have been implemented. New approaches in computer modeling of the flow and movement of contaminants have appeared. The problem of modeling the heterogeneity of the geological formations and their effects on flow and transport continues to be a challenge and has been the object of many papers. This information is dispersed in many professional and scientific journals, textbooks, and reports. This new *Handbook of Groundwater Engineering* attempts to synthesize these facts and to provide this information in an easily accessible form. It contains a blend of professional practice and scientific information.

The typical readers of the handbook are expected to be geohydrologists and engineers possessing Bachelors or a Masters degrees in applied sciences or in engineering, possibly with some years of experience.

The first two chapters provide an introduction to the geological setting and the analysis of simple groundwater flow and contaminant transport problems. Chapters 8 and 9 are concerned with the production of groundwater and hydraulics of wells. The last six chapters (22 to 27) provide practical information on the solution of engineering problems concerned with pollution remediation, landfills, and the associated legal and regulatory aspects. The remaining 18 chapters deal with the analytical tools for the description and modeling of flow and transport phenomena in the saturated and in the unsaturated zones. Special attention is given to the effects of heterogeneity and the need to use probabilistic methods to account for the variability of the parameters.

I am grateful to have an excellent group of authors who are leaders in their respective specialties. Their names and affiliations appear in the headings of the individual chapters. The chapters were reviewed in a manner similar to papers in scientific and professional journals. I am also very grateful to the reviewers. They made a special effort to ascertain that the information presented is sound. Some reviewers read two chapters.

REVIEWERS

Pedro J. Alvarez, Department of Civil and Environmental Engineering, University of Iowa, Iowa City, IA 52242-1527.

Mary P. Anderson, Department of Geology and Geophysics, University of Wisconsin, Madison, WI 53706-1692.

Greg Batchelder-Adams, Woodward-Clyde Consultants, Denver, CO 80237.

Ralph Cady, U.S. Nuclear Regulatory Commission, Office of Nuclear Regulatory Research, Washington, DC 2055-0001.

Barry Christopher, Consultant, 210 Boxelder Lane, Roswell, GA 30076.

Derek Clark, Institute of Irrigation Studies, University of Southampton, Southampton, United Kingdom, SO17 1BJ.

Gedeon Dagan, Faculty of Mechanical Engineering, Tel Aviv University, Ramat-Aviv, Israel.

David E. Elrick, Department of Land Resources Sciences, University of Guelph, Guelph, Ontario, Canada, N1G 2W1.

Steven J. Fritz, Department of Earth and Atmospheric Sciences, Purdue University, West Lafayette, IN 47907.

Dimitry A. Grivas, Center for Infrastructure and Transportation Studies, Rensselaer Polytechnic Institute, Troy, NY 12180-3590.

Ahmed E. Hassan, Water Resources Center, Desert Research Institute, Las Vegas, NV 89119.

Paul Jen, Ground Water Protection Council, Republic, WA 99169.

Cindy Kao, Department of Civil and Environmental Engineering, Worcester Polytechnic Institute, Worcester, MA 01609.

Jeff Moore, GeoSyntec Consultants, Columbia, MD 21044.

John E. Mylroie, Department of Geosciences, Mississippi State University, Mississippi State, MS 39762.

Jean-Yves Parlange, Department of Agricultural and Biological Engineering, Cornell University, Ithaca, NY 14853-5701.

Eileen Poeter, Department of Geology and Geological Engineering, Colorado School of Mines, Golden, CO 80401.

Ward E. Sanford, U.S. Geological Survey, Reston, VA 22092.

Stephan E. Silliman, Department of Civil Engineering and Geological Sciences, University of Notre Dame, Notre Dame, IN 46556-0767.

Raghavan Srinivasan, Blackland Research Center, Texas A & M University System, Temple, TX 76502.

John Thackston, Woodward-Clyde Consultants, Denver, CO 80237.

Jane Wallace, GeoSyntec Consultants, Columbia, MD 21044.

F. M. Wilhelm, Department of Civil and Environmental Engineering, Northeastern University, Boston, MA 02115.

William Yeh, Department of Civil and Environmental Engineering, UCLA, Los Angeles, CA 90095-1593.

Wouter Zijl, Department of Geo-Energy, Netherlands Institute of Applied Geoscience TNO, National Geological Survey, 2600 JA Delft, The Netherlands.

Contents

Introduction

Since Darcy's pioneering work well over a century ago, our qualitative and quantitative understanding of the flow of water and transport of chemicals in the subsurface has skyrocketed. Much of our current understanding of flow processes has been developed over the past 30 years, and our understanding of transport processes has developed mainly over the last 15 years. Substantial recent advances have been largely driven by two factors: a) the development of high-speed computing, and b) the international focus on environmental quality.

Much of the groundwater work performed earlier in this century revolved around very simplistic conceptualizations of the subsurface, simplifications that were so drastic as to allow analytical solutions to flow problems and the subsequent development of type-curves. Consequently, well installation and design, and groundwater maintenance revolved around such idealizations as homogeneity, isotropy, and infinite domains, which, as we all know, are completely unrealistic assumptions.

With the advent of computing in the 1960s, through its modern refinements in software and hardware, and derivation of new sophisticated numerical algorithms, we have made a quantum leap forward in our ability to predict the behavior of our groundwater resources. This, coupled with the tremendous amount of research dollars, tied mainly to the environment, has led to a flood of research manuscripts, new engineering tools, and commercially available groundwater flow and transport software.

Unfortunately, while we have made tremendous advances in our ability to model the subsurface, our ability to cost-effectively measure material coefficients that go into the models has not kept pace. Thus, modern models are full of uncertainty in input data, and hence, their output is uncertain. This uncertainty has led to the recent introduction of stochastic tools to the field of subsurface hydrology.

This reference text is a compilation of many classical and more modern approaches to understanding and modeling subsurface flow and transport. Chapters 1 through 3 provide an introduction to the subject. In Chapter 1 the basic functions and terminology requisite to the study of flow and transport are laid out. This chapter also provides a historical background of the field. Subsequently, steady one- and two-dimensional flow and transport (Chapter 2) and deterministic two- and three-dimensional flow (Chapter 3) are discussed.

Flow and transport in geologic media are characterized by tremendous uncertainty in the geological environment. One can safely say that in no other field of engineering are we faced with more complexity and uncertainty than in groundwater flow and transport engineering. It is required that the groundwater engineer have at his or her disposal the tools (Chapter 4) which account for uncertainty in designs of dams, impoundments, etc. In addition, the engineer should have access to tools to characterize the uncertainty in heterogeneous aquifers (Chapters 10, 11), i.e., aquifers with material parameters which are assumed to vary randomly in space. It is also necessary to understand how to interpolate these data and their uncertainties (Chapter 12), and how to efficiently manipulate enormous amounts of spatial data (Chapter 21). Because of the randomness in geophysical data, stochastic analytic models often provide considerable insight into uncertainty propagation. In this regard, Chapter 14 deals with conser-

vative chemicals with special emphasis on uncertainty in dilution as contrasted with spreading. Chapter 15 introduces both lagrangian and eulerian stochastic-analytic models for chemicals undergoing equilibrium or nonequilibrium, and linear or nonlinear reactions.

Often it is observed that when a chemical is introduced in the subsurface, its travel time to some control plane follows a bimodal distribution, and in many cases the first arrival time of the chemical is far sooner than could be explained by theories involving homogeneous media. Such problems are discussed in Chapters 7, 17, and 18. Chapters 7 and 17 discuss flow in preferential paths of dual porosity type. Chapter 7 focuses on preferential paths induced by living organisms and shrinkage such as is encountered in natural soils, while Chapter 17 focuses on large-scale geologic fractures, such as those found at Yucca Mountain. Chapter 18 deals with solution channels such as those found in Karst media.

For water or a contaminant to reach an aquifer, it often must pass through an unsaturated medium. Chapters 5 and 6 deal with such flow problems, while Chapter 16 examines the transport of chemicals in nonrandom unsaturated media.

Groundwater wells have been around for thousands of years. However, only over the last 50 or so years have their siting and design (Chapters 8 and 9) significantly improved, and, unlike in earlier times, many wells today are constructed not for drinking or irrigation purposes, but rather for monitoring groundwater depletion and chemical migration (Chapter 24).

As environmental concerns have advanced, so too have the associated legal issues, and laws and regulations (Chapter 22) have also continuously evolved. Many of these have led to extensive litigation associated with landfills (Chapter 23) and various other contamination scenarios. This, in turn, has led to continued refinements in flow modeling (Chapter 20) and attempts to holistically engineer systems (Chapter 19). Litigation often results in one or more parties being liable for cleanup of groundwater (Chapter 25), which has led to significant research in biodegradation (Chapter 26) and geosynthetic material development (Chapter 27).

This text has been under development for approximately two years. The authors of the chapters are widely considered to be among the best in their subject areas. The chapters have been peer reviewed and revised accordingly.

<div align="right">

John H. Cushman
Center for Applied Mathematics
Purdue University
January 1998

</div>

1

Geological Occurrence of Groundwater

Darrell I. Leap
Purdue University

1.1   Introduction

Subsurface water is generally divided into two major types: phreatic water or soil moisture in the unsaturated zone, and groundwater in the saturated zone. This division is made mainly because of the

differences in the physics of flow of water in the saturated versus the unsaturated zone; these differences are discussed later in this chapter.

In order to understand the growth of importance of groundwater and the science explaining it in a modern, industrialized nation and economy, it is well to consider the experience of the United States. In the 18th century when the U.S. was being formed into a new nation, little emphasis was given to groundwater because in the humid East, where most of the population was located, surface water supplied nearly all water needs. In the 19th century and into the first half of the 20th, groundwater was investigated as a water source, especially as the population moved westward and new farms, ranches, mines, and cities were created in the semi-arid West. Lack of water caused intense competition among the various uses and users which continues to this day at an even greater pace.

Anticipating a need for natural resource evaluation of all kinds, the U.S. government in the waning years of the last century sponsored a number of exploratory and investigative studies of water resources, especially in the West. These studies continued until well into this century. Notable were the field efforts of T.C. Chamberlin (1885), N.H. Darton (1909), C.H. Lee (1915), W.C. Mendenhall (1905), and others. In addition, laboratory studies and theoretical studies added to our understanding of the physics and chemistry of groundwater. Important breakthroughs in physical understanding were made by several investigators who are listed with representative significant references: A. Hazen (1911), F.H. King (1899), C.V. Theis (1935), C.E. Jacob (1940), M.S. Hantush (1956), M.K. Hubbert (1940), and many others. The work of O.E. Meinzer (1923, 1939) of the U.S. Geological Survey is significant in that he was the prime motivator for many government-sponsored investigations.

The work of these scientists was built upon a solid foundation of theory developed in Europe during the 19th and early 20th centuries by such luminaries as H. Darcy (1856), J. Boussinesq (1904), J. DuPuit (1863), P. Forcheimer (1914), and G.Theim (1906).

In the latter half of this century, several refinements of physical groundwater concepts have been made, and there has been a considerable body of knowledge built up as the result of efforts to more safely dispose of hazardous and radioactive wastes. Numerous advances in remote sensing by satellites have helped in delineating water-bearing entities such as fracture zones, springs, and to a lesser extent, aquifers. Geophysical techniques and methods have improved significantly in their use as groundwater exploration tools. Pumping tests and other hydraulic tests have been developed which greatly improve our understanding of the hydraulics of aquifers and their potential for groundwater supplies.

Perhaps the most significant tool developed in this century for the study of groundwater has been the digital computer. As a consequence, numerical modeling of groundwater flow and contaminant transport has become a commonplace effort in nearly all groundwater studies.

Problems of data quality and data quantity in quantitative assessment of groundwater properties are mainly functions of the heterogeneity of most geologic formations and their hydraulic properties and the difficulty and expense of drilling for good data. With the recent advent of stochastic methods for analyzing hydrogeologic data, there are now better ways of interpolating, extrapolating, and predicting trends in hydraulic parameters.

With all the advances, both theoretical and technical, there is still one overriding fact in hydrogeologic analysis and modeling that must not be forgotten. Equations of flow and transport are differential in nature and require boundary conditions for their complete solution. The goodness of the solution to a partial differential equation of groundwater flow or transport is directly dependent upon the accuracy of the boundary conditions used. Boundary conditions are geological in nature and are only as accurate as the knowledge of the geology of the area being investigated.

This may appear simplistic to some, but the facts stand, not only from theory, but from the experience of thousands of investigators working for the past several decades on this problem. A thorough understanding of groundwater hydrology, correct interpretations of results of hydraulic tests, and correct interpretations of resulting model predictions can only be made after a thorough understanding of the geology is attained.

Still, high accuracy is often difficult to obtain in the quest for hydraulic parameter values of aquifers. There are variations in geologic structure which are often too small to be determined, but which do affect

hydraulic conductivity on a significant scale. Even in fairly homogeneous systems, this problem still frustrates the analyst and the modeler, and much more research is needed. Recent work by Eggleston et al. (1996) gives an example of this problem in the glacial outwash of Cape Cod.

Perhaps the most crucial area of groundwater research today is the search for new and improved methods to acquire, quantify, and properly utilize good subsurface hydrogeological data with minimal expenditure of time and money.

1.2 Importance of Groundwater

1.2.1 Groundwater and the World's Available Fresh Water

The amount of water on, under, and above the earth's surface remains essentially constant. Although a minor amount of water vapor may escape into space, additional new water is constantly created as *juvenile* water by chemical reactions during volcanic emanations. Table 1.1 shows estimates of the volumes of various kinds of water on, above, and under the surface of the earth as reported by Maidment (1993). Examination of the fig\ ^{3} in contrast to the 1,338,000,000 km^3 of sea water. Thus, fresh water makes up only 2.5% of all the water on earth, but not all of this water is available for human use. The water in polar ice caps, other forms of ice and snow, soil moisture, marshes, biological systems, and the atmosphere are not readily available. As a result, only the 10,530,000 km^3 of groundwater, 91,000 km^3 of fresh water in lakes, and the 2,120 km^3 of water in rivers are considered attainable for use and comprise a total of 10,623,120 km^3. Consequently, groundwater comprises 99% of the earth's available fresh water.

TABLE 1.1 Estimates of Relative Volumes of Water of Various Kinds on Earth

Item	Area 10^6 km^2	Volume km^3	Percent of total water	Percent of fresh water
Oceans	361.3	1,338,000,000	96.50	
Groundwater:				
Fresh	134.8	10,530,000	0.76	30.10
Saline	134.8	12,870,000	0.93	
Soil moisture	82.0	16,500	0.0012	0.05
Polar ice	16.0	24,023,500	1.7	68.6
Other ice and snow	0.3	340,600	0.025	1.0
Lakes:				
Fresh	1.2	91,000	0.007	0.26
Saline	0.8	85,400	0.006	
Marshes	2.7	11,470	0.0008	0.03
Rivers	148.8	2,120	0.0002	0.006
Biological water	510.0	1,120	0.0001	0.003
Atmospheric water	510.0	12,900	0.001	0.04
Total water	510.0	1,385,984,610	100	
Fresh water	148.8	35,029,210	2.5	100

Source: Adapted from Maidment, D.R. 1993. Hydrology, in Maidment, D. R., ed. *Handbook of Hydrology,* 1.1-1.15. McGraw-Hill, Inc., New York, NY. Data from UNESCO. 1992. *International Glossary of Hydrology.* 2nd Ed. WMO Rept. 385, World Meteorological Association. Geneva, Switzerland.

1.2.2 Groundwater Availability in North America

Heath (1984) describes the groundwater regions of the United States. Heath (1988) has made a thorough compilation of groundwater areas in all of North America, including the U.S., Canada, and Mexico. Figure 1.1 shows the major groundwater regions of the conterminous United States.

The United States, because of its size and varied and complex geology, contains groundwater in many different kinds of rocks and unconsolidated materials. Table 1.2 summarizes the hydraulic characteristics of each of the regions in Figure 1.1. A major aquifer system in the U.S. is worth a brief discussion because

FIGURE 1.1 Ground water regions in North America (Adapted from Heath, R.C. 1984. *Ground-Water Regions of the United States.* U.S. Geological Survey Water-Supply Paper 2242, 1-78. U.S. Government Printing Office, Washington, DC.).

TABLE 1.2 Hydraulic Characteristics of Groundwater Regions in the United States

Region No.	Name	Unconfined: Hydrologically insignificant	Unconfined: Minor aquifer or not very productive	Unconfined: Dominant aquifer	Confining beds: Hydrologically insignificant	Confining beds: Thin, discontinuous, or v. leaky	Confining beds: Interlayered with aquifers	Confined: Hydrologically insignificant	Confined: Not highly productive	Confined: Multiple productive aquifers	Confined: The dominant productive aquifer	Presence: Single unconfined aquifer	Presence: Two interconnected aquifers	Presence: Unconfined aquifer, confining bed, confined aquifer	Presence: complex interbedded sequence	Primary: Pores in unconsolidated dep.	Primary: Pores in semiconsolidated rocks	Primary: Tubes and cooling cracks in lava	Secondary: Fractures and faults	Secondary: Solution-enlarged openings	Solubility: Insoluble	Solubility: Mixed soluble and insoluble	Solubility: Soluble	Porosity: Large (>0.2)	Porosity: Moderate (0.01–0.2)	Porosity: Small (<0.01)	Transmissivity: Large (>2,500 $m^2 day^{-1}$)	Transmissivity: Moderate (250–2,500 $m^2 day^{-1}$)	Transmissivity: Small (25–250 $m^2 day^{-1}$)	Transmissivity: Very small (<25 $m^2 day^{-1}$)	Recharge: Uplands between streams	Recharge: Losing streams	Recharge: Leakage through confining beds	Discharge: Springs and surface seepage	Discharge: Evaporation and basin sinks	Discharge: Into other aquifers
1	Western Mountain Ranges		X		X				X				X			X			X		X				X				X		X	X		X		
2	Alluvial Basins			X		X			X						X	X					X			X				X				X			X	X
3	Columbia Lava Plateau		X			X			X						X	X		X	X		X			X			X					X		X		
4	Colorado Plateau and Wyoming Basin	X				X			X						X	X			X		X				X				X		X				X	X
5	High Plains			X	X			X				X				X					X			X				X				X		X		
6	Nonglaciated Central Region		X			X			X						X	X	X		X	X		X			X			X			X		X	X	X	X
7	Glaciated Central Region		X			X			X						X	X	X		X	X		X			X			X					X			X
8	Piedmont and Blue Ridge		X		X				X				X			X			X		X				X					X	X			X		
9	Northeast and Superior Uplands		X			X			X				X			X			X		X				X				X		X			X		
10	Atlantic and Gulf Coastal Plain		X			X				X					X	X				X	X			X				X			X		X	X		X
11	Southeast Coastal Plain		X			X								X		X				X			X	X				X			X		X	X		X
12	Alluvial Valleys			X		X		X				X				X					X			X				X			X			X		
13	Hawaii			X		X		X				X						X			X			X				X			X	X		X		
14	Alaska			X		X		X							X	X			X			X			X			X				X		X		
15	Puerto Rico and Virgin Islands			X		X			X						X	X				X			X	X				X			X			X	X	X

(Adapted from Heath, R.C. 1984. *Ground Water Regions of the United States.* U.S. Geological Survey Water-Supply Paper 2242, I-78, U.S. Government Printing Office, Washington, DC.)

of its importance in supplying water for major agricultural uses and because of problems associated with heavy pumpage from it.

The High Plains Aquifer occupies the area marked as "High Plains" in Figure 1.1. It is composed of several hundred feet of Tertiary sediments that were laid down by streams flowing eastward from the present site of the Rocky Mountains as they were being uplifted. Sandy layers in the present aquifer provide irrigation water for several states in the High Plains

Although there are a number of separate geological formations present, the Ogallala formation, consisting of alluvium, is the principal water-yielding unit and, thus, is the one most often tapped by wells. Gutentag et al. (1984) reported that in 1978 there were 170,000 wells extracting water from this aquifer, and the pumping rate is from 2 to 100 times the recharge rate.

As a result, the water table has dropped 100 feet or more, causing an increase in pumping expense. Careful management practices are now being put into place to better manage the resource. This is a significant example of ground-water mining and shows the effects of overuse of water resources.

Detailed descriptions of the hydrogeology of each region and in all of North America can be found in Heath (1988). In addition, the U.S. Geological Survey has performed detailed studies on several major aquifer systems in the U.S. through a major investigative effort known as the Regional Aquifer Systems Analysis (RASA). Specific reports of these studies are available from the U.S. Geological Survey National Headquarters in Reston, Virginia. Summaries of results of the RASA studies can be found in Bennet (1979), Sun (1986), and chapters 25, 26, and 27 of Back et al. (1988).

1.2.3 Brief History of Groundwater Usage

It is uncertain when mankind first started extracting groundwater by artificial means such as wells or infiltration galleries. Early humans most likely drank from surface streams. They may also have discovered groundwater through the discharge of natural springs in some parts of the world, and used this source in addition to surface streams. As streams dried up in hot weather, people learned to dig into the alluvium to find water below the surface.

Much speculation arose about the origin of water emanating from springs, and early theories proposed a hollow earth filled with water. The ancient Greeks were the first to record their theories, germinated no doubt from their observation of disappearing rivers and emerging springs in the limestone *karst* areas of Greece and the rest of the Balkan Peninsula. Davis and DeWiest (1966) summarize the history of groundwater use and the development of the body of knowledge now surrounding it.

The earliest recorded use of groundwater is found in the 16th chapter of the *Book of Genesis*, in which it tells of a spring in the desert. In chapter 26, mention is made of the digging of wells. Egyptians are reported to have perfected core drilling by 3000 BC. Ancient Chinese invented the churn drill (similar to the modern cable-tool drill) and drilled to a depth of 1500 meters (Tolman, 1937). Western Europe did not engage in drilling until the 12th century when churn drilling was developed there. Much of the impetus for development of drilling methods was generated in the area around Artois, France, where artesian water was in demand. In the latter part of the 19th century, the rotary drilling technique was developed, and by 1890 the *hydraulic rotary method* had been perfected. This was probably the most significant single advance in drilling technology (Davis and DeWiest, 1966).

As the water-well and oil-well drilling businesses grew in various parts of the world, rotary drilling methods were continuously improved and introduced worldwide. Today, it is the most widely used method and has been the primary reason that groundwater is now available to even primitive tribes. Present groundwater availability to people on every continent has been a prime reason for increasing the health and standard of living for untold millions. Groundwater is now recognized as a national treasure and a most important natural resource. For example, approximately half the water used in the U.S. is groundwater. In some areas, groundwater is the sole water supply for all purposes.

1.3   Earth Materials

1.3.1   Minerals and Rocks

The inorganic part of the earth is made up of rocks and their weathering products. A rock, by definition, contains one or more *minerals*, and a mineral is defined by Hurlbut (1970) as follows:

"In addition to being natural and inorganic, a mineral must meet another requirement: it must be a chemical element or compound. It cannot be a random mixture of elements; the atoms that make it up must have definite ratios to each other, so that its composition can be expressed by a chemical formula. Not only are the proportions of the various atoms of a given mineral fixed, but so are their relative positions. These attributes give to each mineral a set of properties that characterize it so uniquely that one can distinguish it from all other minerals".

A *rock* can be composed of one mineral or a mixture of several. For example, sandstone may contain grains of the mineral quartz (silica or SiO_2), and a cement between the quartz grains composed of the mineral calcite ($CaCO_3$). Granite is normally composed of crystals of the minerals feldspar, quartz, mica, and others. By definition then, frozen water (ice) is both a mineral and a mono-mineralic rock.

The most fundamental of the three classes of rocks are *igneous rocks,* which are formed as cooling products from the molten state. Igneous rocks are primordial in many parts of the world and are some of the world's oldest, exceeding three billion years in age. One way they can be formed is when molten rock is intruded into other rock formations and then cooled to the solid state. If the cooling is slow enough, various minerals will crystallize into an interlocking solid mass that is characteristic of particular rocks such as granite.

Extrusive igneous rocks like dense basalt are forced from fissures in the earth's crust and harden into vast sheets of solid material, usually containing very small crystals (due to the rapid cooling), or perhaps no crystals at all (obsidian glass). Other extrusives like lighter lava or pumice are ejected during volcanic eruptions and are highly charged with gases to form very porous and even frothy glasses resembling a sponge. Some of the ejecta may fall from the air to settle as a sediment. This particular kind of deposit is known as a *pyroclastic* rock, i.e., both igneous and sedimentary.

A *sedimentary rock*, the second class, is deposited from either air or water as grains of rocks and minerals. These sediments in turn may have been derived from the weathering of igneous, metamorphic, or other sedimentary rocks.

If any kind of rock, igneous, metamorphic, or sedimentary, is subjected to intense heat and pressure, such as exists at great depths in the earth's crust, at the edge of tectonic plates, or in rising mountain ranges, the parent rock will be transformed into the third class of rock — a *metamorphic rock*. A metamorphic rock may contain the same chemical composition as the parent rock, but the mineral composition and structure may be changed drastically from the parent. For example, limestone containing amorphous or cryptocrystalline calcite ($CaCO_3$) is often metamorphosed into marble that has a definite crystal structure and is much harder than the original limestone. Granite may be metamorphosed into gneiss with the same overall chemical composition as the original granite, but with new minerals and mineral structures.

Groundwater can be found in all three classes of rocks, but in general, the sedimentary rocks contain by far the greatest amounts of water due to their greater *porosity*.

1.3.2   Unconsolidated Materials

Unconsolidated materials are those earth materials which have not been indurated. That is, the grains have not been fused together by heat and pressure, as in the cases of igneous granite or metamorphic gneiss; or cemented together, as in the case of sedimentary rocks. Unconsolidated materials can be the non-indurated products of weathering of all three classes of rocks, or sediments laid down by running water, ponded water, the sea, or ejecta from volcanoes.

Most unconsolidated materials are young, geologically speaking, and are at or near the earth's surface. Thus, they have not been exposed to pressure, heat, and migrating cementing fluids long enough to become consolidated or hard. Hence, they generally have high porosity and are the sources for much groundwater.

A common unconsolidated deposit in glaciated areas of the world is *glacial drift*. It is any kind of earth material that was deposited directly by glaciers or by meltwater from glaciers. As such, it can range in size from the finest silt to the largest boulders, and can be mixtures of all sizes. The name "drift" was given to this material when it was believed that it was depositional material, or "drift," from the great Noachian Flood, described in the Book of Genesis. Drift can generally be subdivided into the more specific lithologies such as *till* and *outwash*.

Glacial till is a generally heterogeneous mixture of many different lithologies and particle sizes. Typically in the midwestern U.S., till contains a preponderence of clay and silt with additional amounts of ground-up rocks and boulders that may vary in size from small pebbles to erratics the size of a house or larger. On rare occasions, geologists find tills composed of one lithology, indicating local sources of material. Glacial till is not generally utilized as a source of groundwater because of its low *permeability*

Glacial outwash is material deposited from high-energy streams of water that originated from melting glaciers. This process can be seen today at the toe of any mountain glacier on different continents.

Glacial ice, more often than not, contains entrained rock and soil material which it has eroded from the surrounding valley sides or the ground beneath it. This material varies widely in size from clay to large boulders. As the ice melts and leaves the toe of the glacier, normally in great flow rates and very turbulent, this material is moved with the flowing water. As the stream loses energy, materials settle out, with larger material coming out first, followed by gradually smaller material. Therefore, along a stretch of an outwash stream one may find coarse gravel and boulders settling out first, followed by finer gravel, then sand and gravel mixtures, then sand, silt, and finally clay (in still water).

Many present-day stream courses were glacial spillways for outwash water and sediment during the Pleistocene Epoch (Ice Age). For example, the Wabash River Valley in Indiana contains outwash sand and gravel deposits in excess of 300 feet in thickness which were deposited by the melting of two and possibly three different ice sheets. The Big Sioux River valley in South Dakota and Iowa is another example. A very extensive deposit of outwash sand is the Cape Cod peninsula which was deposited in an interlobate outwash between two lobes of ice — one to the east and one to the west of the site. Outwash sand and gravel deposits are frequently exploited for groundwater because of their high porosity and permeability. Well yields in many of these deposits often exceed 5500 m^3/d.

Other water-lain deposits not directly deposited from glacial meltwater may be comprised of reworked glacial detritus (glacially transported material), or they may be found in areas where glaciers never occurred. The most common example of such deposits is *alluvium*. Flood plains along large streams are created of this material as the streams flood over their banks and deposit the material. Stream beds also contain alluvium.

In fast-flowing streams with high gradients, as in mountainous areas, alluvium may be absent because the stream is eroding rather than depositing material. If alluvium is found in and along such streams, it is generally very coarse-grained gravel with large boulders. On the other hand, mature streams such as the Ohio, Missouri, and Mississippi Rivers, deposit their loads of fine silt and clay over broad flood plains.

Alluvium deposits may serve as important groundwater sources, but in large river valleys, the yield of such deposits may be low to moderate, depending on the grain size and the resulting permeability and porosity.

Lacustrine materials are silts and clays that are deposited from relatively still bodies of water such as lakes and lagoons. This material, being so fine-grained, is not utilized for groundwater supplies because of its low permeability.

Peat is the remains mostly of water plants that die and accumulate in ponded water and marshes over long periods of time. The top part of a peat deposit is very porous and permeable, but it becomes more compact with depth. The lower layers of peat are often sticky masses of black organic material with little resemblance to the original plant material.

Peat is not generally utilized as a groundwater source, but it can serve, under the right conditions, as a natural cleansing agent to remove organics and heavy metals from water that passes through it — the large and complex organic molecules in the peat attract such contaminants.

Chemical precipitates of most importance include limestone and marl which precipitate directly from sea water or even from fresh water bodies. Major deposits of limestone were deposited in many parts of the world during the Cretaceous Period (the "Chalk Period") of the Mesozoic Era. Examples include the Chalk Beds of Dover, England; the limestones of the Balkan Peninsula; and limestones of the High Plains in the United States. Other vast limestone deposits were formed in earlier times and are found across most of the midwestern U.S. and the Appalachians.

Limestone is composed primarily of calcite. Entrained silt and clay and other materials may also be present. After deposition, a process known as *diagenesis* often takes place in which the rock incorporates magnesium to become dolostone, or $CaMg(CO_3)_2$; this is also the formula for the mineral, dolomite. Pure $MgCO_3$ is the mineral, magnesite.

Collectively, limestone and dolostone are called "carbonate rocks," or "carbonates." Carbonates, especially limestone, often undergo solution along bedding planes and fractures to form caves and sinkholes, which in an interconnected system, is known as *karst* terrane, after the Karst region in Yugoslavia. Networks of such caves and tunnels may exceed hundreds of miles in length and may contain large streams which emanate from springs in the rock.

Examples of such systems are found in the Balkan Peninsula, the Mediterranean area, France, Kentucky (Mammoth Caves), and New Mexico (Carlsbad Caverns), to name just a few. Karst systems are often sources of very large quantities of groundwater, and due to the very high permeability, can be productive aquifers. A drawback, though, is the ease with which water in karst systems can be contaminated by surface sources. Thus, care must be taken to protect such sources.

Aeolian deposits are fine-grained materials, such as silt and sand, which may have been deposited originally from water, but which have been reworked and redeposited by wind. Examples of such active deposits today can be found in sand-dune areas of the Sahara, the Middle East, New Mexico, Nevada, and many other places.

Ancient dunes from the geological past are often found as sandstone bodies and may have some potential for groundwater extraction if coarse enough to allow sufficient porosity and permeability. Generally, wind-blown deposits are fine-grained, and when cemented with precipitates from circulating groundwater, may possess low porosity and permeability. The finest-grained aeolian deposits are composed of silt or "rock flour" known as *loess*. The silt source is usually a wide river bed with braided channels where large dry areas of fine-grained materials are exposed to the wind. In most cases, the silt was deposited in such rivers as the end product in the long chain of deposition of glacial outwash. Prevailing winds then pick up the silt and transport it downwind where it is deposited on the lee sides of river valleys.

Significant loess deposits are found on the east side of the Missouri River, which acted as a glacial spillway during the Pleistocene. North of Sioux City, Iowa, this material forms bluffs which are tens of meters high. Other noteworthy deposits are found along the Mississippi River (another spillway) and in the Gobi Desert of China.

Loess, after deposition, will often be reworked by frost action to form columnar structures with vertical fractures. The grains of silt are then oriented with their long axes vertically to form such features. With this alteration, it will allow fast vertical movement and drainage of water and is fairly solid material to build upon, but its permeability is too low to utilize it for a groundwater supply.

1.4 Important Hydraulic Properties of Earth Materials and Groundwater

1.4.1 Porosity

Porosity (n) is defined as the volume of the pores of a rock or soil sample (V_p) divided by the total volume (V_t) of both pores and solid material. That is,

$$n = V_p \big/ V_t \qquad (1)$$

When a rock is first formed by precipitation, cooling from an igneous melt, induration from loose sediments, or when a soil is first formed by weathering of rock materials and possibly subsequent biological action, the new entity will contain a certain inherent porosity known as *primary porosity*. This porosity may later be reduced by cementation from precipitates from circulating groundwater, or from compaction accompanying burial by later sediments. However, fractures or solution cavities formed in the rock, or root tubes or animal burrows in soils may later form and are known as *secondary porosity*. Thus, the total porosity of a sample will be the sum of the primary and secondary porosities.

Porosity of a consolidated sample can be determined quite simply by first cutting the sample to a known dimension such as a cylinder or cube and measuring the total volume. Next, the sample is submerged in a known volume of water and allowed to saturate. After saturation, the volume of water displaced will be the volume of solids in the sample. The volume of voids is simply the difference between the total volume and the volume of solids, and porosity can be calculated by the above formula. More accurate and more sophisticated methods for testing earth materials for various properties are given in various publications of the American Society for Testing and Materials (ASTM).

If all the pores in a rock are not connected, only a certain fraction of the pores will allow the passage of water, and this fraction is known as the *effective porosity*. An example is pumice, a glassy volcanic ejecta (a solidified froth) which may float in water because the total porosity is so high and because it contains so much entrained gas that its bulk density is less than water. It may take some days to sink because the effective porosity is so low that water cannot easily pass through it. Coarse gravel may have a porosity of only 25%, but that porosity will practically all be effective. Thus, gravel is an excellent conductor of water.

Porosity of a rock or soil is determined largely by the packing arrangement of its grains and the uniformity of the grain-size distribution. The greatest ideal porosity that could be attained in a material with uniform spherical grains is 47.65%; this is known as *cubic packing* (Figure 1.2A) because the centers of eight such grains form the vertices of a cube. The least porosity that can be attained with the same grains is found in *rhombohedral packing* with a value of 25.95% (Figure 1.2B). The centers of the eight adjacent spheres form the vertices of a rhombus.

These are ideal extremes because no natural sediment or rock contains spherical grains nor are the grains of uniform size. If small grains are situated in the spaces between large grains, the porosity will be reduced below that for the large grains only. This is demonstrated by the *uniformity coefficient* (C_u) which is a measure of how well or how poorly the sample grains are sorted,

$$C_u = D_{60} \big/ D_{10} \qquad (2)$$

where D_{60} is the grain diameter below which 60% of the grains are finer, and D_{10} (the effective grain size) is the diameter below which 10% of the grains are finer. If the C_u is less than 4, the sample is well sorted; if it is greater than 6, it is poorly sorted and will have a lower porosity than the first sample.

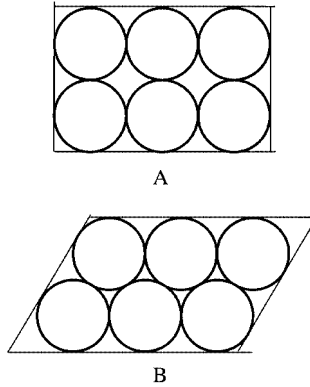The void ratio, e, used in soil mechanics, is defined as

FIGURE 1.2 (A) Cubic packing; (B) Rhombohedral packing.

$$e = V_p / V_s \qquad (3)$$

where V_s is the volume of solids, so that

$$n = 1 / \left(1 + 1/e\right) \qquad (4)$$

Typical values of porosity are given in Table 1.3.

TABLE 1.3   Important Physical Properties of Soil and Rock

Lithology	Porosity (percent)	Hydraulic conductivity (cm/sec)	Compressibility, α (m²/N or Pa⁻¹)
Unconsolidated			
Gravel	25–40	$10^{-2}–10^{2}$	$10^{-8}–10^{-10}$
Sand	25–50	$10^{-4}–1$	$10^{-7}–10^{-9}$
Silt	35–50	$10^{-7}–10^{-3}$	no data
Clay	40–70	$10^{-10}–10^{-7}$	$10^{-6}–10^{-8}$
Glacial Till	10–20	$10^{-10}–10^{-4}$	$10^{-6}–10^{-8}$
Indurated			
Fractured Basalt	5–50	$10^{-5}–1$	$10^{-8}–10^{-10}$
Karst Limestone	5–50	$10^{-4}–10$	not applicable
Sandstone	5–30	$10^{-8}–10^{-4}$	$10^{-11}–10^{-10}$
Limestone, Dolomite	0–20	$10^{-7}–10^{-4}$	$< 10^{-10}$
Shale	0–10	$10^{-11}–10^{-7}$	$10^{-7}–10^{-8}$
Fractured Crystalline Rock	0–10	$10^{-7}–10^{-2}$	$—10^{-10}—$
Dense Crystalline Rock	0–5	$10^{-12}–10^{-8}$	$10^{-9}–10^{-11}$

(Adapted from Domenico, P.A. and Schwartz, F.W. 1990. *Physical and Chemical Hydrogeology*. John Wiley & Sons, Inc., New York; Freeze, R.A. and Cherry, J.A. 1979. *Groundwater*. Prentice-Hall, Inc., Englewood Cliffs, NJ.; Fetter, C.W. 1994. *Applied Hydrogeology*, 3rd ed. Macmillan College Publishing Co. Inc., New York; Narasimhan, T.N., and Goyal, K.P. 1984. Subsidence due to geothermal fluid withdrawal, in *Man-Induced Land Subsidence,* Reviews in Engineering Geology, v. VI, Geological Society of America, 35-66.)

1.4.2 Moisture Content

Moisture content (Θ) can be measured and described either *gravimetrically* or *volumetrically*. The gravimetric equation is

$$\Phi = W_w / W_t \tag{5}$$

where W_w is the weight of the water contained in a sample and W_t is the total weight of the solids and water in the sample. This definition, although useful for some purposes, does not indicate the degree of saturation of the rock or soil. Thus, the volumetric definition is more widely used,

$$\Phi = V_w / V_t \tag{6}$$

where V_w = volume of water in the sample, and V_t = the total volume of water and solids. Closely related terms are the *saturation ratio,* which is the volume of contained water divided by the volume of voids, and the *degree of saturation,* or *saturation index*, which is the saturation ratio times 100, $(V_w/V_v) \times 100$. If the saturation ratio is less than unity, or the degree of saturation is less than 100%, the sample is unsaturated and the pores are then partially filled with air.

Gravimetric moisture content can be determined by weighing a sample to obtain the total weight and then drying it in an oven to drive out the moisture. The dry sample is weighed to obtain the weight of the solids. The difference is the weight of the entrained water.

1.4.3   Capillarity

Capillary forces play a major role in the movement of water through unsaturated materials. Water is attracted to solid grains by *adhesion.* The familiar example of water rising in a soda straw is also a good example for capillary rise, h_c, in a small tube of radius, r, (Figure 1.3) and the rise is calculated as

$$h_c = 2\sigma \cos\alpha / \gamma_w r \tag{7}$$

where σ is the surface tension of water (0.0756 N/m at 0°C), γ_w is the specific weight of water (9.805 kN/m^3 at 0°C), $\alpha = 0°$ for water. Thus, for water at 0°C in a clean tube, Domenico and Schwartz (1990) give the formula for capillary rise in cm as,

$$h_c = 0.153 / r \tag{8}$$

Mavis and Tsui (1939) developed the following equation to estimate the height of capillary rise in mm in soils,

$$h_c = \left(2.2 / d_H\right)\left[\left(1-n\right)/n\right]^{2/3} \tag{9}$$

where d_H is the harmonic mean grain diameter in millimeters and n is the porosity. Capillary rise in coarse gravels may be only a few millimeters, but in clay it may be as much as three or four meters (Davis and DeWiest, 1966; Lohman, 1972).

In the root zone, normally in the upper part of an unsaturated soil, there is a natural competition for water between capillary forces in the soil and osmotic suction in the plant roots. The moisture content at which the capillary forces become greater than osmotic forces is known as the *wilting point*, i.e., the moisture content at which the plants will cease to take in water and will start to wilt. Due to the fact that capillary forces in coarse-grained materials are not as great as those in fine-grained materials, the wilting point in sand is only about 5%, whereas in clay it is around 25 to 30% (USDA, 1955). The wilting point will be reached more rapidly during hot, dry weather, especially if rainfall is insufficient, because capillarity causes water in the soil to rise where evapotranspiration from the surface will accelerate the loss of water from the soil.
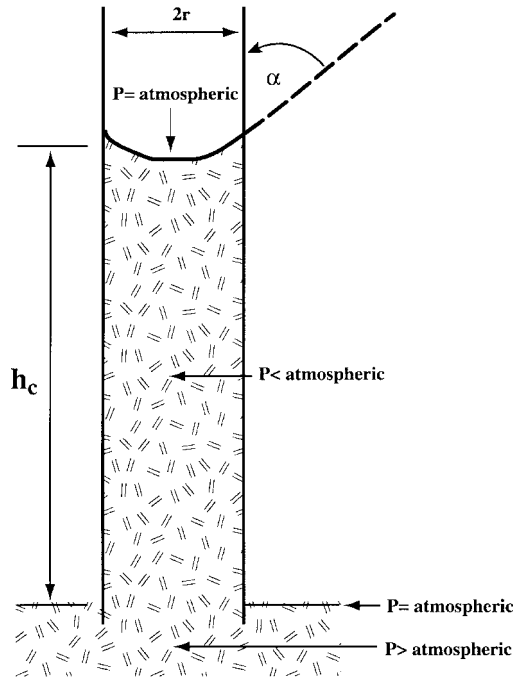
FIGURE 1.3 Capillary rise in a tube (Adapted from Lohman, S.W. 1972. *Ground-Water Hydraulics.* U.S. Geological Survey Prof. Paper 708, US Government Printing Office, Washington, DC.).

1.4.4 Electrostatic Forces of Attraction and States of Water in Pores

Capillarity is caused by a combination of two forces: (1) *molecular attraction,* which is responsible for water adhering to soil or rock particle surfaces, and (2) *surface tension,* which is due to cohesion of water molecules toward each other when water is exposed to air. In saturated systems, the first force is balanced in all directions and thus, canceled out, and surface tension is also balanced.

An unsaturated rock or soil, if allowed to drain under the force of gravity, will not lose all its water by this means (Figure 1.4). Instead, some water will be held inside the pores by these forces. The drained water is known as *gravitational water,* and the retained water as *capillary water* (Todd, 1980). Older literature may refer to the latter as *pellicular water.* The capillary or pellicular water will stay in the pores unless it is subjected to additional stresses such as centrifugation in the laboratory or excessive heat as during a hot, dry period.

If moisture reduction continues, the minimal water that will be held will exist as thin films around grains and will be held there by adhesive forces of molecular attraction; this water is known as *hygroscopic water* and is unavailable to plants. The wilting point is always a higher moisture content than the hygroscopic water content.

Often, initially dry soils near the surface will first attract and hold hygroscopic water if the air humidity increases significantly. It is only after the hygroscopic moisture content exceeds the *hygroscopic coefficient* (the maximum water content that can be held hygroscopically) that any additional water from precipitation or irrigation travel through the soil. Adhesion and cohesion are strongly affected by the mineralogy of soil or rock and chemical content of soil and water.

1.4.5 Compressibility of Water

Water is only very slightly compressible. At conditions of constant temperature and mass, Domenico and Schwartz (1990) define the *isothermal compressibility* of water as
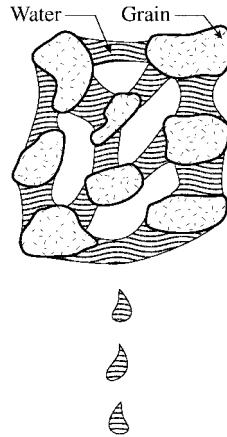
FIGURE 1.4 Porous medium under drainage.

$$\beta_w = 1/K_w = -\left(1/V_w\right)\left(\partial V_w/\partial P\right) \tag{10}$$

where β_w is the fluid compressibility in units of reciprocal pressure, K_w is the bulk modulus of compression for water, V_w is the bulk volume of water, and P is the pressure. At 25°C, groundwater possesses a β_w of 4.8×10^{-10} m²/N or 2.3×10^{-8} ft²/lb.

1.4.6 Compressibility of Solid Earth Materials

The compressibility of water-bearing rock and soil at some internal point is affected by both external and internal stresses and pressure of entrained water within the pores. The stress-balance equation is given as

$$\sigma_t = \sigma_e + P_p \tag{11}$$

where σ_t is the total vertical stress acting downward on the point of interest and includes the pressure of overlying soil or rock and its contained water as well as that from buildings, trees, etc., on the surface. The effective stress or resisting stress from the skeleton of the solid grains, i.e., the matrix, is σ_e, and P_p is the pore pressure of the water in the pores. Figure 1.5 illustrates this relationship.

Any increase in total vertical stress must be balanced by the same increase on the right side of the equation; this occurs when depth increases because total vertical stress naturally increases with depth. If a well is pumped, the extraction of water will suddenly decrease the pore pressure and the porosity with an attendant increase in the effective stress exerted by the matrix. As a result, the matrix will compress. If water is injected into the well, the reverse occurs.

Such expansion and compression of the matrix, at constant temperature and assuming incompressible grains, can be quantified as shown by Domenico and Schwartz (1990) as

$$\beta_b = 1/K_b \tag{12}$$

$$= -\left(1/V_b\right)\left(\partial V_b/\partial \sigma_t\right) \tag{13}$$

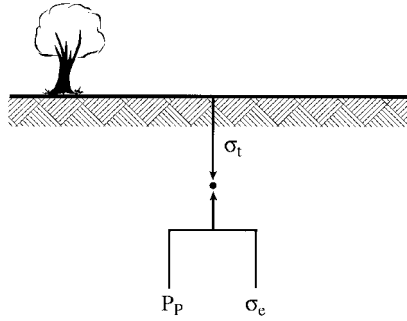$$= -1/V_b\left(\partial V_p/\partial \sigma_t\right) \tag{14}$$

FIGURE 1.5 Stress-balance diagram.

$$= \beta_p = 1/H_p \tag{15}$$

where β_b is the bulk matrix compressibility in units of the reciprocal of pressure, K_b is the rock bulk modulus of compression, β_p is the vertical compressibility, H_p is a modulus of vertical compression related only to pores, V_b is the bulk volume, and V_p is the pore volume. The negative sign in the equation refers to the fact that volume decreases with increasing pressure. Equations 9 through 12 show that if the grains are incompressible then β_b is equivalent to β_p.

Table 1.3 lists the compressibility of several lithologies.

1.4.7 Hydraulic Head

Hydraulic head at a point in a groundwater system (Figure 1.6) is expressed as,

$$h = Z + P/\rho g \tag{16}$$

where Z is the elevation head or the distance of the reference point above a datum plane (normally mean sea level), P is fluid pressure at the point exerted by the column of water above the point, and ρg is the specific weight of water, γ, or more simply stated

$$h = Z + \Psi \tag{17}$$

where Ψ is also called the pressure head. The relationship has the dimension of length. Development of this relationship can be found in Chapter 2 of this publication and in Hubbert (1940).

Thus, hydraulic head has the dimension of length which makes it convenient for calculations based on the elevation of water above sea level, the generally accepted zero datum. Fluid potential and hydraulic head are equivalent and both are used, except head is the most popular entity for groundwater studies. It can be shown, often to the surprise of the unsuspecting, that water can flow from a region of lower pressure to a region of higher pressure if the total head at the starting point is greater than at the ending point.

In the field of petroleum and natural gas engineering, pressure is generally used in place of head because pressures at great depths are normally so great that elevation heads are often insignificant.

1.4.8 Storage

A simple mass balance equation for groundwater flow through a unit volume of porous medium is given as

$$Q_{\text{out}} \, \Delta t = Q_{\text{in}} \, \Delta t \pm \Delta S \, \Delta t \tag{18}$$
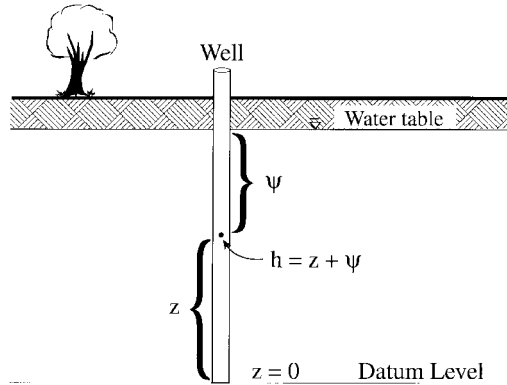
FIGURE 1.6 Illustration of hydraulic head.

where Q is total flow rate in volume per unit time, and ΔS is the volume per unit of time going into or coming out of storage in time, t.

In a saturated porous medium that is confined between two transmissive layers of rock or clay, water will be stored in the pores of the medium by a combination of two phenomena; these are water compression and aquifer expansion. As water is forced into the system at a rate greater than it is being extracted, the water will compress and the matrix will expand to accommodate the excess. In a unit volume of saturated porous matrix, the volume of water that will be taken into storage under a unit increase in head, or the volume that will be released under a unit decrease in head is called *specific storage*, and is shown as

$$S_S = \rho g \left(\alpha + n\beta \right) \tag{19}$$

where α is aquifer compressibility, ρ is fluid density, g is gravitational acceleration, n is porosity, and β is water compressibility. This unit has the dimension of 1/L and is quite small, usually 0.0001 or less.

The storage coefficient of an aquifer, or simply, the *storativity*, S, is given as

$$S = S_S b \tag{20}$$

where b is the saturated thickness of the aquifer. Storativity is defined as the volume of water per unit aquifer surface area taken into or released from storage per unit increase or decrease in head, respectively (Figure 1.7). It is a dimensionless quantity. In confined aquifers the value of storativity ranges from 0.005 to 0.00005.

In unconfined porous media, that is, where there is no overlying confining cover, storage of water in its upper part is defined as specific yield, S_y. This is the ratio of the volume of water that drains from a saturated porous matrix under the influence of gravity to the total volume of the matrix, per unit drop in the water table. Specific yield is normally much greater than specific storage, as water released from elastic storage leaves the pores still saturated. Specific yield is often in the range of 0.2 to 0.3, or three to four orders of magnitude greater than elastic storage.

Gravity drainage will proceed until the forces of surface tension and molecular attraction to the matrix grains become equal to the force of gravity. The ratio of the volume of water retained in the pores to the total matrix volume is known as *specific retention*, or S_r. In the upper parts of an unconfined porous medium, where elastic storage is not significant, the sum of specific yield and specific retention equals porosity,
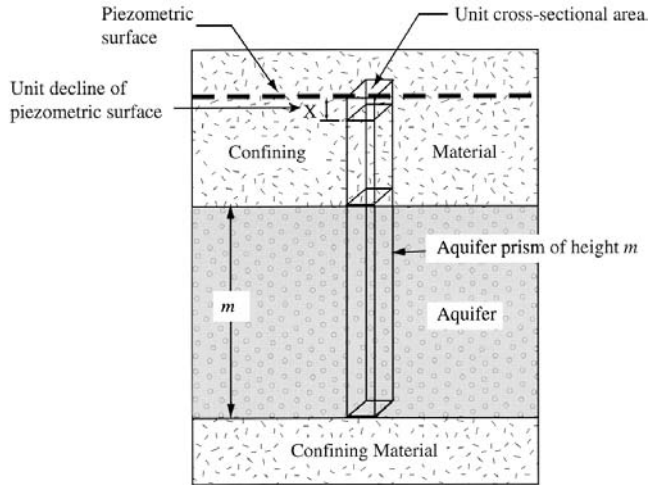
FIGURE 1.7 Illustration of specific storage and storativity (Adapted from Ferris, J. G., Knowles, D.B., Browne, R.H., and Stallman, R.W. 1962. *Theory of Aquifer Test.* U.S. Geological Survey. Water-Supply Paper 1536E.)

$$S_y + S_r = n \tag{21}$$

When an unconfined porous medium is very thick, the lower parts of the medium may also contain water under elastic storage, owing to the increase of pressure and consequent water compressibility and matrix expansion with increasing depth. In this case the total storativity of the medium is expressed as

$$S = S_y + b S_S \tag{22}$$

where b is the saturated medium thickness.

Field capacity is used to describe essentially the same phenomena as specific retention, but it is normally used in agricultural soil-moisture studies. It is a function not only of specific retention, but also of the evaporation depth and the unsaturated permeability of the soil (discussed below). A good discussion of all the above relationships can be found in De Marsily (1986).

1.4.9 Intrinsic Permeability

There are several basic properties of both porous media and fluid that will determine the ease with which the medium will transmit a fluid. The most fundamental of all these properties is known as *intrinsic permeability*, k. It is simply a function of the average pore size of the medium and is related to this property as follows,

$$k = C d^2 \tag{23}$$

where d is the average pore diameter, and C is an empirical constant which depends upon packing, sorting, and other factors.

Intrinsic permeability is strictly a function of the medium and has nothing to do with the temperature, pressure, or fluid properties of a particular fluid passing through the medium. It is commonly measured in terms of the *darcy*, after Henri Darcy who developed the relationship known today as Darcy's Law, (discussed later). The millidarcy is the commonly used unit, but many authors also use the units of cm^2.

Although independent of the fluid, the darcy is defined with a fluid standard in order to quantify it as the area through which a fluid with a dynamic viscosity of one centipoise will flow at a rate equivalent

to one cubic centimeter per second per square centimeter under a pressure gradient of one atmosphere per centimeter; or

$$1 \text{ darcy} = \left[\left(1 \text{ centipoise} \times 1 \text{ cm}^3/\text{sec}\right)/1 \text{ cm}^2\right]/(1 \text{ atmosphere}/1 \text{ cm}) \tag{24}$$

One centipoise is 0.01 dyne-sec/cm^2, and one atmosphere is 1.0132 x 10^6 dynes/cm^2. If these values are substituted in to the above equation, the result is that

$$1 \text{ darcy} = 9.87 \times 10^{-9} \text{ cm}^2 \tag{25}$$

and the same value will apply to the flow of gases or newtonian liquids of any kind through the medium.

1.4.10 Viscosity

The resistance of a flowing fluid to shear is known as *dynamic viscosity*. Figure 1.8 shows the conceptual and mathematical relationships in the development of the term. Note that a liquid fills the space between two plates; the bottom plate is stationary and the top plate moves unidirectionally at a velocity, U. The liquid is undergoing laminar flow and thus there is a linear change of velocity upward from zero on the bottom plate to U on the top plate. The rate of change of velocity with vertical direction is then dU/dy which is also called the vertical rate of strain. This is caused by the shearing stress, τ, which is

$$\tau = F/A \tag{26}$$

where *F* is the force that pulls the upper plate along at the velocity, *U*, and *A* is the surface area of the upper plate.
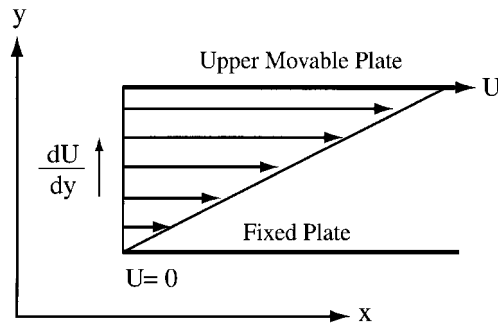

FIGURE 1.8 Illustration of dynamic viscosity as resistance of a fluid to shear.

From these observations it is apparent that

$$\tau \propto dU/dy \tag{27}$$

and we assume that there is a constant, μ, that will convert this proportionality into an equation, namely

$$\tau = \mu\left(dU/dy\right) \tag{28}$$

and

$$\mu = \tau \big/ \left(dU \big/ dy \right) \tag{29}$$

where the constant, is known as the dynamic viscosity. It is specific to the fluid and to temperature, generally decreasing with increasing temperature. The dynamic viscosity has the units of [FT/L^2].

Dynamic viscosity is defined in terms of the *poise* (0.1 Ns/m^2), after the French mathematician and fluid dynamicist, Poiseuille. The centipoise (0.001 Ns/m^2) is the most commonly used unit. At 15°C, water possesses a dynamic viscosity of 0.011404 poise, or 1.1404×10^{-3} Ns/m^2, or approximately one centipoise. In contrast, the dynamic viscosity of the earth's mantle is estimated at 10^{23} centipoise.

Dynamic viscosity is not to be confused with *kinematic viscosity*, ν, which is dynamic viscosity divided by fluid density,

$$\upsilon = \mu \big/ \rho \tag{30}$$

The kinematic viscosity has the units of L^2/T. The stoke, named after G.G. Stoke, the British physicist, has the units of cm^2/s, but the centistoke (0.01 cm^2/s) is a more convenient unit. Water at 15°C has a kinematic viscosity of 1.139×10^{-6} m^2/s or 1.139 centistokes.

1.4.11 Hydraulic Gradient

The hydraulic gradient is simply the loss of head per unit length of flow along a stream line, and is given as

$$I = dh \big/ dl \tag{31}$$

a dimensionless number. Development of this relationship can be found in Chapter 2 of this publication and in Hubbert (1940).

1.4.12 Hydraulic Conductivity of Saturated Media

Although the intrinsic permeability describes the water-transmitting property of a porous medium, it does not completely describe the relative ease with which a particular liquid will flow through the medium. The complete description is given by the *hydraulic conductivity, K,* which combines both medium and fluid properties,

$$K = k\rho g \big/ \mu \tag{32}$$

This parameter has the dimension of velocity, generally cm/sec or ft/day, and is a second-order tensor quantity. Its physical meaning is stated as, "The volume of liquid flowing perpendicular to a unit area of porous medium per unit time under the influence of a hydraulic gradient of unity." Earlier literature described this phenomenon as *field coefficient of permeability* with units of gallons per day per square foot. This name and definition are now only rarely used. Table 1.3 lists the hydraulic conductivities of several lithologies.

In the mid-19th century, a French engineer, Henri Darcy, was experimenting with sand filters for a water supply system for the city of Dijon, France. Through a series of experiments with sand-filled tubes, (Figure 1.9), he was able to determine the factors that controlled the flow rate of water through the sand (Darcy, 1856).

He discovered one of the most important physical relationships in the science of porous-media hydrodynamics, which became known as *Darcy's Law,*

$$q = -K \left(dh \big/ dl \right) \tag{33}$$
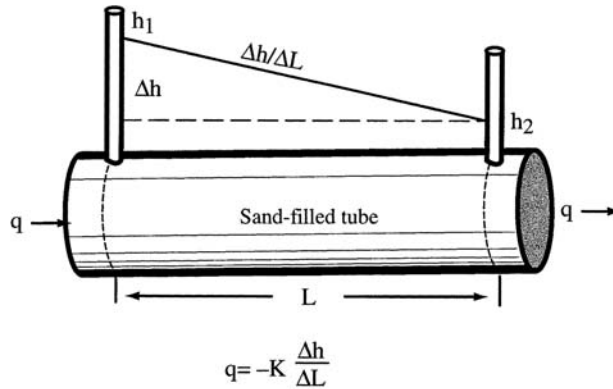
$$q = -K \frac{\Delta h}{\Delta L}$$

FIGURE 1.9 Figure illustrating tube experiments of Henri Darcy.

where q is the *specific discharge* or volumetric flow rate per unit area of porous medium perpendicular to the direction of flow, and dh/dl is the hydraulic gradient along the flow path, l. By convention, the negative sign implies that flow is along the direction of decreasing gradient.

It should be noted that q has the dimension of velocity and is called *darcian velocity*. This term or its usage is not the same as the true *average linear velocity* or *average pore velocity* or *seepage velocity* of flow through a porous medium; the latter property is given as

$$v = q/n_e \tag{34}$$

where n_e is the effective porosity. This difference should always be kept in mind when estimating the true velocity of groundwater flow and solute transport.

Although it is common to employ hydraulic conductivity in a general sense in studying or describing the hydraulic properties of a porous medium, it is more advantageous to use the term, *transmissivity* (defined below) to describe the ease with which water moves through a large porous medium body such as a horizontal or layered aquifer. Transmissivity, T, (sometimes called transmissibility) is simply the product of hydraulic conductivity and saturated thickness of the aquifer,

$$T = Kb \tag{35}$$

and has the dimensions of L^2/T. This dimensional characteristic derives from the definition of transmissivity as "the volume of water per unit time passing through a unit width area of aquifer perpendicular to flow integrated over the thickness of the aquifer," or $[L^3/(TL^2)]$L (Figure 1.10).

Transmissivity is usually reported in units of square feet per day or square meters per day. The total rate of flow (Q) through any area (A) of the aquifer perpendicular to the flow direction under the gradient (I) is then given as

$$Q = TIA \tag{36}$$

1.4.13 Hydraulic Conductivity of Unsaturated Media

Water in unsaturated media is subject to the influences of not only the hydraulic gradient, but also of molecular attraction and surface tension, as described previously. Therefore, the unsaturated hydraulic conductivity is a function of the pressure head (Ψ), which, in unsaturated media, will always be negative. Consequently, we write unsaturated hydraulic conductivity as K(Ψ). This is discussed later in the section on the unsaturated zone.
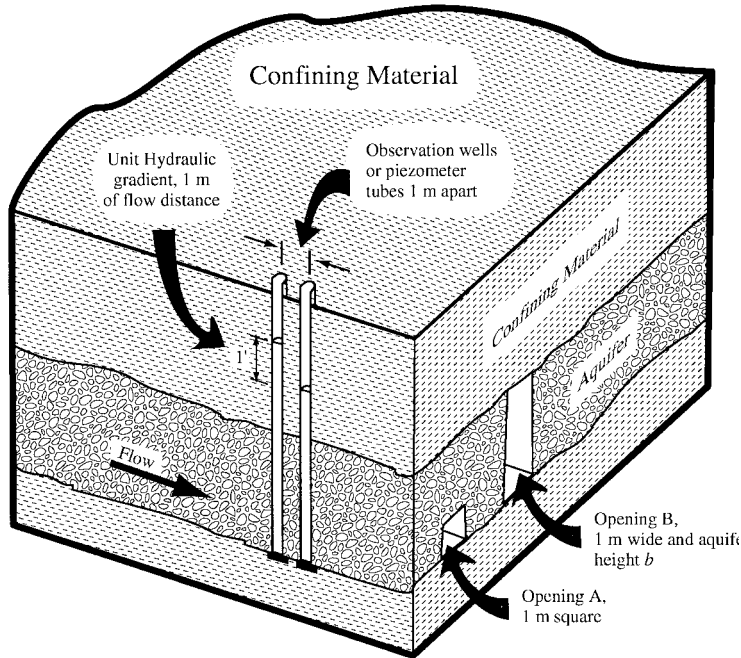
FIGURE 1.10 Illustration of the concept of transmissivity (Adapted from Ferris, J.G., Knowles, D.B., Browne, R.H., and Stallman, R.W. 1962. *Theory of Aquifer Tests.* U.S. Geological Survey. Water-Supply Paper 1536E.)

1.4.14 Anisotropy and Heterogeneity

The discussion to this point has assumed that the porous media is *isotropic*, i.e., the permeability and hence, the hydraulic conductivity and transmissivity, are equal in all directions at any point in the porous medium. If the parameters differ in value directionally at a point, the medium is then said to be *anisotropic*.

These directional properties are described three-dimensionally in cartesian tensor notation using nine general terms, a_{ij}, which can represent permeability, hydraulic conductivity, or transmissivity, but transmissivity is a two-dimensional term that is applied only in a horizontal sense.

$$A_{ij} = \begin{matrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{matrix} \tag{37}$$

If the principal directions of the tensor coincide with the major cartesian axes, the off-diagonal terms will cancel out and the only terms of interest will be a_{11}, a_{22}, and a_{33}. Specific discharge, velocity and hydraulic conductivity are vectors. Thus, for example, in a three-dimensional porous medium with anisotropic hydraulic conductivity, specific discharge will be given as

$$q = q_1 i + q_2 j + q_3 k \tag{38}$$

where i, j, and k are unit vectors in the x, y, and z directions respectively, and q_1, q_2, and q_3 are specific discharge components in the x, y, and z directions, respectively.

Using K_{xx}, K_{yy}, and K_{zz} to represent the main diagonal terms of the hydraulic conductivity tensor, the above equation can be expanded to the following form,

$$q = -K_{xx}\left(\partial h / \partial x\right)\mathbf{i} - K_{yy}\left(\partial h / \partial y\right)\mathbf{j} - K_{zz}\left(\partial h / \partial z\right)\mathbf{k} \tag{39}$$

If the principal directions are along the cartesian directions, it is also common to express anisotropy graphically by use of the equation and figure for an ellipsoid in three dimensions, and an ellipse in two dimensions. Using a two-dimensional example, the semi-axes of an ellipse represent the square root of the property along those axes, and the length in any given direction, r, is the square root of the property in that direction (Figure 1.11).
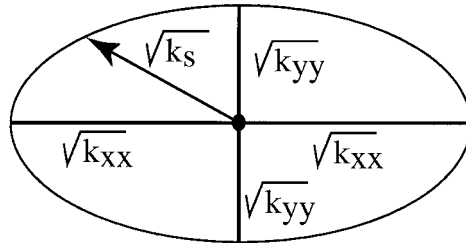
FIGURE 1.11 Elliptical representation of permeability and hydraulic conductivity.

The equation for an ellipsoid is

$$\left(x^2 / a^2\right) + \left(y^2 / b^2\right) + \left(z^2 / c^2\right) = 1 \tag{40}$$

where *a*, *b*, and *c* are the semi-axes. If the medium is isotropic, the ellipsoid degenerates into a sphere for which the equation is

$$x^2 + y^2 + z^2 = r^2 \tag{41}$$

Isotropy or anisotropy of hydraulic conductivity and transmissivity is controlled specifically by the isotropy or anisotropy of the intrinsic permeability. The latter parameter, being a function strictly of the porous medium, is usually determined by the medium structure, which in turn, is the result of the medium's geological origin and subsequent alterations.

For example, sedimentary deposits almost always possess a lower permeability in the vertical direction than in the horizontal. This is due to the fact that as sediments are laid down from water they assume a more stable position if possible where the longer axes of the grains and pebbles are oriented horizontally (Figure 1.12).
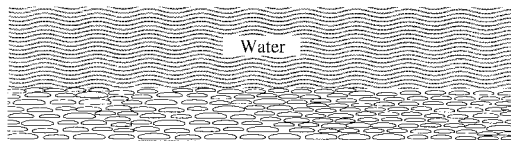
FIGURE 1.12 Natural hydraulic anisotropy in water-lain sediments.

Subsequent fracturing by tectonic forces in a given direction will add a secondary permeability to the medium that is oriented by the forces and may result in greater vertical permeability than horizontal (Figure 1.13).

If the condition of directional equality of properties is the same from point to point anywhere in the medium, the medium is termed *homogeneous*. If the condition of either isotropy or anisotropy varies from point to point, the medium is then said to be *heterogeneous.* Figure 1.14 demonstrates these four
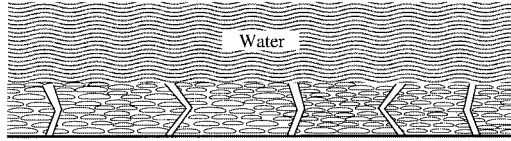
FIGURE 1.13 Anisotropy due to generation of secondary porosity by fracturing.
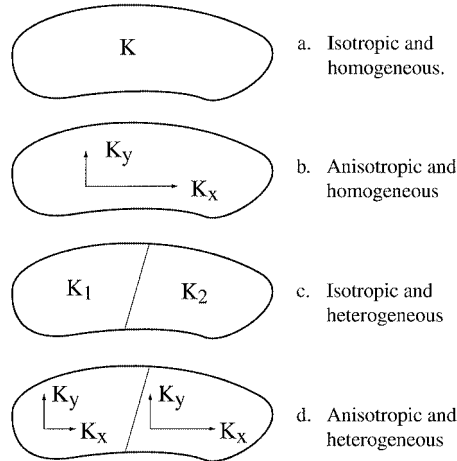

FIGURE 1.14 Summary of possible combinations of isotropy, anisotropy, homogeneity, and heterogeneity.

possible descriptions of a medium — (1) isotropic and homogeneous, (2) anisotropic and homogeneous, (3) isotropic and heterogeneous, and (4) anisotropic and heterogeneous.

The average hydraulic conductivity (K_p) perpendicular to a layered sequence of m beds (Figure 1.15A), each of which is either isotropic or anisotropic, can be determined using a weighted harmonic average as

$$K_p = d \bigg/ \sum_1^m d_m \bigg/ K_{pm} \qquad (42)$$

where d is the total thickness, d_m is the thickness of each layer, and K_{pm} is the perpendicular hydraulic conductivity of each bed.

The average horizontal hydraulic conductivity parallel to the beds is given as a weighted linear average (Figure 1.15B)

$$K_h = \sum_1^m K_{hm} d_m \bigg/ d \qquad (43)$$

where K_{hm} is the horizontal hydraulic conductivity of each bed.

If two isotropic beds of porous media of differing hydraulic conductivities are separated by a plane surface, then a particle of water traveling through one bed will refract upon entering the other (Figure 1.16) according to the *Law of Tangents*,

$$\tan\theta_1 \big/ \tan\theta_2 = K_1 \big/ K_2 \qquad (44)$$

This phenomenon occurs because of the laws of conservation of energy and mass, i.e., one side of a particle will move faster or more slowly than the other at the interface, depending upon whether or not
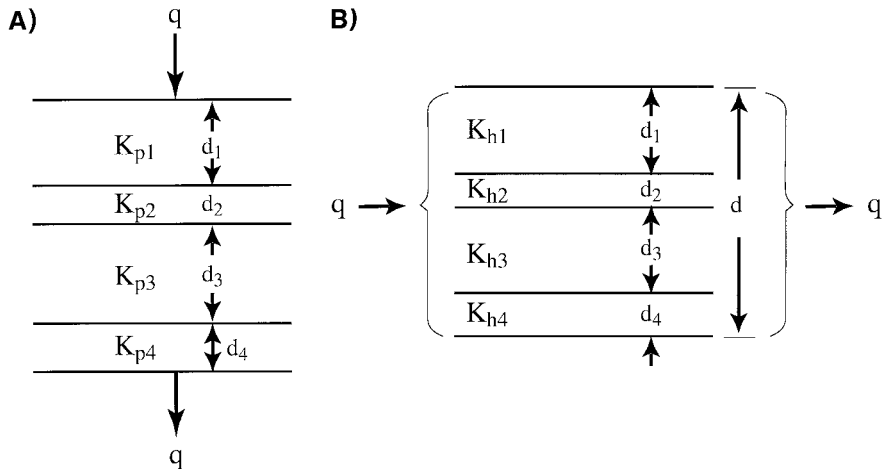
figure 1.15 Hydraulic conductivity of layered systems. A) flow perpendicular to layers; B) flow parallel to layers.
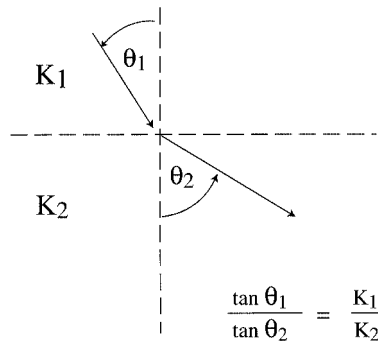

FIGURE 1.16 The Law of Tangents.

a more permeable or less permeable medium, respectively, is encountered. This phenomenon is discussed in Bear (1972).

1.4.15 Aquifers

An *aquifer* is defined by Davis and DeWiest (1966) as "… natural zone (geological formation) below the surface that yields water in sufficiently large amounts to be important economically." This definition is very relative and subjective, for a thin bed of sandstone may economically yield water to a well at a rate of 5.5 m^3/d for a home but would not be sufficient to supply an irrigation well that required 2700 m^3/d. Yet, it could be called a aquifer by strict definition.

The most productive aquifers are generally deposits of glacial outwash, karstic carbonates, permeable sandstones, and highly fractured rocks of all kinds.

An aquifer that is sandwiched between two impermeable layers or formations that are impermeable is called a *confined aquifer* if it is totally saturated from top to bottom (Figure 1.17A).

If the recharge area for the aquifer is located at a higher elevation than the top of the aquifer, and a well is drilled into the aquifer, the water level will rise above the top as shown. Such an aquifer is known as an *artesian* aquifer; it is named after Artois, France, where such wells are common. It should be noted, however, that the well does not have to be flowing to be termed "artesian," although that is the popular conception. A flowing well is known as a "flowing artesian well."
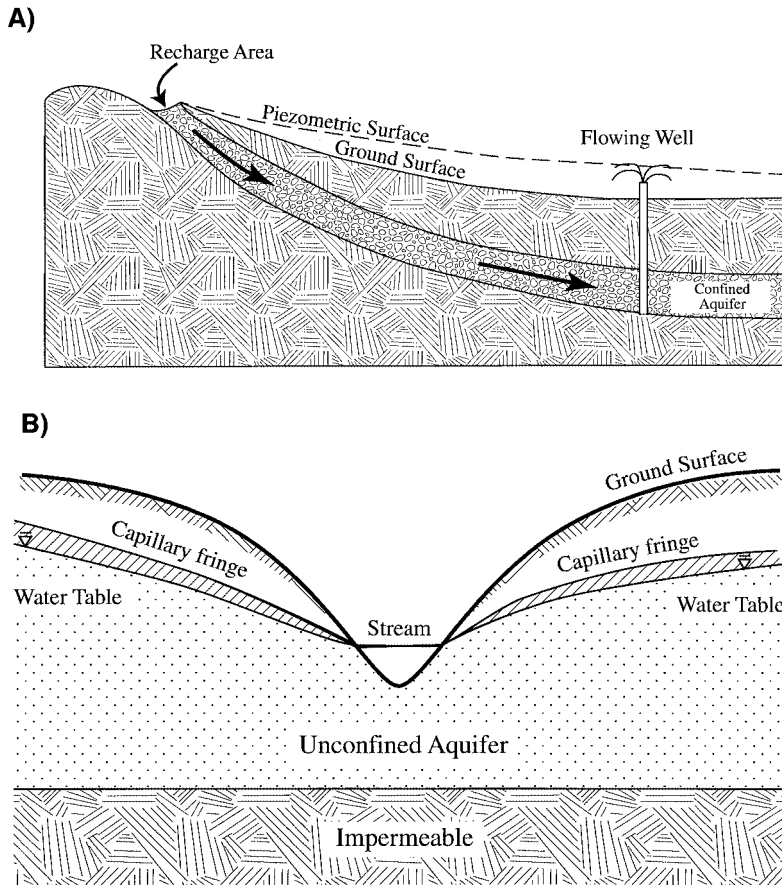
FIGURE 1.17 Kinds of aquifers. A) Confined aquifer; B) unconfined aquifer.

The water level above the top is known as the *piezometric surface* (pressure surface), which is the locus of the piezometric head, and it is not to be confused with the water table discussed below. The piezometric surface occurs above the ground surface because the higher elevation of the recharge area causes the pressure head to rise to such an elevation. The water within the aquifer will be partly under elastic storage. Pumping a well or allowing it to flow will release the water from storage.

Artesian or confined aquifers are common in glaciated regions of the world where a body of outwash sand and gravel may have been covered by clay-rich till or lacustrine sediments from a subsequent glaciation. They may also occur in layered bedrock such as the famous Dakota aquifer of Cretaceous age which rises in the west to lap onto the crystalline rocks of the Black Hills in South Dakota and Wyoming where it is recharged by rain and snowmelt, and extends for a few hundred miles east to Minnesota. Along the way, it is sandwiched between impermeable or slightly permeable shales which maintain its confined and artesian condition. Wells drilled into this aquifer, even hundreds of miles from its recharge area, often flow from the pressure within the aquifer.

An *unconfined aquifer* possesses no overlying confining layer, but may sit upon an impermeable or slightly permeable bed. Therefore, the top of the unsaturated zone of an unconfined aquifer is most often the ground surface, and the top of the saturated zone is usually under negative pressure or tension. This latter property gives rise to the definition of the *water table* which is simply the surface where the relative pressure is zero, i.e., the absolute pressure is atmospheric (Figure 1.17B). Immediately above the water table, the medium is still saturated but the water is held by capillary forces, thus creating a negative pressure head, or tension. This tension can exist even though the pores may be saturated between the
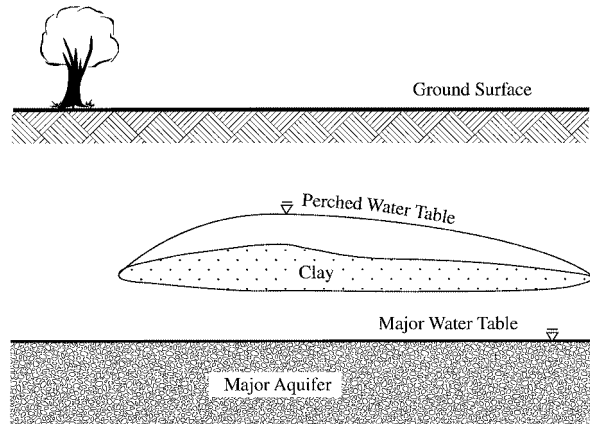
FIGURE 1.18 A perched aquifer.

water table and the top of the capillary fringe. Below the water table, the water pressure increases with depth.

Figure 1.18 shows the cross section of a perched aquifer. These are pockets of sand or gravel or other material that sit on top of impermeable materials such as clay. During periods of high recharge rates, these aquifers may become saturated and actually contain enough water for usage on a temporary basis. Often though, after the onset of dry weather, and after pumping, the perched aquifers will become dry.

Many people make the mistake of finishing a well in a perched aquifer, thinking they are saving money by not having to drill deeper for water, but they end up running out of water when the perched aquifer is depleted. Perched aquifers are common in glaciated regions because a till sheet may contain numerous sand and gravel deposits at various depths that are actually perched aquifers.

A map of the water table in a region can be constructed from water levels measured in wells in the region. In most cases the water table will be a subdued replica of the surface topography. Although the term, "piezometric surface," is generally reserved for confined aquifers with a high pressure head, both water table and piezometric surface can be grouped into the term, *potentiometric surface*.

There is yet another kind of unit between the permeable and the impermeable, known as an *aquitard*, which is semipermeable. A totally impermeable bed is termed an *aquiclude*. An aquifer sandwiched in between an aquitard and an aquiclude, or between two aquitards, is called a *semi-confined* or *leaky* aquifer. It usually possesses properties common to a confined aquifer, such as a high piezometric surface, but some water will flow into or out of the aquifer through the aquitard(s) (Figure 1.19).

In reality, few rocks or formations are entirely impermeable, and the above definitions may be relative to the time period and intensity of pumping stresses imposed on the aquifer. For example, a low-permeability stratum overlying an aquifer may indeed serve as a confining layer under normal and natural pressures. However, when the added stress of heavy pumping is added to the aquifer, the change in head within the aquifer may be great enough to cause some water to move through the overlying bed and the system will then be semiconfined.

Pumping a well in a confined aquifer will release water from elastic storage determined by the elastic specific storage, as discussed above. Pumping from an unconfined aquifer releases water from gravity drainage of the actual media itself, and this is controlled by the specific yield. Recalling that specific yield may be as much as three or four orders of magnitude greater than the elastic storage coefficient implies that much more water can be obtained from an unconfined aquifer per unit of energy expended than from a confined aquifer, providing that the permeability and thickness of the two different aquifers are the same.
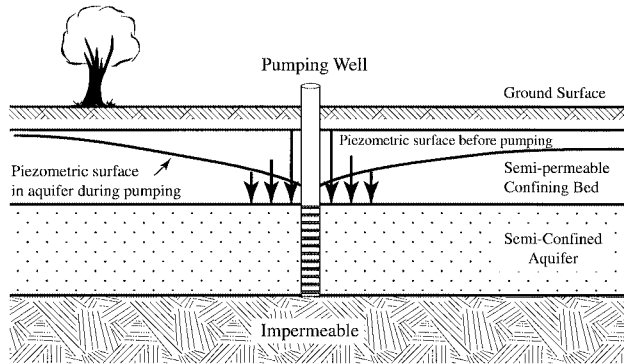
FIGURE 1.19 A semiconfined or leaky aquifer. Vertical arrows indicate relative rates of leakage into aquifer as a function of the distance from the well.

Well yields of a semiconfined aquifer with the same properties as the confined and unconfined aquifers will generally yield more water per unit of energy expended than the confined aquifer, but less than the unconfined, because the semiconfined will yield water from elastic storage plus some from leakage through the semiconfining bed(s), but neither source is as great as the yield from an unconfined aquifer in most cases.

Although more water can be obtained from unconfined aquifers than from the other types, they are more vulnerable to contamination from surface sources than the others because of the lack of a protective confining bed above them. It is for this reason that special care must be taken to ensure the removal and/or remediation of contamination sources above an unconfined aquifer in use or being considered for use as a water supply. This is discussed later under "Well-Head Protection Programs."

1.5   Water in the Unsaturated Zone

1.5.1   Moisture Content Versus Depth

Figure 1.20 shows the general profile of the saturation index from the surface to the water table in an ideal homogeneous and isotropic aquifer. For purposes of discussion, this figure assumes that there have been no recent rains or snowmelt events and that at the surface the moisture content is zero.

From the ground surface to the top of the capillary fringe, the saturation ratio increases from zero to unity and will remain so to the bottom of the aquifer. Below the capillary fringe is the water table at atmospheric pressure which marks the surface delineating the tension zone above from the pressure zone below. The relationships are all the result of a dynamic equilibrium of several forces and conditions —
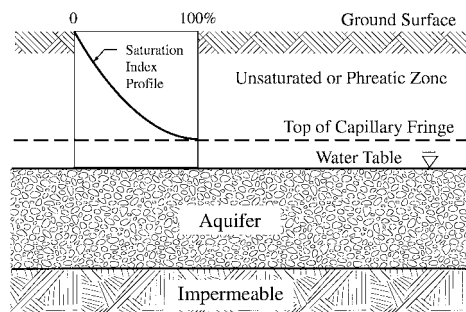

FIGURE 1.20 Saturation index profile versus depth in an unsaturated soil.
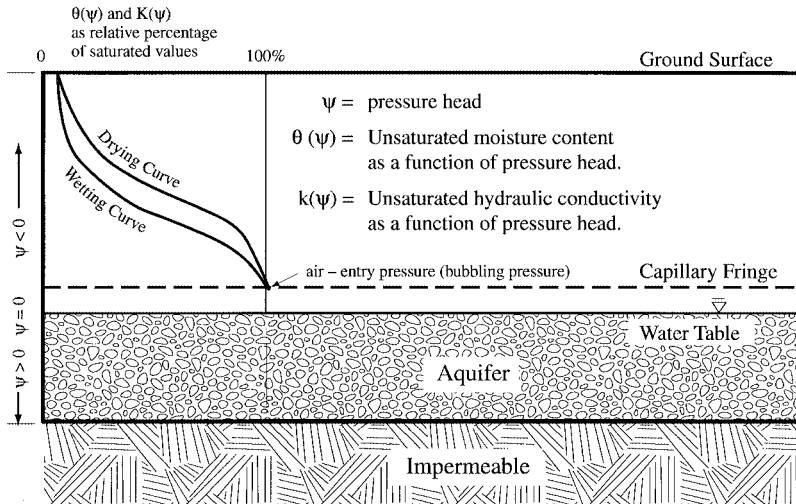
FIGURE 1.21 Soil moisture characteristic curves.

capillary attraction, porosity, initial moisture recharge at the surface, gravitational acceleration, vertical permeability, and evapotranspiration from the surface.

The functional relationship of the moisture content and hydraulic conductivity of the unsaturated profile can be demonstrated by the use of *characteristic curves* as shown in Figure 1.21.

At very low (negative) values of the pressure head, Ψ, both the moisture content and hydraulic conductivity are at minimal values for the system. With increasing values of Ψ, they increase to become constant at the top of the capillary fringe where the saturation ratio is unity, indicating full saturation. Here also, the hydraulic conductivity is no longer a function of the pressure or moisture content, but attains a maximum value at which it remains, providing no changes in the porous media occur. The pressure head at this point is called the *bubbling pressure* or the *air-entry pressure*; at a lower pressure, air will enter the system causing loss of saturation. As the pressure increases with increasing depth in the

range, marking the water table. With still-increasing pressure, both moisture content and hydraulic conductivity remain constant.

The lag between the wetting and drying curve shapes is called *hysteresis*. This is due to the fact that initial wetting of the porous medium is enhanced by strong capillary and adhesive forces, but these same forces tend to hold water in the unsaturated pores and it requires more suction or greater negative pressures to drain the pores, as illustrated by the offset of the drying curve from the wetting curve.

It should be noted that characteristic curves of fine-grained materials such as silt and clay will be much steeper in shape than those of coarse sand and gravel. This is due to the fact that in the coarse-grained material with relatively large pore diameters, the effects of capillarity and adhesion are not as great as in the fine-grained material. It is for this reason that the capillary fringe in clay is much thicker (one or two meters) than in coarse materials (a few centimeters at most).

In reality, soil moisture is not distributed evenly in the unsaturated zone, but varies both vertically and horizontally due to variations in soil types, crop types, infiltration rates, etc. Recently, new work has been done in expressing the variability of soil moisture in unsaturated chalk with electrical tomography (Andrews et al., 1995). This method, although in the experimental stage, does allow visual imaging of soil moisture contents based on subsurface electrical resistivity of various zones.

1.5.2 Recharge and Infiltration Capacity

Recharge is the actual entering of water into an aquifer, whereas *infiltration* is movement of water from the surface into the ground. Usually, the recharge source is surface water from precipitation and to a lesser degree, irrigation or artificially constructed recharge pits or losing streams (discussed below). The actual recharge rate is controlled by several factors: (1) the amount and rate of precipitation not lost to surface runoff and evapotranspiration, (2) the initial soil moisture content or saturation ratio of the soil, (3) the elevation of the recharge surface relative to the discharge area, (4) the horizontal hydraulic conductivity of the aquifer being recharged and its hydraulic gradient, which determine the rate at which recharged water will be carried out of the recharge area, (5) the vertical hydraulic conductivity of the soil being recharged, and (6) the presence of man-made alterations to the subsurface, such as drainage tiles that carry water away to run off in surface streams.

People are often surprised to learn how little precipitation is actually recharged in even moist climatic zones. For example, the state of Indiana receives approximately 91.4 cm of precipitation water per year. Yet a statewide average of only about 10% (9.14 cm) is actually recharged into the groundwater system, the rest being lost to evapotranspiration, surface runoff, and subsurface drainage tiles. Recharge is not evenly distributed, with perhaps eight to ten times the amount going into more permeable sands and gravels than into till.

Figure 1.22 shows a hypothetical cross section of a recharge/discharge system with recharge taking place in the higher elevation and discharge in the lower. The stream lines are perpendicular to the equipotential lines. Discharge areas are often seen on the ground as wet spots with seeps or springs and often with lush vegetation.
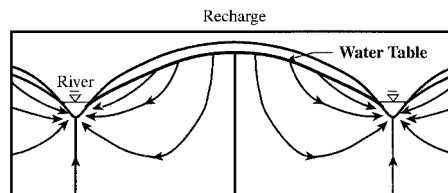

FIGURE 1.22 Hypothetical recharge/discharge system. (Adapted from Lohman, S.W. 1972. *Ground-Water Hydraulics.* U.S. Geological Survey Prof. Paper 708, U.S. Government Printing Office, Washington, DC.).

Recharge rates are not easy to determine due to the variability of soils and the necessity of extensive field instrumentation. One very effective method has been the actual tracking of infiltrating water by analyzing the concentrations of atmospheric tritium at depth in the soil. This procedure was very successful in Indiana (Daniels et al., 1991), and in Minnesota (Kanivetsky and Rumynin, 1993). Tritium in the atmosphere, due to atmospheric nuclear weapons tests in the '50s and '60s, is incorporated into precipitation which is eventually recharged into the ground. The concentration of tritium in a vertical drill core can be used to estimate the rate of recharge.

In addition to elevation and soil type and its variability, recharge is controlled by the *infiltration capacity* of the soil. This phenomenon was studied by Horton (1933) and is the maximum rate at which a soil will permit the entry of water. It is generally a time-dependent parameter because it depends upon the rate of change of the saturation ratio and hence, the moisture content.

This relationship is illustrated with Figure 1.23. Here, an initially dry soil is rained upon at a constant rate. Initial infiltration into the soil is rapid as the empty pores begin filling. As infiltration continues, the pores gradually lose capacity for additional water at the initial rate. As a result, the infiltration capacity decreases to an eventual steady-state rate where the infiltrating water enters at the same rate as that at which it is transported downward by porous media flow.

Once the infiltration rate falls below the precipitation rate (or snowmelt rate), the excess water will become runoff known as *overland flow*. Generally, this will happen after all surface storage requirements are met, e.g., leaf storage in trees, crops, and grasses, and depression storage in hollows in the ground. From these considerations, it is obvious that the most effective rain for watering plants and recharging
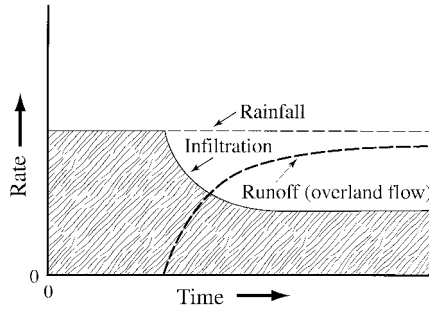
FIGURE 1.23 Illustration of infiltration capacity.

groundwater supplies is the slow steady kind, and not the intense, high-rate storm, most of which runs off and does not infiltrate.

1.5.3 Hydraulic Conductivity and Specific Discharge Through the Unsaturated Zone

When the infiltration capacity is reached, the unsaturated hydraulic conductivity is then at its maximum. When the soil is initially dry, water will not flow any significant distance until the water content in the soil is sufficient and the pressure head becomes less negative, as demonstrated with the characteristic curves above. The specific discharge is then given as

$$q = -K(\Psi)\,\partial h/\partial l \tag{45}$$

or

$$q = -K(\theta)\,\partial h/\partial l \tag{46}$$

where θ is the water content of the soil. Three methods for computing unsaturated hydraulic conductivity as a function of water content are given below and are summarized in Rawls et al. (1993). The equations below are dimensionless, making them applicable to any consistent set of units.

Brooks and Corey (1964) give the formula for hydraulic conductivity as

$$K(\theta)/K_s = \left[(\theta - \theta_r)/(n - \theta_r)\right]^m \tag{47}$$

where K_s is saturated hydraulic conductivity, θ_r is residual water content, n is porosity, and

$$m = 3 + 2/\lambda \tag{48}$$

where λ is the pore-size index (Brooks and Corey, 1964).

Campbell (1974) presents the following formula:

$$K(\theta)/K_s = (\theta/n)^m \tag{49}$$

where

$$m = 3 + 2/\lambda \tag{50}$$

Van Genuchten (1980) presents the formula,

$$K(\theta)\big/K_s = \left[(\theta-\theta_r)\big/(n-\theta_r)\right]^{1/2} \left\{1-\left[\left((\theta-\theta_r)\big/(n-\theta_r)\right)^{1/m}\right]^m\right\}^2 \qquad (51)$$

where

$$m = \lambda\big/(\lambda+1) \qquad (52)$$

1.5.4 Residence Time

Water exists in different forms and in different places and for widely varying times in the hydrologic cycle. Freeze and Cherry (1979) state that soil moisture may stay in place for two weeks to one year. This rate is an average and depends on a wide variety of conditions such as permeability, infiltration rate, plant use, hydraulic gradient and anthropogenic withdrawal rates.

1.5.5 Subsurface Stormflow or Interflow Zone

In many regions of the world, especially in tilled crop land and forest litter on sloping ground, infiltration into the upper one or two meters of soil will be quite rapid, but continued infiltration into the substratum will be much lower. The upper part often possesses much greater permeability than the lower part due to tilling, root activity, and movement of soil by worms, insects, and small mammals.

If the hydraulic conductivity of the lower part is less than the upper, then subsurface runoff will occur in the upper part, or the *subsurface stormflow* or *interflow zone*, as it is called. This zone may carry water all the way to a stream, or it may discharge it along the way at a lower elevation. Figure 1.24 illustrates the interflow zone.
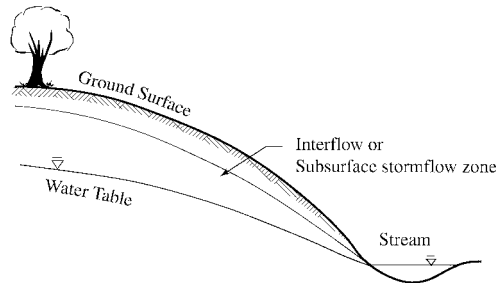
FIGURE 1.24 The interflow or subsurface stormflow zone.

1.5.6 Atmospheric Discharge and Seepage Face

In a discharge area, at the juncture point where groundwater enters a stream or lake, the water f wi ll coincide with the elevation of the surface water. Beyond this point, the hydraulic gradient is zero and groundwater has completed its journey from recharge to discharge.

Also at this point, there will generally be a zone of wet soil extending above the juncture point (Figure 1.25). This zone, called the *seepage face,* is wetted by capillary attraction of the water above the water table, and it is essentially the surface exposure of the capillary fringe. A small fraction of the discharge from the groundwater system will escape into the atmosphere at this point. The height of the seepage face will depend upon the grain size of the soil — greater in fine-grained and lesser in coarse-grained sediments.
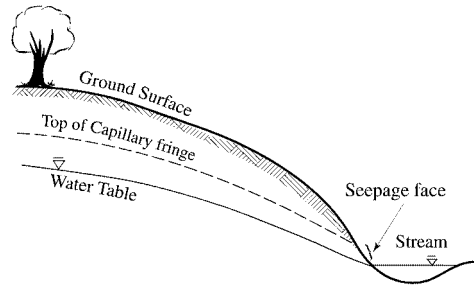
FIGURE 1.25 The seepage face.

1.5.7   Discharge to the Saturated Zone

Water in the unsaturated zone will percolate downward from the surface and gradually increase in content to a maximum, where the saturation ratio is unity at the top of the capillary fringe. At this point, it will flow downward to the water table under saturated flow conditions. The rate of flow to the saturated zone will depend to a large degree on the vertical permeability and the rate at which the groundwater in the saturated zone is being carried away by lateral groundwater flow.

1.6   Water in the Saturated Zone

In saturated pores, the forces of adhesion are equal in all directions, so there is no directional attraction. The only dynamic or motive forces in action are gravity and the force represented by the gradient of the fluid potential. Resistive forces due to viscosity (resistance to shear) work in opposition to the dynamic forces.

Darcy's Law is assumed to be in effect in porous media flow, and the flow is assumed to be laminar, i.e., the *Reynolds number*

$$R_e = \rho v d / \mu \tag{53}$$

has a value from 1 to 10, where ρ is water density, v is average pore velocity, d is average pore diameter, and μ is dynamic viscosity of water at a given temperature.

The average pore diameter is probably the best estimate of pore properties that can be attained on a large scale. However, the pore size in a medium can range over even a few orders of magnitude in very heterogeneous media with high uniformity coefficients (greater than 4). Therefore, it is unlikely that laminar flow occurs in all pores at the same time. Large pores (including fractures and karst conduits) may allow passage of water at velocities high enough to be nonlinear laminar or transitional ($4 \leq R_e \leq 100$), and in some cases, turbulent ($R_e > 100$). In small pores, the frictional resistance may be too great to allow flow faster than laminar. In spite of these uncertainties, hydrogeologists generally accept the assumption of laminar flow throughout a granular porous media with $R_e < 100$ for the sake of simplifiying the mathematics of flow.

Groundwater may flow for some distance over a wide area in the laminar state, but as it approaches a discharge point (e.g., a spring or well) much narrower than its upgradient flow field, the streamlines will crowd together and the flow velocity will increase to maintain the same volumetric discharge rate. Near or at the point of discharge the velocity will often be great enough to be actually turbulent.

Under such conditions, the specific discharge will be expressed as

$$q = -K \left(\partial h / \partial l \right)^m \tag{54}$$
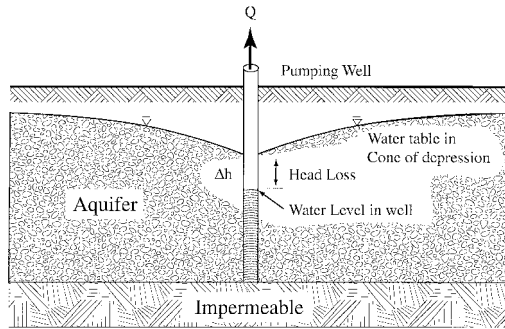
where $m > 1$.

FIGURE 1.26 Flow to a pumping well with head loss.

Figure 1.26 illustrates a situation where flow can change from laminar to turbulent as it approaches a pumping well. Assuming that the aquifer is homogeneous and isotropic and that the flow is steady-state, the flow lines converge as they approach the well.

At some distance from the well the equation of steady-state flow is given as

$$q = -K\left(h/r\right) \tag{55}$$

where r is the radial distance from the well center. At or near the well, the equation becomes

$$q = -K\left(h/\left(r\right)\right)^m \tag{56}$$

where m > 1 and generally closer to 2.

Steady-state laminar flow into the well will require a gradient such that the water level in the well casing is the same elevation as that immediately outside the well casing in the aquifer. Turbulent flow into the well requires more energy and thus a steeper gradient than laminar flow. In order to accomplish this, the water level in the well will have to drop below the level required for laminar flow as demonstrated in Figure 1.26.

This extra head drop is known as head loss or *well loss*, and is undesirable. Well loss means extra energy is needed which reduces the efficiency of the pumping system and drives up the cost of pumping water. In addition, because a greater pressure drop occurs with head loss, gasses will be more likely to come out of solution and cause precipitation of lime ($CaCO_3$) and iron and manganese oxides (Fe_2O_3 and MnO_2, respectively) onto the well screen which causes clogging of the screen and further reduces the efficiency of the well.

1.6.1   Residence Time

Groundwater occurs over a wide range of depths, from near the surface to even thousands of meters below. It has been estimated that the residence time of groundwater can vary from as little as two weeks to more than 10,000 years (Freeze and Cherry, 1979). Some water presently being extracted from aquifers in the north central United States is believed to have been recharged during the Pleistocene Epoch more than 10,000 years ago.

Tóth (1962) introduced the concept of flow system lengths and applied it to the prairie pothole regions of the Canadian high plains. The same concept (the "Prairie Profile") applies in many other regions and to many other groundwater regimes. Tóth's three systems are the (1) local systems, (2) intermediate systems, and (3) regional systems, as demonstrated by Figure 1.27.
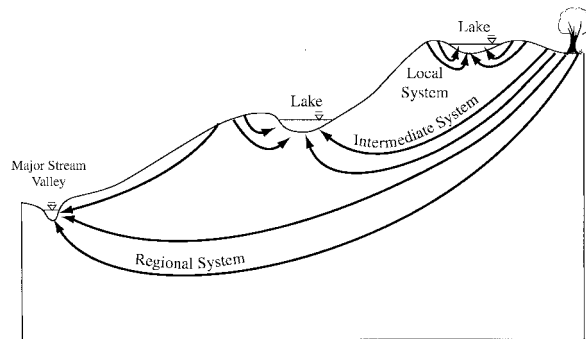
FIGURE 1.27 Local, intermediate, and regional flow systems.

The local flow system consists of relatively short cells with discharge a few tens to a few hundreds of meters from the recharge areas. Such systems generally contain water of the best quality, i.e., the lowest concentration of total dissolved solids, because of the water's short residence time and the short flow paths. Residence times may be only a few months or a few years. Such systems are commonly recharged on the interfluve high ground between glacially created ponds, and they discharge to the ponds a few tens or hundreds of meters away.

Intermediate systems cover a larger area of perhaps a few hundred to a few thousand meters and residence time may be tens of years . Dissolved solids concentrations are greater than in the local systems.

Regional systems may cover many kilometers and be as large as a state. A typical example is the Dakota Aquifer which is recharged in the Black Hills of South Dakota and yet extends for a few hundred miles eastward.

Most regional systems discharge into major rivers or large lakes, such as the Mississippi or the Great Lakes, respectively. In regional systems, the total dissolved solids (TDS) can be quite high. For example, it is not uncommon to find as much as 2500 to 3000 mg/l of TDS in the Dakota Aquifer. Residence times can exceed several thousands of years.

Fossil water is that water trapped deep in the earth and incorporated with original sediments. Generally, it is of geologic time in age and contains such high concentrations of dissolved minerals that it is often considered a brine. *Connate water* is groundwater which has circulated so deeply in the flow system that it too is excessively old, approaching geologic time in many cases. It also is briny in most cases. Both kinds of water may contain large concentrations of magnesium, calcium, sodium, potassium, chloride, sulfate, carbonate, and bicarbonate. Often these brines contain dissolved matter in concentrations exceeding 100,000 mg/l, and as such are known as *bittern brines* (Hem, 1985).

1.6.2 Surface Discharge

When streams are incised into aquifers, or even into saturated soil or rocks which cannot be classified as aquifers, the streams normally serve as discharge lines or sinks for groundwater as it leaves the groundwater system. Groundwater is seldom static and streams carry water back to lakes and oceans where it is naturally recycled by evaporation.

Between storms and runoff events, stream flow is maintained by groundwater discharge known as *baseflow,* as long as the water table remains above the stream bottom. After a storm or snowmelt event, the stream will flow at its highest rate and will continue to flow at an ever-decreasing rate until the next precipitation event. During dry weather, the water table may even drop below the stream bottom which will cause stream flow to cease. The history of these events and relationships is shown in a *hydrograph* in Figure 1.28.

The falling limb of the hydrograph is a record of both overland flow and baseflow. Over a period of many years, such falling-limb records of a particular stretch of stream can be analyzed by regression
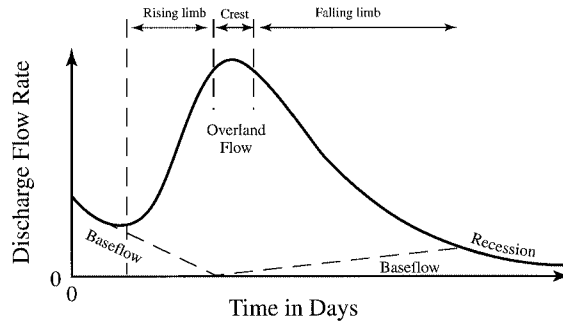
FIGURE 1.28 Hydrograph of stream flow.

analysis to derive a number unique to that particular stretch and its upstream contributing basin. This number, a, the *base flow recession constant*, is used in the following equation to compute flow rates,

$$Q = Q_o e^{-at} \qquad (57)$$

where Q is the flow rate at any time, t, Q_o is the initial flow rate, and t is elapsed time since time of the initial flow rate. The reader may note that this equation, a decay-curve equation, is of the same form as the classical decay curve equation which is also applicable to radioactive decay, using different units.

The baseflow recession constant of a stream or a certain stretch of a stream may change only slightly over a long period of time if neither the environment nor the climate changes appreciably. However, anthropogenic alterations such as timbering, farming, installation of field drainage tiles, construction of roofs and pavements, and damming of the stream or diversion of its water can significantly change the recession constant in a fairly short time.

In water-management and water-budget studies, it is often necessary to separate the various components of the hydrograph to determine the contribution of groundwater, surface water, etc. This is not an easy task, but three commonly employed methods are described in the following paragraphs:

1. *Direct-measurement method*: Accurate knowledge of the groundwater hydraulic gradient on both sides of the stream and an accurate estimate of the hydraulic conductivity must be obtained beforehand. The first can be gotten by installation of piezometers, in which the water levels are regularly measured. Hydraulic conductivity can be determined with pumping tests or from laboratory determinations from grain-size distributions. Direct measurement of stream flow is carefully taken, and evapotranspiration is estimated by proper formulas. Direct precipitation is also recorded.

   Using Darcy's law, the specific discharge from groundwater can be calculated to yield an estimate of groundwater contribution to total stream flow. If groundwater discharge, precipitation, and evapotranspiration are then known, they can be subtracted from total stream flow at various times in the time period of the hydrograph to determine the overland flow component during these times.

2. *The curve tangent method*: Pilgrim and Cordery (1993) describe a graphical estimation method in which tangents are drawn to recession curves at the points where overland flow starts and is assumed to end (breaks in the hydrograph slopes). Under conditions where maximum baseflow discharge is below 10% of the maximum discharge, a straight line can be drawn between the two tangent points which will be an acceptable estimate of the baseflow discharge component.

3. *The basin-area method*: Linsley et al. (1975) give the formula for calculating D, the number of days between the storm peak of a hydrograph and the end of overland flow, for a basin of area, A (in square kilometers), as,

$$D = 0.827 A^{(0.2)} \tag{58}$$

The hydrograph recession limb that existed before the storm is extended until it is under the hydrograph peak. A straight line drawn from this point to the point on the hydrograph corresponding to D is then considered to be the graph of the baseflow component.

4. *The chemical and isotope method*: Stream water and groundwater supplying the stream's base flow can be analyzed for a variety of chemicals and isotopes during periods of normal flow to obtain normal background levels. Typical analytes include ^2H, ^3H, ^{18}O, Na^+, Ca^{2+}, Mg^{2+}, Cl^-, SO_4^{-2}, and HCO_3^-. Electrical conductance has also been used. During periods of high flow, the stream water is sampled again and analyzed for the same isotopes (Freeze and Cherry, 1979).

If the chemicals or isotopes that were high in concentration in the groundwater are also high in the stream water, then it can be assumed that a significant percentage of the surface flow is actually being contributed by groundwater. On the other hand, if the groundwater isotopes are low in the stream water, then it can be assumed that most flow originates from surface runoff.

Highly permeable soils will contribute significantly to surface flow during precipitation. Low-permeability soils will contribute less, and the isotope analyses will show these differences.

A *spring* is a natural discharge point for groundwater to the surface and the atmosphere. Several geological factors can cause springs of various types to occur, as listed below.

1. *Geological contact springs*: When a saturated permeable formation sits on a stratum of low permeability, it will often result in a spring at the contact. This is a common occurrence in glaciated regions where saturated sand and gravel sits on till. If the water has not been contaminated by surface sources, the water from such springs is often of good quality because it has not traveled far before discharge and therefore, little mineral matter has been dissolved.

2. *Karst springs*: In limestone and dolomite formations, springs will often discharge from the points on a valley side where caves have been intersected by incising stream valleys. These are common in areas of the world where the valleys are deep enough and precipitation and recharge are great enough to allow a water table to exist well above the streams.

The Fontaine de Vaucluse in southern France discharges between 8 and 150 m^3/s and is one of the largest springs in the world. It is supplied by precipitation on the fissured carbonates of the Vaucluse Plateau. The extensive network of karst conduits ends at the point of discharge.

Another kind of karst spring is the sinkhole spring. This feature is simply a spring where a sinkhole intersects karst conduits. Springs of this kind are very common in the limestone of the state of Florida, with names like Silver Springs, Tarpon Springs, and others. Many large lakes in central Florida are actually large sinkhole springs.

3. *Structural springs*: When there is significant movement along faults such that permeable and water-bearing formations are juxtaposed against impermeable rocks; or fracturing opens pathways into aquifers, water can emanate from these features to discharge at the surface. Such springs are not uncommon in the Great Basin of the U.S. where tension in the earth's crust has produced a large area of block faulting Nevada, Utah, and parts of Arizona and Idaho. One such example is the Ash Meadows Springs in the Amargosa Desert of Nevada, just east of Death Valley, where discharge along a fault spring in carbonates is approximately 2.1×10^7 cubic meters per year.

4. *Depression springs*: If the land surface is lowered below the water table, a depression spring will form and will actually create a small local flow system, with the spring being the discharge point. Such depressions can be created by wind in areas of fine sand, or by landslides, tectonic activity, or even the collapse of land into underlying sinkholes or mine tunnels.

Many springs are known for their size and have been utilized for water supplies, spas, and tourist attractions. Examples are Steamboat Springs, Colorado; Hot Springs, Arkansas; Big Springs, Texas; White Sulfur Springs, West Virginia; and many others.

Ponds and lakes can form in many ways. Commonly in glaciated regions like the midwestern United States, Canada, and northern Europe, present-day lakes and ponds are the results of ice blocks that were covered with drift and later melted to form permanent depressions in the new landscape. Glaciers in mountain valleys may leave large debris dams at their termini which cause ponding of water. Sinkhole lakes are common in Florida, Kentucky, Indiana, and other karst areas. Landslides can dam up rivers. Tectonic forces can cause certain areas, as in the Basin and Range areas of the United States, to subside relative to the surrounding areas, e.g., Death Valley, which contained a large lake during the Pleistocene Epoch. Volcanic craters may contain ponded water like Crater Lake, Oregon. Oxbow lakes can form when large rivers meander and eventually, a meander loop is cut off from the main stream.

Surface bodies of water are nearly always groundwater discharge areas. Very large lakes, for example the Great Lakes in North America, serve as discharge areas for regional, intermediate, and local systems upgradient from them. The smallest ponds of a few acres in area discharge water from local systems.

The dynamic role of lakes and ponds in the movement of groundwater is most significant for two reasons. In the first place, the generally low topographic positions of these bodies in relation to the upgradient flow system feeding them cause water to flow toward them under the force of gravity, and the size of the lake and its elevation relative to other bodies of water in the area can determine the relative size of the groundwater system feeding it; i.e., very large and low-elevation lake surfaces can be sinks for regional systems, whereas small ponds at relatively high elevations will generally serve only local systems.

Second, lakes, especially those with large surface areas, can allow significant evaporation to take discharge water out of the system and put it back into the atmosphere. This process is especially important in hot summer months when actual evapotranspiration can equal potential evapotranspiration. For example, Leap (1986) discusses an area on the Coteau des Prairies in northeastern South Dakota where 6% of the surface area is covered with glacially derived ponds and lakes, ranging from a few hundred meters to a few kilometers in diameter, that drive local and intermediate systems, respectively. The potential evapotranspiration rate from May through September, when transpiration is highest, was calculated to be 0.61 m.

Assuming, for example, that in one square kilometer of the area, 6% of the area, or 60,000 m^2 is covered with water, then the total amount lost from the lakes is approximated as 60,000 $m^2 \times 0.61$ m or 36,600 cubic meters of water during the period from the first of May through the end of September. This is a conservative estimate because minimal evapotranspiration occurs before and after this time before the winter freeze sets in.

For illustrative purposes, assume that the one-square-kilometer area is underlain by an aquifer five meters thick with a porosity of 25% (characteristic of this area). If fully saturated, this aquifer would then contain 1,250,000 m^3 of water. Thus, the 36,600 m^3 of water lost by evapotranspiration would equal 3% of the total water contained in the aquifer.

This amount does not seem like much over the entire one-square-kilometer area, but the lowering of the water surface of many ponds by even a small amount by evapotranspiration sets up gradients sufficient to drive many local flow systems. The same amount of loss in one large lake of the same area can cause gradients sufficient to drive even intermediate systems.

Meyboom (1967) discusses the role of evapotranspiration in prairie pothole lakes in western Canada. His study showed that the role of vegetation, especially water-using phreatophytes like willow, around the edge of a lake can be quite significant. In many areas, the water table will actually be depressed beneath these trees.

The depression of the water table below the root zone can be studied quantitatively with a method described by White (1932). Using his equation, one can estimate the amount of water removed per day from the groundwater system by evapotranspiration. An example is given in Todd (1980).

Thus, it is the interaction of the dynamics of gravity pulling rain and snow to the surface and then moving groundwater to the lakes (and streams), and evapotranspiration removing it from the earth's surface that maintains the circulation of groundwater. If this were not so, the residence time in the ground would be so long that groundwater would dissolve so much mineral matter that it would become
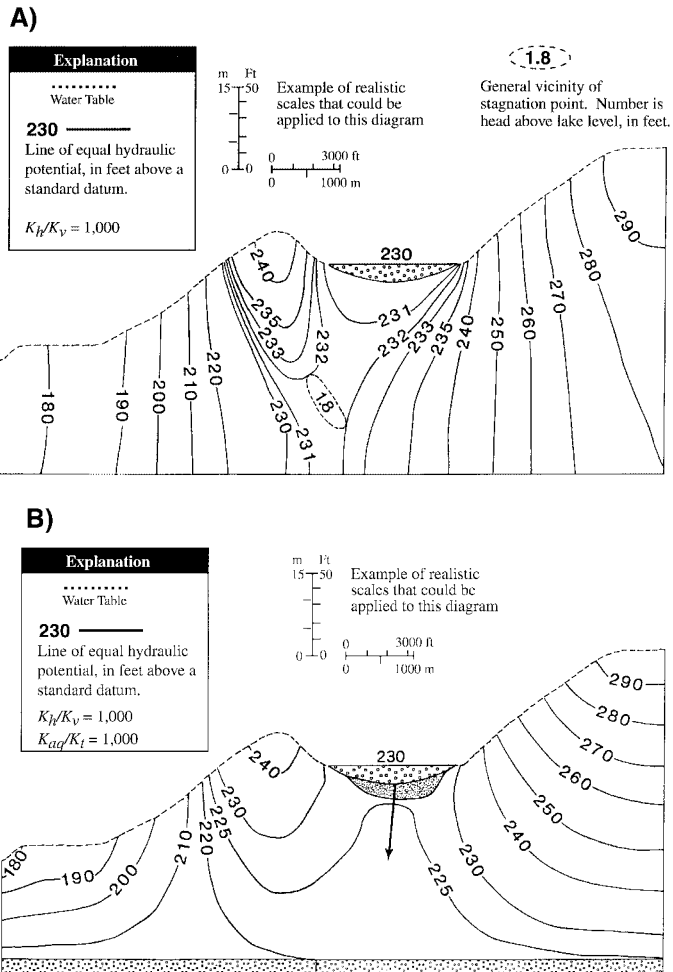
FIGURE 1.29 Cross section showing lake/aquifer interactions. A) Anisotropic and homogeneous aquifer; B) Anisotropic and heterogeneous aquifer. (Adapted from Winter, T.C. 1976. *Numerical Simulation Analysis of the Interaction of Lakes and Ground Water.* U.S. Geological Survey Prof. Paper 1001. U.S. Government Printing Office, Washington, DC.).

unpotable. Through this interactive process, water on earth is constantly cleansed before being reintroduced into the system.

Definitive modeling studies of the three-dimensional physical dynamics of groundwater flow to and from lakes were made by Winter (1976). Two contrasting examples of several modeled scenarios are shown in Figures 1.29A and B in order to illustrate the importance of permeability contrasts in the subsurface below lakes.

In Figure 1.29A, the subsurface is anisotropic, but homogeneous. Figure 1.29B shows the case where the subsurface is anisotropic and heterogeneous. Given that there is an ample supply of recharge, a high-permeability layer at depth can be great enough to divert much of the flow toward the lower layer. Thus, high-permeability layers at depth are significant in causing water from the surface to recharge the subsurface. Knowledge of the positions and hydraulic characteristics of such layers is critical to estimating recharge rates for water-management purposes, and for estimating the potential for contaminants from the surface to reach potable groundwater supplies.

Ideally, there will be a point beneath a lake undergoing both recharge and discharge, to and from the subsurface, respectively, where the hydraulic gradients will be in opposition to the extent that flow does

not take place. This is called the *stagnation point* and it is shown in Figure 1.29A. The stagnation point is in dynamic equilibrium with the head and gradients that exist at any one time, but will change position with changes in recharge rates from the surface, or discharge from the subsurface. Therefore, any attempts to dispose of waste at a stagnation point, as has been suggested in the past, would not be wise because the location will change under most circumstances and soluble waste will then migrate.

An additional factor that can greatly affect the inflow/outflow characteristics of lakes is the presence of low-permeability sediments on the bottom. This is often a problem in artificial recharge pits where such sediments can lower the infiltration rate of water from the pits to the subsurface. The influence of the sediments can be quantified by Equation 42, from which it can be shown that even a thin layer of low-permeability sediments can make a very large difference in the infiltration rate.

1.6.3 Gaining and Losing Streams

Streams also gain and lose water in the same manner as lakes and ponds (Figure 1.30A and B). *Perennial streams*, i.e., streams that flow year-round, are usually *gaining streams*. They gain water from base flow through the sides of the streams as it flows through from the groundwater system. This kind of stream is situated such that the water table always rises above the stream surface.

Gaining streams in the cross section (Figure 1.30A) will usually show a stagnation point in the subsurface similar to that of lakes. However, in three dimensions, it is in reality a stagnation line that runs under the stream along its length.

Most streams are gaining streams; losing streams (Figure 1.30B) are generally those *intermittent streams* which flow only after significant precipitation and during runoff periods. They lose water to the subsurface
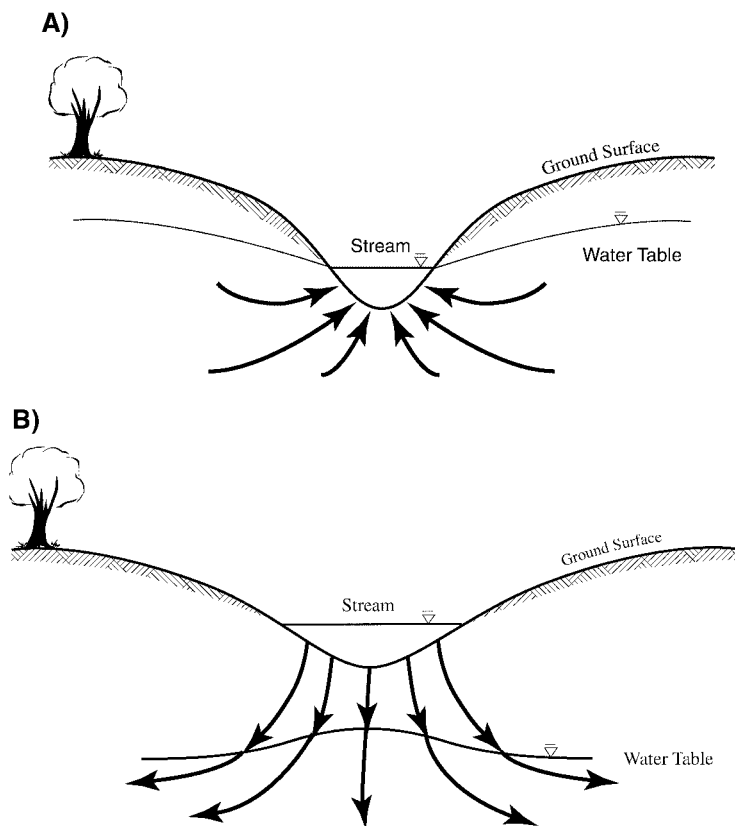

FIGURE 1.30 Gaining and losing streams. A) Gaining stream; B) losing stream.

because the water table is below the stream surface. Therefore, losing streams are generally found in mountainous areas on alluvial fans debouching onto pediments, on sand and gravel surfaces where the water table is low, or on steep slopes. In nearly all cases, the precipitation rate is insufficient and/or the subsurface is too permeable to maintain a water table high enough to support gaining streams.

Streams in certain geological settings may contain stretches which are gaining for some distance where the water table is above the stream surface, and then become losing stretches when they flow over a low-water-table area, or vice-versa. These streams are common in karst terranes where a stream may be losing in its upstream reaches, and may even disappear beneath the ground surface into a cave; downstream, it may later emerge to flow on the surface as a gaining stream. In glaciated areas of the midwestern United States, streams may flow over till with high water tables as gaining streams, but after crossing a contact between till and outwash sand and gravel, the streams may become losing streams as their water infiltrates downward into the permeable outwash.

In areas where there are heavy demands for water for irrigation, industry, and large cities, damming of streams and pumping of aquifers has dramatically altered natural stream/groundwater interactions. Damming often raises the water table upstream of the dam, causing some losing streams to become gaining streams, and lowers the water table downstream, causing gaining streams to become losing. Heavy pumping has caused some small gaining streams to become losing because of drastic lowering of the water table. When a gaining stream does become losing, it can introduce surface contaminants into the groundwater system.

1.6.4 Bank Storage

Figure 1.30A shows a gaining stream during normal baseflow when the water table slopes toward the stream, and Figure 1.31 shows the same stream after a sudden rise in the surface elevation following an intense precipitation event. In the latter case, the water table near the stream has risen above the normal water-surface elevation and now slopes away from the stream to force water to move from the stream into the banks and beyond.

After the stream stage falls back to its normal position, the water table near the stream may not fall as fast and may remain unusually higher than the stream surface for several days or hours, depending on the permeability of the ground and the amount of water that was forced into the banks. This excess water in the stream banks is known as *bank storage* and is common along rivers that flood.

In bottom lands along alluvial stretches, bank storage can slow natural drainage back into the stream, and thus, can cause farm fields along the stream to remain soggy for extended periods. If the stream is contaminated, the contaminants can be forced into the surrounding soil, and they may not be flushed
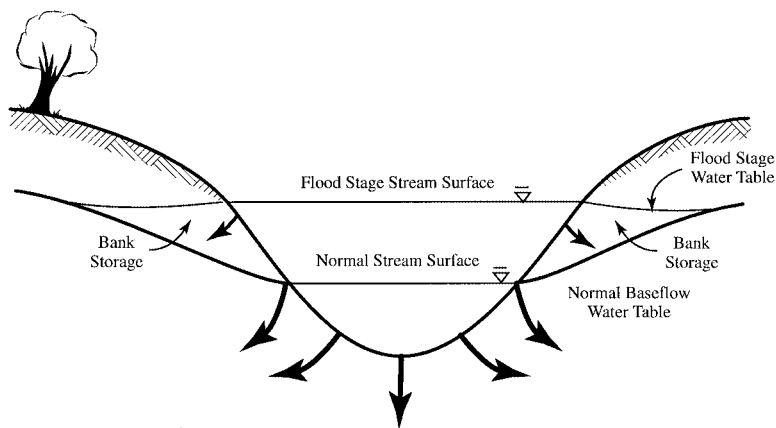

FIGURE 1.31 Bank storage.

out for some time, depending on soil permeability, the amount of precipitation, and the frequency of flooding.

1.7   Protection of Groundwater Supplies

With ever-increasing demands on our groundwater supplies, more and more thought is being given to their protection, but the water demands of growing populations, especially in large cities, are often in conflict with established and traditional uses such as farming, ranching, and industry. In the eastern U.S. where the precipitation may approach 180 cm/year, the problem is not as great as in the west where cities have been built in arid areas with precipitation of perhaps 15 cm/year. In these areas, not only has water from rivers in distant states been diverted (e.g., the Colorado River), but also extensive pumpage of groundwater has been taking place that withdraws water faster than it can be replenished by inflow or natural recharge. This is known as *groundwater mining.*

The concept of *safe yield* was evidently suggested early in this century (Lee, 1915). The definition put forward by Todd (1959) is "… the basin draft on a groundwater supply which can be continued indefinitely without harming the supply or basin landowners." This definition is basically the same as that given by most hydrologists, but the concept of "harm" has changed over the years.

Early in this century, harm meant overdraft of the aquifer and adverse effects on the supplies of adjacent landowners who tapped the same aquifer. It came to include the breaking of water laws eventually. In the last 20 years, it has included adverse effects on ecosystems, especially if pumping would cause a lake or wetland to lose water and thereby endanger certain species, or if profitable and popular recreational sites and activities were adversely affected.

The added concerns complicate the "safe yield" definition beyond the rather simple concern of overdraft which can be calculated with the Equation of Hydrologic Equilibrium (Todd, 1980), which is given as,

[Surface inflow + Subsurface inflow + Precipitation + Imported water

+ Decrease in surface storage + Decrease in groundwater storage]

= [Surface outflow + Subsurface outflow + Consumptive use + Exported water

+ Increase in surface storage + Increase in groundwater storage]   (59)

Another complicating factor is that the added concerns above are subject to temporal and locational interpretations and values; what may be a politically and socially acceptable definition today may change in 10 years with changes in demographics and land use. An example is a present problem in Park County, Colorado, where a proposed *conjunctive use* plan by the city of Aurora, Colorado, is meeting stiff opposition from residents of the county.

Conjunctive use is defined simply as the use of both groundwater and surface water conjunctively. The city of Aurora, approximately 100 miles (150 km) east of Park County, bought water rights from several ranchers in the county some years ago when ranching started to become less profitable. The city is expanding and now has proposed to recharge aquifers with surface water during periods of higher runoff, and then to pump the water out later and use it conjunctively with surface water from the same area. It appears that the amount to be pumped will exceed the recharge and groundwater mining will occur. Although the greater population of Aurora is in favor of the plan, the much smaller population of Park County is opposed.

This situation is only one of many such conflicts that have occurred and are still occurring in the western United States. Similar plans have been put forward by Colorado Springs, Colorado, and Las Vegas, Nevada. Again, the mining of vast quantities of water from remote areas would be necessary, even with conjunctive use to supply ever-expanding cities with no practical limits to growth. Therefore, safe

yield may be defined differently by developers and city planners than by landowners from whose area the water is to be taken.

1.7.1   Water Laws

Numerous laws have been passed since the founding of the United States that regulate the actual use of water. These vary from state to state and from east to west. A good overview of these laws can be found in Domenico and Schwartz (1990), Fetter (1994), Todd (1959), and Walton (1991).

Within the past two decades, the U.S. Congress has passed several laws aimed at protecting water supplies, making them safer to drink and more available to everyone; they include the *National Environmental Policy Act of 1969 (P.L. 91-190), Federal Water Pollution Control Act of 1972 (P.L. 92-500), Clean Water Act Amendments of 1977 (P.L. 95-217), Safe Drinking Water Act of 1974 [P.L. 93-523) and Amendments, Resource Conservation and Recovery Act of 1976 (RCRA — P.L. 94-580), Comprehensive Environmental Response, Compensation and Liability Act of 1980 (CERCLA — P.L. 96-510), Superfund Amendments and Reauthorization Act of 1986 (SARA), Surface Mining Control and Reclamation Act (SMCRA — P.L. 95-87), Uranium Mill Tailings and Control Act of 1978 (UMTRCA — P.L. 95-604 and later amendments], Toxic Substances Control Act [TOSCA — P.L. 94-469 and amendment], and Federal Insecticide, Fungicide and Rodenticide Act (FIFRA — P.L. 92-516 and later amendments).*

To many individuals and industries, this number of recently passed environmental laws seems onerous and restrictive. Yet the lack of such laws earlier in the history of the United States resulted in severe contamination of both surface water and groundwater from mining, industry, agriculture, and municipal and domestic sewage. For the most part, the laws are accomplishing a good deal of the protection and remediation for which they were written. For further details, see Chapter 22 on laws and regulations.

1.7.2   Natural Protective Barriers and Waste Disposal

Groundwater contamination is now a worldwide problem, especially in rapidly urbanizing areas, not only from trash and other wastes, but also from fertilization of lawns and gardens. For example, Cox et al. (1996) discuss the effects of urban growth on groundwater supplies of Australia; and Howard et al. (1996) discuss similar problems in Canada.

Much progress is being made today in the ability to map land for groundwater pollution vulnerability, and also to determine the areas that are protected by natural barriers, such as till and clay (Vrband and Zaporozec, 1994). Today, there is a greater emphasis on utilization of natural protective barriers for groundwater protection than ever before. Impermeable or low-permeability natural barriers are normally clay (glacial till and lacustrine sediments), shale, salt deposits, and unfractured granitic rocks and basalt.

If waste, hazardous or otherwise, can be placed in proper containers and placed so that a natural barrier is between it and a groundwater supply, the chances are good that the water will be protected, providing proper engineering methods are followed. Unfortunately, such practices have not always been the case. In the past, abandoned gravel pits and rock quarries were often used as disposal places of convenience. Today, many of them leak and have seriously contaminated groundwater resources under and around them.

An example is the Tippecanoe County Landfill north of Lafayette, Indiana, which was constructed approximately 25 years ago in a gravel pit in the glacial outwash sand and gravel bordering the Wabash River. A three-meter-thick layer of clay on the bottom of the pit was removed to deepen the depression to allow more volume for refuse. Thus, the natural barrier was removed and leachate seeped into the aquifer below.

The landfill was used to dispose of not only domestic trash, but also industrial chemicals, including polychlorinated biphenyls (PCBs). The EPA and the Indiana Department of Environmental Management forced the landfill to close and cease operations and declared it a CERCLA (Superfund) site. Remediation is expected to cost the Principally Responsible Parties (PRPs) and perhaps the citizens of the county at least $14 million, and the estimate continues to grow.

Another example in the same county, also on the outwash along the Wabash River, is the disposal of waste on the permeable surface. Sludge from municipal sewage processing combined with substrate sludge from myacin drug manufacturing has been disposed of for many years on the surface of the outwash. The outwash is perhaps 30 meters thick and has a hydraulic conductivity of more than 200 meters per day. The applied mixture is exceedingly rich in nitrate, phosphate, and potassium, which is welcomed by farmers as free fertilizer.

Since the soil is very permeable and retains little water, leachate from the sludge not only rapidly enters the soil, but also center-pivot irrigation from wells in the outwash is necessary to maintain adequate soil moisture. As a result, the nitrate level in the groundwater below the site and downgradient from the site is high, exceeding by a factor of four or five the maximum permissible concentration (MPC) of 45 mg/liter for drinking water set by the EPA. Not only is this water unhealthy to drink, but in addition, the property values of homes in the immediate area have dropped significantly, and perhaps as much as eight square kilometers of valuable and productive aquifer area is now seriously contaminated.

If the sludge were spread a few kilometers away on till 10 meters thick with a hydraulic conductivity of 10^{-7} cm/sec, not only would it serve as inexpensive fertilizer, but the natural barrier properties of the till would probably prevent groundwater contamination.

1.7.3 Artificial Protective Barriers

When natural barriers are not present, two kinds of artificial barriers can be constructed — solid and hydraulic. Extensive information about the use of such barriers can be found in EPA publications (1977, 1984).

Solid barriers may consist of grout curtains, slurry walls, or sheet piles constructed around a contaminated site such as a landfill to prevent the spread of contaminants to the surrounding groundwater. Figure 1.32 shows a slurry wall downgradient from a landfill which is blocking the movement of contaminants from moving downgradient at that location. This same figure can also demonstrate the utility of a sheet pile with interlocking pieces, or a grout curtain constructed at the same location.

A slurry wall is constructed by digging a trench down to an impermeable stratum and then filling it with a mixture of water, clay, and *bentonite*. The latter material is a clay mineral formed from the chemical alteration of volcanic ash and is often used as a sealant because when wet, it will expand eight to ten times its dry volume and is quite impermeable. If the depth to an impermeable stratum is too deep for construction of a slurry wall, a grout curtain can be constructed by drilling several wells close together in a line and then injecting a variety of sealing compounds into them which will infiltrate the soil pores to form an impermeable wall around the source of contamination.

A hydraulic barrier can be implemented by injecting water from a line of wells as illustrated in Figure 1.33. This kind of barrier is utilized to prevent the encroachment of saline water into freshwater aquifers along a seacoast and has been used successfully in Los Angeles, California, and in Long Island, New York. It is also employed to prevent the downgradient migration of contaminants from landfills.
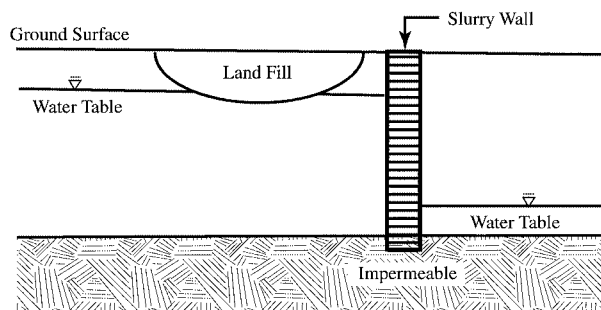

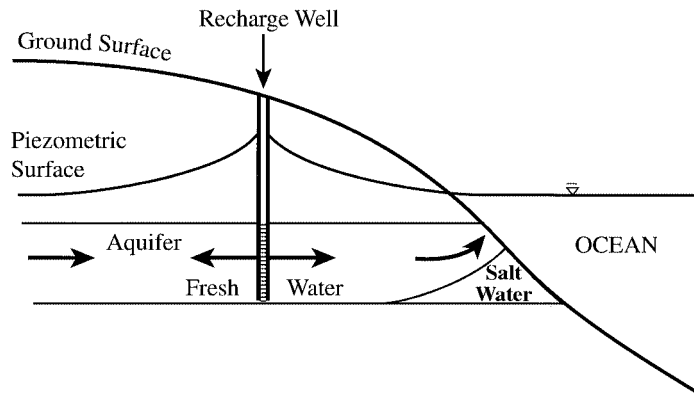**FIGURE 1.32** Slurry wall to prevent contaminant migration.

FIGURE 1.33 A hydraulic barrier.

1.7.4 Well-Head Protection Programs

Several state environmental agencies have implemented Well-Head Protection Programs under directives from the U.S. EPA. These programs are designed to protect the recharge areas of municipal and public wells from surface contamination by limiting the kinds of activities that can take place within a certain distance from a well or well field within the *well-head protection zone*. Generally, this zone is the same as the *capture zone* for a well or well field in an unconfined aquifer. The distance is computed as the distance of travel of water to the well within a given time limit. For example, the Indiana Well-Head Protection Program, currently under revision, requires that the minimum distance limit be set for a travel time of five years, but strongly encourages a ten-year limit.

The argument for setting travel-time limits is twofold. In the first place, an arbitrary distance limit would not make much sense without taking into account the geology, conductivity, and gradients specific to each site. Second, it is assumed that a five- or ten-year travel time will more likely guarantee that under most situations, any contaminants will have traveled far enough through the aquifer, and enough water from outside the contaminated zone will have been pumped that sorption and dilution, respectively, will have rendered them innocuous.

The consequence of this law will be restriction from the recharge areas, or at least the strong regulation, of many activities that generate waste or require chemical applications. These will include livestock feedlots, fertilization of farm fields and golf courses, spreading of sludge or nitrate-rich water from sewage processing, landfills, oil refineries, chemical plants, and others.

A simple capture zone is illustrated in Figure 1.34 and is computed analytically as (Todd, 1980),

$$x = -y \big/ \tan\big[2\pi K b i y \big/ Q\big] \tag{60}$$

where x and y are coordinate directions, Q is the pumping rate, K is the hydraulic conductivity, b is the initial saturated thickness of the unconfined aquifer, i is the hydraulic gradient, and the argument of the tangent is in radians. The units have been chosen so that the argument is dimensionless.

Bakker and Strack (1996) present a sophisticated analytical-element modeling approach to determining the capture zone under conditions of irregular and more complicated geology and boundary conditions.

Well-head protection plans have been in effect in The Netherlands for several years. The implementation of such rules in the U.S. has met with considerable opposition from some municipalities and well owners who realize that it will necessitate an extra expense. In order to adequately define the capture zone, there may be a need in some cases to perform test drilling and surface geophysical studies to determine the geology of the area. In addition, modeling studies may be necessary to completely understand the hydraulics of the aquifer. Small public water operations may find these studies to be financially
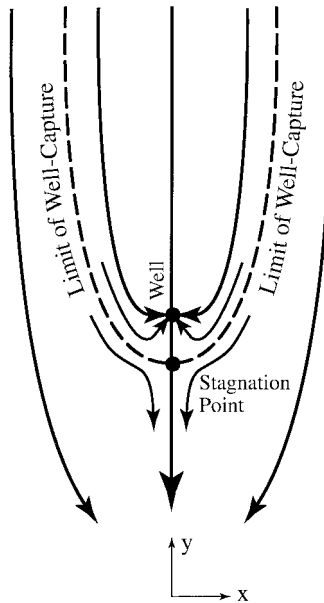
FIGURE 1.34 A capture zone around a pumping well.

burdensome. However, wellhead protection plans, when implemented, are expected to provide a high level of protection for public water supplies as the population grows and more demands are put on groundwater resources.

For further details see Chapters 25 and 27 on site remediation and geosynthetics.

1.8   Possible Effects of Climate Change on Groundwater Flow

It is known that the world's climate has changed significantly over geologic time. Evidence from fossil plants, pollen, rodent middens, tree rings, and even clay-mineral formation has shown that precipitation and temperature have varied widely (Rasmussen et al., 1993). For example, near the end of the Pleistocene Epoch (the Ice Age), the climate was much wetter in the American southwest than it is now, and the water table in southern Nevada was 100 meters higher than at present (Winograd and Doty, 1980). It is not entirely clear what caused climate change in the past, although volcanic eruptions and tilting of the earth's rotational axis have been suggested.

Many climatologists believe that anthropogenic factors are now effecting a slow, but perceptible change in the atmosphere that is causing gradually warming temperatures. These factors include changes in land use with overgrazing and burning of forests, and the introduction into the atmosphere of abnormal amounts of certain gases.

The gases, commonly known as *greenhouse gases*, include carbon dioxide (CO_2) and methane (CH_4) as the prime suspects in maintaining the *greenhouse effect*. This effect is the atmospheric entrapment of heat from the sun which, in turn, is expected to cause greater evapotranspiration in some areas and greater precipitation in others. As a result, the climate is expected to change significantly within the next century, causing a reduction of precipitation in the Midwest and a great increase of precipitation in the southwestern United States.

Although much research has been accomplished in estimating the potential effects of climate change on surface water supplies, little attention has been directed toward the possible effects on groundwater except to estimate the increase or decrease in recharge rates and amounts that can be expected. Recent work by Leap and Belmonte (1992), Leap (1993), and Reichard (1995) illustrates the role of increased pore pressure in the opening of fractures in rocks with consequent increase in hydraulic conductivity.

It has been known for some time that a slight increase in fracture width (aperture) can cause the flow rate through a fracture to increase by a power of three; this is known as the *Cubic Law* (Witherspoon, et al., 1980), and is given as

$$Q = \left(\rho g b^2 / 12\mu \right) \left(bw \right) \left(\partial h / \partial l \right) \tag{61}$$

where Q is the volumetric flow rate, ρ is the density of water, g is acceleration of gravity, b is aperture width, μ is dynamic viscosity of water, w is fracture width perpendicular to flow direction, h is the hydraulic head, and l is the length of the flow path through the fracture. The fracture is assumed to possess smooth sides, and laminar flow exists. A consistent set of units must be used.

Most rocks of aquifer quality may contain tens, hundreds, or even thousands of fractures per cubic meter, and in order to quantify flow through such rocks for practical purposes, it is more convenient to assume that the fracture density is great enough that the flow system can be treated as a porous medium. Jones (1975) performed a series of experiments measuring flow through cores of fractured carbonate rocks under different pressures in order to discover the relationship between intrinsic permeability and effective stress. The rock contained enough fractures that an equivalent porous medium could be assumed, and the relationship sought was determined to be

$$k/k_i = \left\{ \left[\log S_e - 4.602 \right] / \left[\log S_{ei} - 4.602 \right] \right\}^3 \tag{62}$$

where k is intrinsic permeability at any pressure, k_i is initial intrinsic permeability, S_e is effective stress at any pressure, and S_{ei} is initial effective stress.

Leap (1993) postulated that these relationships between permeability, pore pressure, and effective stress could be significant in understanding and predicting changes in permeability if the water table rose during a period of increased recharge consequent to increased precipitation.

Reichard (1995) used this equation as a basis for a modeling study to determine what the effect of a significant recharge-induced water-table rise would be on the hydraulic conductivity of an aquifer after a climate change which would cause significantly increased recharge. His studies showed that under such conditions, the hydraulic conductivity of fractured aquifers could increase as much as 15 to 30%. More research is needed in this area, but the results thus far are significant because they suggest that not only will flow rates be expected to increase, but transport rates of contaminants could also increase as well.

For Further Information

There are many journals and books about hydrology in general, including geology and engineering sources. However, the list below pertains mainly to hydrogeology.

Several very good college-level texts on the broad subject of hydrogeology exist, such as those listed in the **References** by Fetter, Freeze, and Cherry, Todd, and Domenico and Schwartz. These books cover a wide variety of subjects and do not specialize in any one subdiscipline.

Water Supply Papers and Professional Papers by the U.S. Geological Survey are continuously produced on more specialized subjects in hydrogeology and are available from the USGS National Center in Reston, Virginia.

Major journals in hydrology (including hydrogeology) include the following:
Water Resources Research, American Geophysical Union, Washington, DC.
Groundwater, National Association of Groundwater Scientists and Engineers, Dublin, Ohio.
Hydrogeology Journal, co-sponsored by the Geological Society of America in Boulder, Colorado, and the International Association of Hydrogeologists in Hanover, Germany.
Hydrological Science and Technology, American Institute of Hydrology, Minneapolis, Minnesota.

Bulletin of the Geological Society of America and *Geology,* both sponsored by the GSA, often contain papers on hydrogeology.

Special publications by the Geological Society of America and the American Geophysical Union are produced from time to time, specifically on hydrogeological subjects.

Professional Papers and Water Supply Papers are printed continuously by the US Geological Survey, Reston, Virginia, on specialized topics in hydrogeology. Similar publications are produced by the Canadian Geological Survey in Ottawa, and by the surveys of each of the provinces.

State Geological Surveys also produce results of studies of hydrogeology in each state.

References

Andrews, R.J., Barker, R., and Heng, L.M. 1995. The application of electrical tomography in the study of the unsaturated zone in chalk at three sites in Cambridgeshire, United Kingdom. *Hydrogeology Journal.* 3(4), 17-31.

Back, W., Rosenshein, J.S., and Seaber, P.R., eds., 1988. Hydrogeology. *The Geology of North America,* O-2, The Geological Society of America.

Bakker, M. and Strack, O.D.L. 1996. Capture zone delineation in two-dimensional groundwater flow models. *Water Resources Research.* 32(5), 1309-1315.

Bear, J. 1972. *Dynamics of Fluids in Porous Media.* American Elsevier Publishing Co., New York.

Bennet, G.D., 1979. Regional groundwater systems analysis. *Water Spectrum,* 11(4), 36-42.

Brooks, R.H. and Corey, A.T. 1964. Hydraulic properties of porous media, *Hydrology Paper* 3, Colorado State University, Fort Collins, CO.

Boussinesq, J. 1904. Recherches théoriques sur l'ecoulement des nappes d'eau infiltrées dan le sol et sur le débit des sources. *Journal de Mathématiques Pures et Appliquées,* 10, 5-78.

Campbell, G.S. 1974. A simple method for determining unsaturated conductivity from moisture retention data. *Soil Science,* 117, 311-314.

Chamberlin, T.C. 1885. *The Requisite and Qualifying Conditions of Artesian Wells.* U.S. Geological Survey 5th Annual Report, 131-173.

Cox, M.E., Hillier, J., Foster, L., and Ellis, R. 1996. Effects of a rapidly urbanising environment on groundwater, Brisbane, Queensland, Australia. *Hydrogeology Journal.* 4(1), 30-47.

Daniels, D.P., Fritz, S.J., and Leap, D.I. 1991. Estimating recharge rates through unsaturated glacial till by tritium tracing. *Ground Water.* 29(1), 26-34.

Darcy, H. 1856. *Les Fontaines Publiques de la Ville de Dijon.* Dalmont, Paris, France.

Darton, N.H. 1909. *Geology and Underground Waters of South Dakota.* U.S. Geological Survey Water Supply Paper 227.

Davis, S.N. and DeWiest, J.M. 1966. *Hydrogeology.* John Wiley & Sons, Inc., New York, NY.

De Marsily, G. 1986. *Quantitative Hydrogeology — Groundwater Hydrology for Engineers.* Academic Press. Orlando, FL.

Domenico, P.A. and Schwartz, F.W. 1990. *Physical and Chemical Hydrogeology.* John Wiley & Sons, Inc., New York.

Dupuit, J. 1863. *Études théoretiques et pratiques sur le mouvement des eaux dans les canaux découverts et à travers les terrains perméables.* 2nd ed., Dunoud, Paris.

Eggleston, J.R., Rojstaczer, S.A., and Peirce, J.J. 1996. Identification of hydraulic conductivity structure in sand and gravel aquifers: Cape Cod data set. *Water Resources Research.* 32(5), 1209-1222.

EPA 1977. *The Report to Congress; Waste Disposal Practices and Their Effects on Ground Water.* U.S. Environmental Protection Agency. Washington, DC.

EPA 1984. *A Ground Water Protection Strategy for the Environmental Protection Agency.* U.S. Environmental Protection Agency. Washington, DC.

Ferris, J.G., Knowles, D.B., Browne, R.H., and Stallman, R.W. 1962. *Theory of Aquifer Tests.* U.S. Geological Survey. Water-Supply Paper 1536E.

Fetter, C.W. 1994. *Applied Hydrogeology*, 3rd ed. Macmillan College Publishing Co. Inc., New York.

Forcheimer, P. 1914. *Hydraulik.* B.G. Teubner, Leipzig.

Freeze, R.A. and Cherry, J.A. 1979. *Groundwater.* Prentice-Hall, Inc., Englewood Cliffs, NJ.

Gutentag, E.D., Heimes, F.J., Krothe, N.C., Luckey, R.R., and Weeks, J.B. 1984. *Geohydrology of the High Plains Aquifer in Parts of Colorado, Kansas, Nebraska, New Mexico, Oklahoma, South Dakota, Texas and Wyoming.* U.S. Geological Survey Prof. Paper 1400-B.

Hantush, M.S. 1956. Analysis of data from pumping tests in leaky aquifers. Transactions American Geophysical Union, 37, 702-714.

Hazen, A. 1911. Discussion of "Dams on Soil Foundations." Transactions of American Society of Civil Engineers, 73, 199.

Heath, R.C. 1984. *Ground-Water Regions of the United States.* U.S. Geological Survey Water-Supply Paper 2242, 1-78. U.S. Government Printing Office, Washington, DC.

Heath, R.C. 1988. Hydrologic setting of regions, in *Hydrogeology, The Geology of North America, Vol. O-2,* eds. W. Back, J.S. Rosenschein, and P.R. Seaber, 15-23, The Geological Society of America, Inc., Boulder, CO.

Hem, J.D. 1985. *Study and Interpretation of the Chemical Characteristics of Natural Water.* U.S. Geological Survey Water-Supply Paper 2254, U.S. Government Printing Office, Washington, DC.

Horton, R.E. 1933. The role of infiltration in the hydrologic cycle. *Transactions, American Geophysical Union,* 14, 446-60.

Howard, K.W.F., Eyles, N., and Livingstone, S. 1996. Municipal landfilling practice and its impact on grounwater resources in and around urban Toronto, Canada. *Hydrogeology Journal.* 4(1), 64-79.

Hubbert, M.K. 1940. The theory of ground-water movement. *Jour. of Geology,* 48(8), 785-944.

Hurlbut, C.S. 1970. *Minerals and Man.* Random House Inc., New York, NY.

Jacob, C.E. 1940. On the flow of water in an elastic artesian aquifer. Transactions American Geophysical Union. 22, 574-586.

Jones, F.O. 1975. A laboratory study of the effects of confining pressure on fracture flow and storage capacity in carbonate rocks. *Jour. of Petroleum Technology.* 27, 21-27.

Kanivetsky, R. and Rumynin, V.G. 1993. Determination of recharge rates to a glacial aquifer system in Minnesota using environmental tritium. *Hydrological Science and Technology.* (1-4), 62-73.

King, F.H. 1899. *Principles and Conditions of the Movement of Groundwater.* U.S. Geological Survey 19th Annual Report, part 2, 59-294.

Leap. D.I. 1986. *The Geology and Hydrology of Day County, South Dakota.* Bulletin 24, South Dakota State Geological Survey.

Leap, D.I. 1993. Potential increases in contaminant-transport rates from increased aquifer pore pressure following increased precipitation, in *Proceedings of 2nd UAS/CIS Joint Conference on Environmental Hydrology and Hydrogeology, "Impact of Environmental and Climatic Change on Glogal and Regional Hydrology."* 73-91. Washington, DC.

Leap, D.I. and Belmonte, P.M. 1992. Influence of pore pressure on apparent dispersivity of a fissured dolomitic aquifer. *Ground Water.* 30(1), 87-95.

Lee, C.H. 1915. The determination of safe yield of underground reservoirs of the closed basin type. *Transactions, American Society of Civil Engineers.* 78, 148-151.

Linsley, R.K., Jr., Kohler, M.A., and Paulhus, J.L.H. 1975. *Hydrology for Engineers.* McGraw-Hill, Inc., New York, NY.

Lohman, S.W. 1972. *Ground-Water Hydraulics.* U.S. Geological Survey Prof. Paper 708, US Government. Printing Office, Washington, DC.

Maidment, D.R. 1993. Hydrology, in Maidment, D. R., ed. *Handbook of Hydrology,* 1.1-1.15. McGraw-Hill, Inc., New York, NY.

Marsily, de, G. 1986. *Quantitative Hydrogeology,* Academic Press, New York, NY.

Mavis, F.T. and Tsui, T.P. 1939. Percolation and capillary movements of water through sand prisms. *Bull. 18,* University of Iowa Studies in Engineering, Iowa City, Iowa.

Meinzer, O.E. 1923. *The Occurrence of Groundwater in the United States with a Discussion of Principles.* U.S. Geological Survey Water Supply Paper 489.

Meinzer, O.E. 1939. *Groundwater in the United States.* U.S. Geological Survey Water Supply Paper 836-D, 157-232.

Mendenhall, W.C. 1905. *The Hydrology of San Bernardino Valley, California.* U.S. Geological Survey Water-Supply and Irrigation Paper 142, 124 p.

Meyboom, P. 1967. Mass-transfer studies to determine the ground-water regime of permanent lakes in hummocky moraine of western Canada. *Jour. of Hydrology.* 5, 117-142.

Narasimhan, T.N. and Goyal, K.P. 1984. Subsidence due to geothermal fluid withdrawal, in *Man-Induced Land Subsidence,* Reviews in Engineering Geology, v. VI, Geological Society of America, 35-66.

Pilgrim, D.H. and Cordery, I. 1993. Flood Runoff, in *Handbook of Hydrology,* ed. D.R. Maidment. 9.1-9.42. McGraw-Hill, Inc., New York, NY.

Rasmussen, E.M., Dickinson, R.E., Kutzbach, J.E., and Cleaveland, M.K. 1993. Climatology, in *Handbook of Hydrology,* ed. D.R. Maidment, 2.1-2.44. McGraw-Hill, Inc., New York, NY.

Rawls, W.J., Ahuja, L.R., Brakensiak, and Shirmohammadi, A. 1993. Infiltration and soil water movement, in *Handbook of Hydrology,* D.R., Maidment, ed., 5.1-5.51.

Reichard, J.S. 1995. *Modeling the Effects of Climatic Change on Groundwater Flow and Solute Transport.* Ph.D. Dissertation, Department of Earth and Atmospheric Sciences, Purdue University, West Lafayette, IN.

Reichard, J.S. and Leap, D.I. (in press). The effects of pore-pressure changes on fractured aquifers. *Ground Water,* 7pp.

Sun, R.J. ed., 1986. *Regional Aquifer Systems Analysis Program of the U.S. Geological Survey, Summary of Projects, 1978-84.* U.S. Geological Survey Circular 1002, 264 p.

Theis, C.V. 1935. The relation between the lowering of the piezometric surface and rate and duration of discharge of a well using groundwater storage. Transactions of American Geophysical Union, 2, 519-524.

Thiem, G. 1906. *Hydrologische Methoden.* 56, Gebhardt, Leipzig.

Todd, D.K. 1959. *Hydrology.* 1st ed. John Wiley & Sons, Inc., New York, NY.

Todd, D.K. 1980. *Ground Water Hydrology.* 2nd ed. John Wiley & Sons, Inc., New York, NY.

Tolman, C.F. 1937. *Ground Water.* McGraw-Hill Book Co., Inc., New York, NY.

Tóth, J.A. 1962. A theory of ground-water motion in small drainage basins in central Alberta, Canada. *Jour. of Geophysical Research,* 67(11), 4375-87.

UNESCO. 1992. *International Glossary of Hydrology.* 2nd ed., WMO Rept. No 385, World Meteorological Organization, Geneva, Switzerland.

USDA 1955. *United States Department of Agriculture Year Book, 1955.* U.S. Government Printing Office, Washington, DC.

Van Genuchten, M. Th. 1980. A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Soc. of America Journal,* 44, 892-898.

Vrband, J. and Zaporozec, A. 1994, eds. *Guidebook on Mapping Groundwater Vulnerability.* International Contributions to Hydrogeology; International Association of Hydrogeologists, 16.

Walton, W.C. 1991. *Principles of Groundwater Engineering.* Lewis Publishers, Inc., Chelsea, MI.

White, W.N. 1932. *A Method of Estimating Ground-Water Supplies Based on Discharge by Plants and Evaporation from Soil.* U.S. Geological Survey Water-Supply Paper 659, 1-105, U.S. Government Printing Office, Washington, DC.

Winograd, I.J. and Doty, G.C. 1980. *Paleohydrology of the Southern Great Basin, With Special Reference to Water Table Fluctuations Beneath the Nevada Test Site During the Late(?) Pleistocene.* U.S. Geological Survey Open File Report 89-569. U.S. Government Printing Office, Washington, DC.

Winter, T.C. 1976. *Numerical Simulation Analysis of the Interaction of Lakes and Ground Water.* U.S. Geological Survey Prof. Paper 1001. U.S. Government Printing Office, Washington, DC.

Witherspoon, P.A., Wang, J.S.Y., Iwai, K., and Gale, J.E. 1980. Validity of the cubic law for fluid flow in a deformable rock fracture. *Water Resources Research.* 16(6), 1016-1024.

Glossary

Adhesion The attraction of water to solid grains by electrostatic forces.

Aeolian Deposits Materials deposited from the air and moved about by wind (e.g., dune sand).

Alluvium Any lithology or grains size of sediment deposited by flowing streams.

Anisotropy The condition of a porous medium where the permeability is different in different directions.

Aquiclude An impermeable geological formation or body that will prevent the flow of water through it.

Aquifer A permeable geological formation or body that will yield water in economical amounts.

Aquitard A semipermeable geological formation or body that will retard the flow of water through it.

Artesian Aquifer A confined aquifer in which the piezometric surface rises above the top of the aquifer.

Bank Storage Water stored in a stream bank when the stream water level is higher than the adjacent water table, as during a period of high water.

Baseflow The flow of water from the groundwater system to a surface stream or lake that maintains water in the surface body, even between precipitation events.

Cable-Tool or Churn Drilling Drilling with a bit like a large chisel on the end of a cable. It is rapidly raised and lowered to chip a hole in unconsolidated soil or in rock.

Capillary Fringe A zone above the water table where the soil is saturated but under tension (pressure less than atmospheric), as opposed to the zone below the water table where water is under pressure.

Capillary Water or Pellicular Water The water actually held in the soil or rock pores after drainage.

Capture Zone The area around a well or well field from which water will move to the discharge point(s).

Characteristic Curves Curves showing hydraulic conductivity and soil moisture content as functions of the soil pressure head; a set of such curves is "characteristic" of a particular soil.

Confined Aquifer An aquifer overlain by an impermeable layer such that the piezometric head rises above the top of the aquifer.

Darcy's Law The relationship discovered by Henri Darcy between flow rate, hydraulic conductivity, and gradient in a porous medium.

Dynamic Viscosity The measure of the resistance of a fluid to shear.

Effective Porosity The porosity of a rock or soil that is actually connected to provide flow through the rock or soil.

Fluid Potential The potential energy per unit mass of water.

Gaining Stream A stream in which the water surface is above the water table, thus causing water to flow from the stream to the groundwater system below.

Glacial Drift The general term for any and all earth materials deposited from the action of glacial ice or their meltwaters.

Glacial Outwash Material deposited from meltwater from glaciers.

Glacial Spillway Generally a river valley which served as a major outlet for sediment-laden water from melting glaciers (e.g., the Mississippi, Missouri, Big Sioux, and Wabash River valleys).

Glacial Till A mixture of clay, silt, sand, gravel, and boulders deposited in a heterogeneous mass directly from glacial ice through melting or plastering.

Gravitational Water In draining soil, it is the water that actually drains out.

Greenhouse Gases Gases that are believed to trap the escape of heat from the earth and therefore cause warming of the world's climate. The most important of these is carbon dioxide (CO_2), but methane (CH_4) is also of concern.

Groundwater Water at or below the water table in earth materials.

Groundwater Mining Removal of groundwater at a rate exceeding recharge.

Heterogeneous The condition of a porous medium where the isotropic or anisotropic permeability is different in different parts of the medium.

Homogeneous The condition of a porous medium where the isotropic or anisotropic permeability is the same in all parts of the medium.

Hydraulic Conductivity The ease with which a fluid will flow through a porous medium. It is a function of the pore size and fluid properties of viscosity and density.

Hydraulic Head An expression of potential energy of a fluid at a point expressed as a length. It includes the height of the point above a datum plane (normally mean sea level) plus the pressure of the column of water above the point expressed as the pressure-equivalent height of that water column.

Hydraulic-Rotary Drilling Method A drilling method in which a bit on a long shaft, the drill stem, is rotated to drill a hole in the ground. Water, often mixed with clay or other substances, is pumped down the inside of the stem and forced out at the bit to lubricate the bit and to flush cuttings to the surface.

Hydrograph A graph of flow rate of a stream versus time. It also refers to a graph of the water level in a well versus time.

Hygroscopic Coefficient The maximum water content that can be held hygroscopically.

Hygroscopic Water That water held in a thin film around grains. It is not available to plants.

Igneous Rock A rock type that has solidified from the molten state, e.g., basalt, granite.

Induration The process by which mineral grains become fused or cemented together to form a solid rock mass, as opposed to unconsolidated materials.

Infiltration Capacity The maximum rate at which a soil will allow water to infiltrate it from the surface. It is dependent upon the initial moisture content of the soil, the vertical hydraulic gradient, and the rate of precipitation.

Interflow or Subsurface Stormflow The flow of water between the surface permeable zone and a zone of low permeability below it.

Isotropic The condition of a porous medium where the permeability is the same in all directions.

Juvenile Water Water created by chemical reactions in volcanic and other geothermal activities.

Karst A term describing an area of carbonate rocks which contains solution tunnels and caves. This is named after the Karst region of Yugoslavia.

Lacustrine Sediments Generally fine-grained silts and clays deposited from still waters in lakes or ponds.

Loess "Rock flour," or fine silt deposited originally in wide outwash deposits, but later deposited down-wind by the prevailing winds.

Losing Stream A stream in which the water surface is lower than the adjacent water table, thus causing water to flow from the groundwater system into the stream.

Metamorphic Rock A rock type that was formed from either sedimentary rocks (e.g., limestone) or igneous rocks (e.g., granite) when they were subjected to heat and pressure to produce a totally different rock from the parent, but with similar chemical composition (e.g., marble or gneiss, respectively).

Mineral A naturally occurring, nonorganic substance, generally of crystalline form, that has a definite chemical composition or a narrow range of compositions.

Permeability The property of a porous medium to transmit water. It is a function of pore diameter.

Phreatic Water (or soil water) That water above the capillary fringe that is under tension.

Pleistocene Epoch A geological span of time beginning approximately 2 million years ago during which four major glacial advances (Nebraskan, Kansan, Illinoisan and Wisconsinian) advanced over North America. The latest, the Wisconsinian, is contemporaneous with the Wurm advance of Europe.

Pore Velocity or Average Linear Velocity The actual average velocity of flow through a porous medium. It is calculated as the specific discharge divided by the effective porosity.

Porosity The percent ratio of void volume to total volume of a rock or soil.

Precipitates Minerals such as calcite ($CaCO_3$) that are deposited in water from direct precipitation from the water.

Primary Porosity The porosity characteristic of a rock or soil when first formed.

Pyroclastic A rock type formed by the settling of sediment which in turn is the ash and cinders of igneous material from volcanic eruptions.

Reynold's Number The ratio of accelerative forces to viscous forces which determines if a flow is laminar or turbulent.

Rock A naturally occurring stony entity made up of one or more minerals.

Safe Yield The rate of withdrawal of groundwater from an aquifer such that no harm will be caused. The term "harm" is relative to the economics of the area and time.

Saturation Ratio The ratio of contained water in pores to the volume of the pores.

Secondary Porosity The porosity that is added to a rock or soil after formation; it can include faults, fissures, fractures, etc.

Sedimentary Rock A rock formed from the weathered products (detritus) of igneous, metamorphic, or other sedimentary rocks (e.g., sandstone, siltstone, shale).

Seepage Face The zone of wet soil just above the surface of a stream or lake which marks the atmospheric discharge of water from the capillary fringe.

Semi-Confined Aquifer An aquifer overlain or underlain by a semipermeable aquitard which will allow limited flow of water to pass through.

Specific Discharge The volumetric flow rate per unit area of porous media perpendicular to the flow direction. It is sometimes called "Darcian Velocity."

Specific Retention The ratio of the volume of water held by the soil to the total matrix volume.

Specific Storage The volume of water that will be obtained from a unit volume of aquifer upon release of a unit value of head.

Specific Yield The ratio of the volume of water drained from a soil to the total matrix volume.

Spring A natural discharge point for groundwater to the surface.

Storage Coefficient The specific storage times the saturated thickness of the aquifer.

Unconfined Aquifer An aquifer that has no overlying confining impermeable layer.

Wilting Point The moisture content of a soil at which capillary forces become greater than osmotic suction of plant roots, causing the plants to wilt.

2

Elementary Groundwater Flow and Transport Processes

Jacques W. Delleur
Purdue University

2.1   Introduction

Essentially all natural groundwater flows are three-dimensional. That is, the velocity of a percolating water particle is represented by a vector that has three components. A simple example is the three-dimensional radial flow toward a well. (See Chapters 8 and 9.) However, there are many situations in which the velocities are nearly coplanar or there is radial symmetry. In these cases, the flow can be analyzed as two-dimensional with accuracy sufficient for many engineering problems. An example is the infiltration of water into a series of long parallel horizontal tile drains. Away from the pipe extremities, the shape of the water table is independent of the location along the pipe drain. In some cases the flow problem can be further reduced to one dimension. For example, the flow in karst conduit can be regarded as approximately one-dimensional. The existence of symmetry or special assumptions permits the simplification of many problems. It is important, however, to recognize the size of the errors that such simplifications can entail.

Groundwater flow variables, such as velocity and pressure, can vary in time or can be independent of time. For example, when the pumping of a well is started, the water table drawdown increases with time. This is an unsteady or transient flow problem. If the flow variables do not change with time, the flow is steady.

This chapter is concerned with elementary one- and two-dimensional steady groundwater flow and transport problems. Advanced three-dimensional flows, transient flows, transport processes, and modeling are treated in subsequent chapters. The geological setting is discussed in Chapter 1. Chapters 1 and 2 can be regarded as an introduction to groundwater hydrogeology and engineering. The following chapters deal in more detail with a number of advanced or specialized subjects.

The notions of saturated zone, confined aquifer, unconfined aquifer, water table, aquitard, aquiclude, perched aquifer, unsaturated zone or vadose zone, and the physical properties of aquifers are discussed in the first chapter. This chapter starts directly with the discussion of the hydraulics of groundwater or the motion of water below ground. The second part of this chapter deals with the transport of contaminants by groundwater.

2.2   Pressure, Suction, Piezometric Head, and Hydraulic Gradient

Water pressures are generally expressed as *gage pressures*, but they can also be expressed as *absolute pressures.* These pressures are related by the equation (Figure 2.1)

$$\text{Absolute pressure = Local atmospheric pressure + Gage pressure} \qquad (1)$$

Gage pressures are used in the following discussion. The water table is at the local atmospheric pressure and serves as a datum for gage pressures. Point A in Figure 2.1 is in the saturated zone, and the gage pressure is positive and is called the *pore pressure.* Point B is in the unsaturated zone, and the gage pressure is negative. This negative pressure is referred to as a *suction* or *tension.* The suction is expressed as a positive number. Thus a positive suction corresponds to a negative gage pressure. The dimensions of pressure are F/L^2, that is, N/m^2 or pascal (Pa), kN/m^2 or kilopascal (kPa) in SI units and lb/in^2 and lb/ft^2 in U.S. units. The law of hydrostatics states that the pressure, p, can be expressed in terms of the height of liquid, h_p, measured from the water table (assuming the groundwater at rest or moving horizontally). This height is called the *pressure head*

$$h_p = p/\gamma_w = p/\rho g \qquad (2)$$

where $\gamma_w = \rho g$ is the specific weight and ρ is the density of the water. For point A, the quantity h_p is positive, while it is negative for point B. The pressure head is generally expressed in meters of water, but
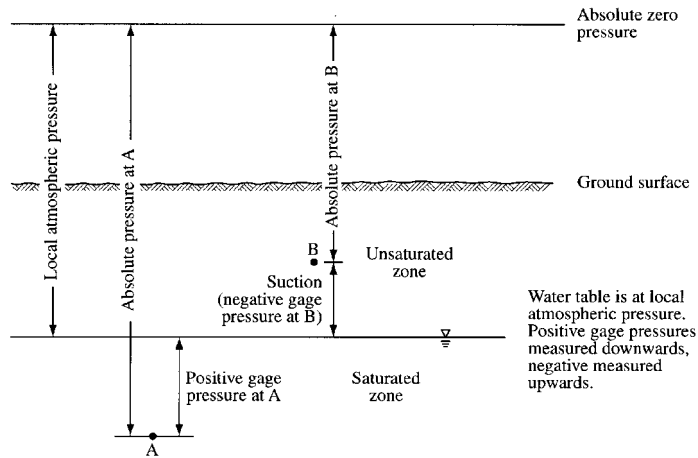
FIGURE 2.1 Absolute and gage pressures.

TABLE 2.1 Conversion Factors for Pressures and Related Terms

To convert from	to	Multiply by
atmosphere	pascal	1.013 E+5
bar	pascal	1.000 E+5
dyne/cm^2	pascal	0.100
ft of water (39.4°F)	pascal	2.989 E+3
inch of mercury (32°F)	pascal	3.386 E+3
inches of water (39.4°F)	pascal	249.1
millibar	pascal	100
millimeters of mercury (0°C)	pascal	133.3
pound per sq. foot	pascal	47.88

Source: Système International d'Unités, Universities Council on Water Resources, 1976.

TABLE 2.2 Approximate Equivalents of Atmospheric Pressure

US	Metric
14.7 psi abs.	101.3 kN/m^2, abs
	1 013 mbars, abs
29.9 in.Hg	760 mm Hg
	0.76 m Hg
33.9 ft. H$_2$O	10.3 m H$_2$O

it can also be expressed in cm of mercury. Some conversion factors for pressure and pressure heads are listed in Table 2.1 and some approximate equivalents of atmospheric pressure are listed in Table 2.2.

If the medium is saturated, the pore pressure, p, can be measured by the pressure head, $h_p = p/\gamma_w$, in a *piezometer*, a non-flowing well. The difference between the altitude of the well, H, (Figure 2.2) and the depth to the water inside the well is the *total head*, h_t, at the well. In fluid mechanics the total head is defined as the sum of the *elevation head*, z, the *pressure head*, p/γ_w, and the *velocity head*, $v^2/2g$, where v is the flow velocity and g is the acceleration of gravity. For groundwater flow the velocity head can generally be ignored because the water moves very slowly. Therefore, the total head at an observation well is taken to be equal to the *piezometric head*, or the sum of the elevation head and the pressure head. The symbol ψ is sometimes used to designate the pressure head.
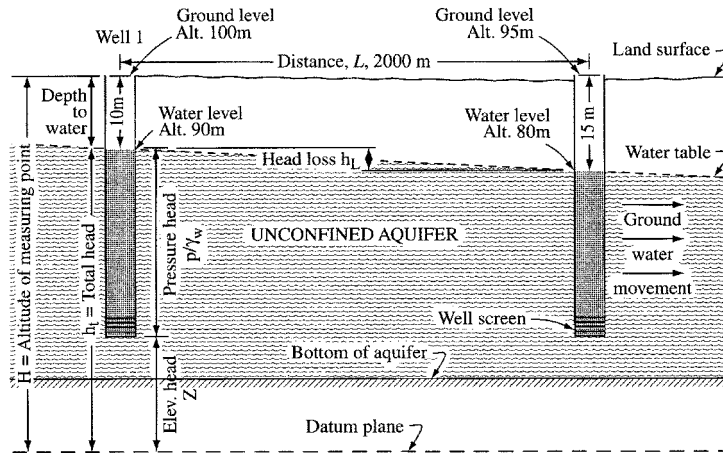
FIGURE 2.2 Heads and gradients. (Adapted from Heath, R. C. 1995. *Basic Ground-Water Hydrology.* U.S. Geological Survey, Water-Supply paper 2200, seventh printing, Denver, CO.)

$$h_t = z + p/\gamma_w = z + \psi \tag{3}$$

The piezometric head is also referred to as the *piezometric potential.* The change of piezometric head per unit distance in a given direction is the *hydraulic gradient.* If the direction is not specified, it is assumed to be in the direction of the maximum gradient. The hydraulic gradient is a dimensionless quantity (L/L) when consistent units are used. (See also Chapter 1, Section 1.4.7.)

Example 1: With the data from Figure 2.2, find the hydraulic gradient**.**

Solution: The hydraulic gradient is: $h_L/L = [(100 - 10) - (95 - 15)]/2000 = (90 - 80)/2000 = 0.005$.

If the piezometric head is known at three observation wells A, B, and C that are not in a straight line, then both the direction of the groundwater flow and the hydraulic gradient can be calculated approximately. For this purpose, with reference to Figure 2.3, (1) select observation well C with the intermediate head. (2) By linear interpolation, find location of the point P having the intermediate head on the line connecting observation wells A and B with the maximum and minimum heads. (3) Connect point P with the intermediate well; this line is a segment of a piezometric contour line. (4) Draw a line perpendicular to this contour that passes either through the maximum head or the minimum head observation well. This line is in the direction of the groundwater movement. (5) The ratio of the head differential between the ends of the perpendicular line and the length of this line is the hydraulic gradient.

Example 2: Using the data of Figure 2.3, find the flow direction and the hydraulic gradient.

Solution: (1) Select well C with the intermediate head of 25.17 m. (2) Find the distance x from well A to point P: $x = (25.30 - 25.17)*230.5/(25.30 - 25.00) = 99.88$ m. (3) Draw the line BD perpendicular to CP; this is the direction of the groundwater movement. Measure the distance BD = 114.29 m. (4) Divide the head differential by the length BD to obtain the hydraulic gradient $(25.17 - 25.00)/114.29 = 0.0015$.

If the medium is unsaturated, the negative pressure or *suction* or *tension* is measured by a *tensiometer.* This instrument is composed of a vertical tube closed at the top and a porous ceramic cup at the bottom (Figure 2.4). The tensiometer is initially filled and the ceramic cup is saturated with water. When the instrument is placed in the soil, the water in the tensiometer is generally at atmospheric pressure. The soil water, which is at negative pressure, produces a suction which drains water from the tensiometer, causing a pressure drop. When equilibrium is reached, the pressure inside the tensiometer is equal to that in the soil and can be measured with a vacuum gage. The practical suction range of a tensiometer
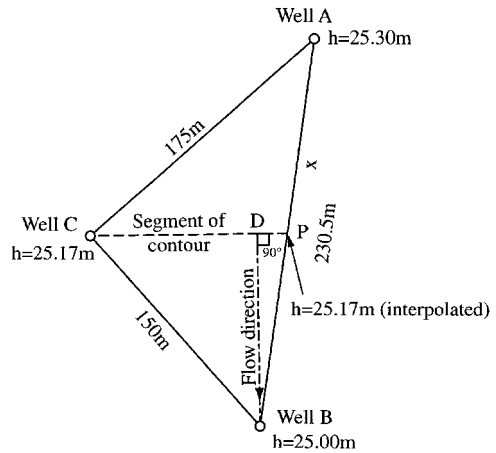
FIGURE 2.3 Finding flow direction and gradient from three observation wells. (Adapted from Heath, R. C. 1995. *Basic Ground-Water Hydrology*. U.S. Geological Survey, Water-Supply paper 2200, seventh printing, Denver, CO.)
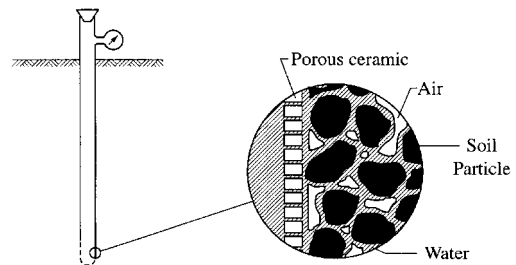
FIGURE 2.4 Tensiometer. (From Bouwer, H. 1978. *Groundwater Hydrology*. McGraw-Hill Inc., New York. With permission.)

is 0 to 0.8 bar or approximately 0 to 8 m of water. Relationships between the soil moisture content θ and the pressure head ψ are presented in Chapters 1, 5 and 6.

The *total pressure* at a point in a porous medium is the weight per unit area of the overburden above this point. This total pressure is the sum of the pore pressure and the *intergranular stress*, i.e., the stress due to forces transmitted from grain to grain in the rock matrix.

2.3 The Motion of Groundwater

The motion of water requires energy. This energy can be expressed as a *head* above a *datum*. The elevation of this datum is arbitrary. This is because the difference in energy or the difference in head is the concern. It is therefore important that the energies be measured with respect to the *same* datum. In groundwater engineering, mean sea level (MSL) is usually taken as the datum. The *hydraulic head* is defined as the energy per unit weight measured relative to the datum.

Water can possess several forms of energy. Perhaps the most obvious is the energy water possesses by virtue of its elevation above the datum. This is the *potential energy*. A mass m of water at an elevation z above the datum has a potential energy mgz, where g is the acceleration of gravity. This is the work necessary to move the mass m from the datum to the elevation z. If ρ is the density of the water, a unit volume of water has a mass ρ and a weight ρg and a potential energy $\rho g z$. The potential energy per unit weight, that is the elevation head, is thus $\rho g z / \rho g = z$. Note that the head has the unit of length.

The energy that water possesses by virtue of its motion is the *kinetic energy*. A mass m of water that moves with a velocity *v* has a kinetic energy $^{1}/_{2}mv^{2}$. Thus the kinetic energy per unit mass is $^{1}/_{2}\rho v^{2}$ and the kinetic energy per unit weight or *velocity head* is $^{1}/_{2}\rho v^{2}/\rho g = v^{2}/2g$. The velocity head has the dimension of length. When groundwater is flowing through the pores of the rock or soil formation, the velocity is very small, perhaps of the order of centimeters per year, and the velocity head is usually negligible with respect to other forms of energy. One exception is near wells, where the velocity increases significantly.
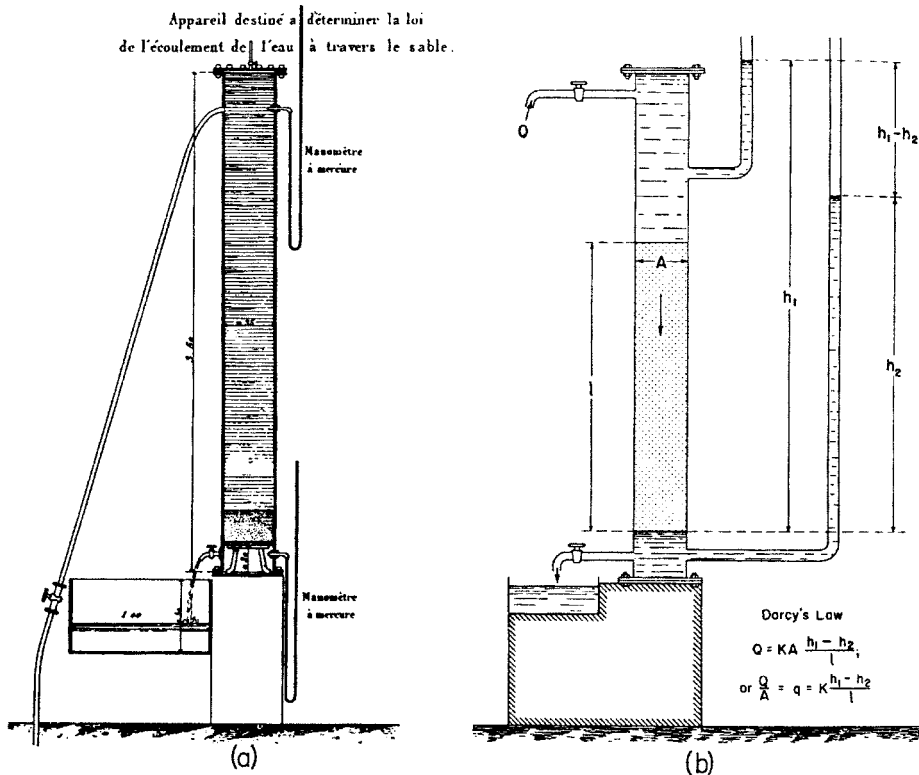
FIGURE 2.5 Darcy's original apparatus with mercury manometer (a) and equivalent apparatus with water manometers (b). (From Hubbert, M. K. 1953. Entrapment of petroleum under hydrodynamic conditions. *Bulletin of American Association of Petroleum Geologists*, Vol. 37, 1954-2026. With permission.)

Another exception is in certain karst conduits where groundwater can flow fast enough that the velocity head is important.

The energy that water possesses by virtue of its pressure is the *pressure energy*. The pressure intensity of the fluid, p, acting on an area dA produces a force p dA. If the area is displaced by a distance ds, in the flow direction, then the force produces an amount of work p dA ds, known as *flow work*. The volume dA ds has a weight ρg dA ds and the flow work per unit weight is p dA ds/ρg dA ds = p/ρg known as the pressure head. The sum of the elevation head and the pressure head is known as the piezometric head h = z + p/ρg.

2.4 Flow Through Porous Media — Darcy's Law

The French engineer, Henry Darcy, performed experiments on the filtration of water through sand columns. His finding that the rate of flow through a sand column is proportional to the loss of head appeared in an appendix to his treatise on the public fountains of the city of Dijon (Darcy, 1856). Figure 2.5 shows the original set-up utilized by Darcy, and Figure 2.6 shows some of his experimental results as plotted by Hubbert (1953) from Darcy's data. Darcy's law states that the volumetric flow rate, Q, $[L^3 T^{-1}]$ across a gross area A of a formation with a hydraulic conductivity K $[LT^{-1}]$ (definition and values are given in Chapter 1), under a hydraulic gradient i = $-\partial h/\partial s$ in the s direction is given by
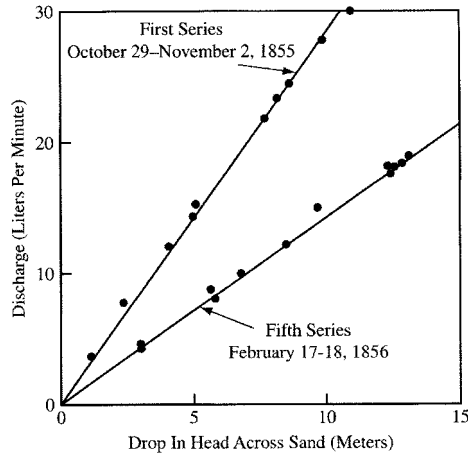
FIGURE 2.6 Darcy's data plotted by Hubbert. (From Hubbert, M. K. 1953. Entrapment of petroleum under hydrodynamic conditions. *Bulletin of American Association of Petroleum Geologists*, Vol. 37, 1954-2026. With permission.)

$$Q = qA = -KA\,\partial h/\partial s = KAi = \frac{k\rho g}{\mu}\,Ai \tag{4}$$

where q is a conceptual velocity called the *specific discharge* or flow rate per unit area $[LT^{-1}]$ also known as the *Darcy velocity*, μ is the viscosity, and k is the *permeability*. The hydraulic head, h, is the sum of the elevation head z and the pressure head p/γ_w. The minus sign in the above equation indicates that the flow takes place from high to low head, namely in the direction of decreasing head. The pore velocity, v $= q/n_e$, where n_e is the effective porosity and v is the average flow velocity in the pores, usually called the *seepage velocity*. It is the average velocity for the transport of solutes that are nonreacting. In the boundary determination of contributing areas for *well head protection zones,* it is often necessary to find the time it takes for the water to move from a point to a bore hole. This can be done using the pore velocity. Pollutants that travel primarily by advection would move at the same velocity as the water, but those that are subject to the effects of diffusion and adsorption move more slowly. These effects are discussed in Section 2.13 and in more detail in Chapters 14 and 15, concerned with transport processes.

The one-dimensional form of Darcy's law is

$$q = K\,\frac{\left(p_1/\gamma_w + z_1\right) - \left(p_2/\gamma_w + z_2\right)}{L} \tag{5}$$

where the subscripts 1 and 2 refer to the points at which the pressure heads and the elevation heads are considered, and *L* is the distance between these points.

Example 3: Find the hydraulic conductivity of the sands used in Darcy's first series of experiments (Figure 2.6), assuming that the height of the sand column is 3 m and the diameter of the stand pipe is 0.35 m.

Solution: Take the flow rate Q = 30 l/min = 0.03 m^3/min. The specific discharge is q = Q/A = 0.03/((π * 0.35^2/4) = 0.312 m/min. From Equation (5) K = qL/Δh. From the graph Δh = 10.5 m. Thus K= 0.312 * 3.0/10.5 = 0.089 m/min or 0.0015 m/sec. This corresponds to a coarse sand.

Example 4: Find the time it takes for a molecule of water to move from a factory to a bore hole located 4 km away in a homogeneous silty sand unconfined aquifer with a hydraulic conductivity of K = 5 ×

10^{-5} m/s or 4.32 m/d, an effective porosity of 0.4, and observing that the water table drops 12 m from the factory to the bore hole.

Solution: Since $v = q/n_e = Ki/n_e$ the pore velocity is calculated as

$$v = \frac{4.32 * \left(12 / 4000\right)}{0.4} = 0.0324 \text{ m/d}$$

It would take $4000/(0.0324*365) \approx 338$ years. If instead, the aquifer was a fractured limestone with a porosity of 0.01 and the same hydraulic conductivity, the pore velocity would be approximately 1.3 m/d and the time to travel the 4 km would be 8.5 years. With a porosity of 0.001, the travel time would reduce to 0.85 years. Pumping at the bore hole will increase the hydraulic gradient and increase the pore velocity and thus decrease the travel time.

2.4.1   Similarity of Darcy's Law and Other Laws of Physics

Darcy's law is similar to Fourier's law of heat transfer, Ohm's law of electricity, and Fick's law of solute diffusion. *Fourier's law* governs the conduction of heat from high temperatures to low temperatures. It states that the heat flux is proportional to the temperature gradient and the constant of proportionality is the thermal conductivity. A number of flow through-porous-media problems have been solved using the heat conduction analogy. The Theis equation for transient flow toward wells (See Chapters 8 on hydraulics of wells) was obtained using an analogous problem in heat flow. *Ohm's law* can be stated as $I = V/R$, where I is the electric current, R is the resistance, and V is the voltage or potential difference across the resistor. For a cylindrical wire of length L, cross-sectional area A and conductivity c, the resistance is $R = L/(cA)$ so that the expression for the current becomes

$$I = cA \frac{V}{L} \tag{6}$$

The similarity between Equations (4) and (6) is now obvious. The flow rate is analogous to the current, and the drop of head to the voltage drop. This analogy forms the basis for the resistance network models of aquifers (Karplus, 1958). *Fick's law* states that diffusion of a solute takes place along the concentration gradient from zones of high concentration to zones of low concentration, and the coefficient of proportionality is the diffusion coefficient. Fick's law is used in Section 2.13 and in Chapters 14 and 15, which are concerned with pollutant transport.

2.4.2   Limitations of Darcy's Law

Darcy's law implies that the flow is laminar, as is generally the case in porous media. The limit of validity can be stated in terms of the Reynolds Number, N_R,

$$N_R = qD / v \tag{7}$$

where q is a velocity, v is the kinematic viscosity of the fluid, defined as its dynamic viscosity μ divided by its density ρ, and D is a representative length. For flow in porous media, q is taken equal to the specific discharge, and the representative length, D, is often taken equal to the pore size or the effective grain diameter, d_{10} (the grain size such that 10% of the material is larger by weight). The Reynolds number measures the importance of the inertia forces relative to the viscous forces. The Reynolds number depends on the viscosity, which varies with temperature as shown in Table 2.3. As a result N_R also varies with temperature. Likewise the hydraulic conductivity $K = k\rho g/\mu$, where k is the intrinsic permeability of the porous medium (see Chapter 1), which also varies with the temperature through μ. Schneebeli (1955),

TABLE 2.3 Density and Viscosity of Water

Temperature °C	Density ρ kg/cm^3	Viscosity $\mu \times 10^3$ Ns/m^2	Kinematic Viscosity $\nu \times 10^6$ m^2/s
0	999.8	1.781	1.785
5	1000.0	1.518	1.519
10	999.7	1.307	1.306
15	999.1	1.139	1.139
20	998.2	1.002	1.003
25	997.0	0.890	0.893
30	995.7	0.798	0.800
40	992.2	0.653	0.658
50	988.0	0.547	0.553
60	983.2	0.466	0.474
70	977.8	0.404	0.413
80	971.8	0.354	0.364
90	965.3	0.315	0.326
100	958.4	0.282	0.294

From Daugherty, L., Franzini, J. B., and Finnemore, E. J. 1985. *Fluid Mechanics with Engineering Applications*. 8th ed. McGraw-Hill Inc., New York. With permission.

using spheres of uniform, diameter found that deviations from Darcy's law start at $N_R \cong 5$ as inertia forces become effective and that turbulent flow started around $N_R \cong 60$. For flows in which the dimension D is large, such as in rocks with large fractures (see Chapter 17) or in karstic limestones (see Chapter 18), the flow can be turbulent and Darcy's law does not apply.

Darcy's law, as given above, applies to *isotropic* media, that is, where the hydraulic conductivity is independent of direction. It also applies to flows where the direction of the hydraulic conductivity corresponds to the direction of the hydraulic gradient. In *anisotropic* media the hydraulic conductivity depends upon the direction of measurement. Then a hydraulic conductivity *tensor* is used, and Darcy's law is expressed as a *tensorial equation* (see Chapter 3).

2.4.3 Laboratory Measurement of Hydraulic Conductivity

Samples of the aquifer must be obtained and returned to the laboratory in undisturbed condition. This is generally possible for consolidated materials but usually impossible for unconsolidated material and rarely possible for fissured aquifers. The samples must be representative of the aquifer. Where the aquifer has horizontally bedded strata, the samples can be collected from bore holes that intersect the several strata. They can also be obtained from cliffs or quarry faces. Where the strata are inclined, the samples likewise can be obtained from bore holes as well as from the outcrop. Bore-hole samples are preferred because the outcrop material may be weathered and consequently unrepresentative of the aquifer. Consolidated aquifer samples typically are cylinders or cubes with diameter and length of 25 to 50 mm. Cubes have the advantage that they permit the determination of the permeability in three directions.

Permeameters are used for the laboratory determination of the hydraulic conductivity making use of Darcy's law. It is best to use groundwater from the formation in the permeameter test because this water will be in chemical equilibrium with the aquifer material. In particular, clays can swell or shrink with changes in water chemistry. There are two types of permeameter: the *constant head* permeameter (Figure 2.7a) used for noncohesive soils such as sands and gravels, and the *falling head* permeameter (Figure 2.7b) used for materials with lower hydraulic conductivity. For the constant head permeameter, the hydraulic conductivity obtained from Darcy's law is

$$K = \frac{\forall L}{Ath} \tag{8}$$
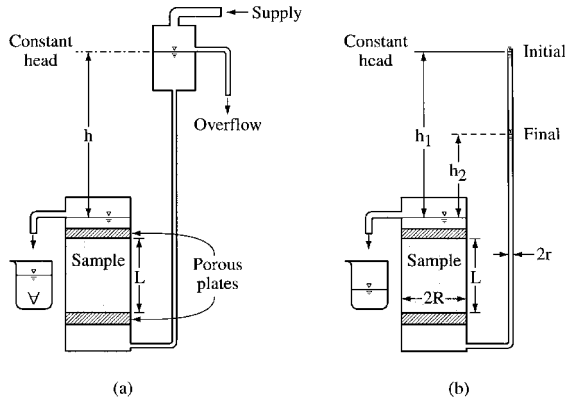
FIGURE 2.7 Permeameters. (a) Constant head. (b) Falling head.

where $\forall$ is the volume of water collected in time t, A and L are the cross section area and the length of the sample, respectively, and h is the constant head. For the falling head permeameter, the hydraulic conductivity is obtained by equating the flow rate through the sample to the flow rate obtained from the observed head drop as

$$K = \frac{r^2 L}{R^2 t} \ln \frac{h_1}{h_2} \tag{9}$$

where R and r are the radii of the sample and of the tube, respectively, L is the length of the sample, h_1 and h_2 are the heads at the beginning and at time t later.

2.4.4   Field Measurement of Hydraulic Conductivity

A dependable method of field determination of the hydraulic conductivity is by pumping tests. The hydraulic conductivity is obtained from observations of the water levels near pumping wells. It yields an integrated value of K rather than the punctual information obtained by laboratory tests. It also has the advantages that the aquifer is not disturbed and formation water is used. The pumping tests are described in detail in Chapter 8 on hydraulics of wells and aquifer tests.

Tracer tests using a dye such as fluorescein or a salt such as calcium chloride can also be used. If there is a drop of water table h in a distance L between the injection test hole and the observation bore hole and t is the observed travel time between the two bore holes, the hydraulic conductivity is obtained by equating the pore velocity obtained by Darcy's law and that obtained by dividing the distance by the time. This results in

$$K = \frac{nL^2}{ht} \tag{10}$$

where n is the porosity of the material. In practice this test is difficult to accomplish because the flow direction must be known exactly, the distance between the bore holes should be small enough that the travel time does not become too long, and it is assumed that the aquifer is not stratified.

Other tests such as the *slug test*, the *auger-hole* test, and the *piezometer test* are based on the observation of the rate of rise of the water in a bore hole after the water level has been abruptly lowered by removal of water with a bailer or bucket. These tests tend to give a more localized value of the hydraulic conductivity than the well tests but are less expensive to conduct. For more detail about these tests, the

reader is referred to Bouwer (1978). Chapters 10 and 11 on aquifer characterization give detailed information on the several methods for the determination of aquifer properties.

2.5   One-Dimensional Hydraulics

The two basic laws of hydraulics are the continuity equation and the energy equation. They are discussed here for the case of *steady incompressible* flow. The continuity equation is a statement of the conservation of mass. For an incompressible fluid such as water the equation becomes a conservation of volumes:

$$Q = A_1 V_1 = A_2 V_2 \tag{11}$$

where Q is the flow rate $(L^3 T^{-1})$, A_1 and A_2 are the cross-sectional areas, and V_1 and V_2 are the mean velocities at sections 1 and 2, respectively. This assumes that there is no inflow or outflow between sections 1 and 2. Figure 2.8 illustrates the continuity in a karst conduit.
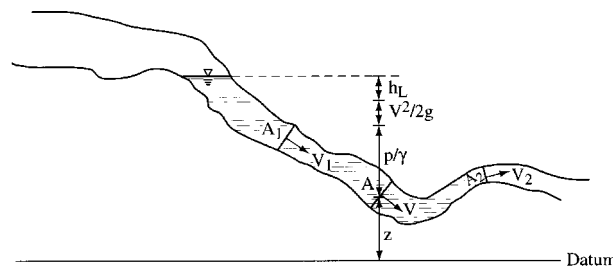
FIGURE 2.8 Continuity, velocity head, pressure head, and elevation head and head loss in a karst system.

The *conservation of energy* for steady flow is:

$$z_1 + \frac{p_1}{\gamma} + \frac{V_1^2}{2g} = z_2 + \frac{p_2}{\gamma} + \frac{V_2^2}{2g} + h_L \tag{12}$$

where z is the elevation head or the elevation above a datum or reference plane, p/γ is the pressure head, (γ is the specific weight of the liquid), $V^2/2g$ is the velocity head, and h_L is the head loss. Each term has the dimension of length, (L), and represents a form of energy per unit weight. As discussed earlier, the elevation head is the potential energy per unit weight, the pressure head is the flow work per unit weight, the velocity head is the kinetic energy per unit weight, and the head loss is the energy loss due to friction or other causes per unit weight. The equation can also be written as energy per unit mass

$$gz_1 + \frac{p_1}{\rho} + \frac{1}{2} V_1^2 = gz_2 + \frac{p_2}{\rho} + \frac{1}{2} V_2^2 + gh_L \tag{13}$$

where ρ is the fluid density. If the head loss is negligible and is made equal to zero, then the equation is known as the *Bernoulli equation.* It states that in an unbranched conduit the sum of the potential, pressure, and kinetic energies remains constant if the losses are neglected. In flow through porous media the velocity is very small, and the velocity head is generally neglected. The friction losses due to viscosity, however, are important. Figure 2.8 illustrates the terms of Equation (12) and piezometric surfaces in a karst system.

In karst terrain, flow usually occurs in larger conduits which, hydraulically, act as pipes or fissures. If the Reynolds number is sufficiently small, the flow is laminar, otherwise it is turbulent. For circular pipes the Reynolds number is defined as

$$N_R = \frac{\rho VD}{\mu} = \frac{VD}{\nu} \tag{14}$$

where D is the diameter, $V = Q/A$ is the average flow velocity, that is the flow rate Q divided by the cross-sectional area A, ρ is the density of the fluid, μ is the absolute viscosity, and ν is the kinematic viscosity. With a consistent set of units, the Reynolds number is a dimensionless quantity. The flow is laminar if the Reynolds number is less than 2000, and it is generally turbulent if the Reynolds number is larger than 4000. For values between 2000 and 4000 the flow is in a transitional regime.

For noncircular conduits such as fractures, the Reynolds number is defined as

$$N_R = \frac{\rho V\left(4R_h\right)}{\mu} \tag{15}$$

where $R_h = A/P$ is the *hydraulic radius* which is the ratio of the cross-sectional area, A, to the wetted perimeter, P. For a circular cross-section $D = 4R_h$.

Friction head losses in conduit are calculated using the *Darcy-Weisbach formula*, which is generally written as

$$h_L = f\frac{L}{D}\frac{V^2}{2g} \tag{16}$$

in which f is a dimensionless quantity called the *friction coefficient*. For laminar flow in circular conduits (also known as *Poiseuille flow*), $f = 64/N_R$ and the flow rate is thus expressed as

$$Q = \frac{\rho g D^2}{32\mu}A\frac{h_L}{L} \tag{17}$$

This equation is seen to be of the same form as Darcy's law (4) with $K = \rho g D^2/(32\mu)$ or $k = D^2/32$ and $i = h_L/L$.

The Darcy-Weisbach formula (16) also holds for turbulent flows. Methods of calculation of the friction coefficient f for turbulent flows are described in fluid mechanics textbooks, such as Franzini and Finemore (1997).

Free surface *open channel flow* can occur in large karst cavities. Free surface streams in karst aquifers are discussed in Chapter 18.

2.6 One-Dimensional Flow through Porous Media — Leaky Aquifers

Simple one-dimensional flows in porous media can be analyzed using Darcy's law as expressed in Equation (5). As an example, consider the case of a *leaky aquifer*. Many aquifers are not fully confined or unconfined. One aquifer may be overlain by another and the stratum separating them is not fully impervious: its hydraulic conductivity is much less than that of either the top or bottom aquifer but it is not zero. The rate of leakage q (volume per unit area per unit time) can be calculated using Darcy's law as
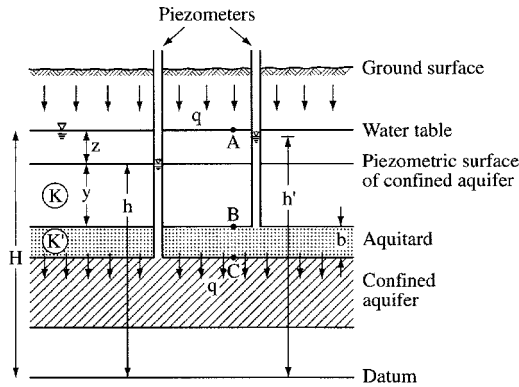
FIGURE 2.9 Leaky aquifer.

$$q = K' \frac{(h' - h)}{b} \tag{18}$$

where K' and b are the vertical hydraulic conductivity and the thickness of the aquitard; h' and h are the heads at the top of the aquitard (bottom of upper aquifer) and at the top of the lower aquifer, just below the aquitard (Figure 2.9). Observe that the level in the right piezometer is slightly lower than the water table due to the downward motion of the flow. This can be verified by writing Darcy's law between point A at the water table and point B at the bottom of the unconfined aquifer: $q = K(H - h')/(y + z)$, from which $H - h' = q(y + z)/K$.

Example 5: As an example of a one-dimensional flow problem, consider the case of the confined aquifer of Figure 2.9 that is recharged from an unconfined aquifer through an aquitard. The recharge rate is 0.3 m/yr or 8.22×10^{-4} m/d. The water table is at H = 30 m above the datum. The aquitard is 2 m thick and its vertical hydraulic conductivity is $K' = 10^{-3}$ m/d. The unconfined aquifer is 20 m thick and has a hydraulic conductivity $K = 10^{-1}$ m/d. Find the piezometric head h' at the bottom of the unconfined aquifer and the difference in elevation between the water table and the piezometric surface of the confined aquifer.

Solution: Let y be the height of the piezometric surface over the top of the aquitard and z the difference in elevation between the water table and the piezometric surface (Figure 2.9). Applying Darcy's law (5) between points A and B yields: $8.22 \times 10^{-4} = 10^{-1} (H - h')/(y + z)$. Thus h' = H $- 8.22 \times 10^{-3} (y + z) = 30 - 8.22 \times 10^{-3} (20) = 29.84$ m. Writing Darcy's equation between the top and the bottom of the aquitard yields: $8.22 \times 10^{-4} = 10^{-3} (h' - h)/b$, thus h = h' $- 8.22 \times 10^{-1} b = 29.84 - 8.22 \times 10^{-1} \times 2 = 28.20$. Hence, z = H $-$ h = 30 $-$ 28.20 = 1.80 m.

When a well discharges from a leaky confined aquifer, the piezometric surface is lowered throughout a wide circular area. This lowering, called draw down, is largest near the well and decreases outward. This variable drawdown changes the head differential between the confined and unconfined aquifers and alters the rate of leakage through the aquitard. From Darcy's law it follows that, at any point, the downward flow is proportional to the difference of elevation between the water table of the unconfined aquifer and the piezometric surface of the confined aquifer. Steady and unsteady flows toward wells in leaky aquifers are discussed in Chapter 8.

2.7   Dupuit–Forchheimer Assumptions

For some of the two-dimensional flow problems, one component of the flow can be neglected with respect to the other. In particular, in some unconfined flows with a free surface, the vertical component

of the flow can be neglected. This approximation pioneered by Dupuit (1863) and utilized later by Forchheimer (1930) is known as the Dupuit–Forchheimer assumption. It gives reasonable results when the depth of the unconfined flow is shallow and the slope of the free surface is small. These assumptions are summarized as follows:

1. The flow is horizontal at any vertical cross-section
2. The velocity is constant over the depth
3. The velocity is calculated using the slope of the free surface as the hydraulic gradient
4. The slope of the water table is relatively small

2.7.1 Steady Flow over a Horizontal Aquiclude

A simple application of the Dupuit–Forchheimer approximations is the analysis of steady flow through an unconfined aquifer overlying an impervious horizontal aquiclude (Figure 2.10). The discharge per unit width is $q_x = qh$, where q is the Darcy velocity or specific discharge and h is the depth of flow. From Darcy's law,

$$q_x = -Kh\frac{dh}{dx} \tag{19}$$
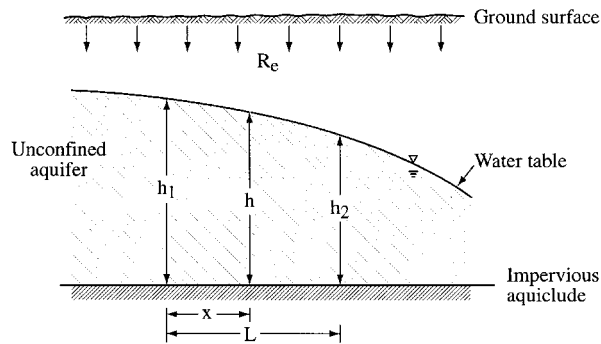
FIGURE 2.10 Steady flow through an unconfined aquifer overlying an impervious horizontal aquiclude.

Equation (19) is integrated from $x = 0$ (where $h = h_1$) to $x = L$ (where $h = h_2$) to obtain the *Dupuit equation:*

$$q_x = K\frac{h_1^2 - h_2^2}{2L} \tag{20}$$

If there is a uniform recharge with a rate R_e, then $q_x = R_e x$, with $x = 0$ at the groundwater divide, and

$$R_e x = -Kh\frac{dh}{dx} \tag{21}$$

from which

$$K\left(h_1^2 - h_2^2\right) = R_e L^2 \tag{22}$$

Example 6. For an unconfined aquifer with a hydraulic conductivity K = 1.75 m/d, an effective porosity of 0.3, and water depths of 10 m and 8 m at two observation wells 200 m apart, calculate the discharge per unit width, the specific discharge, and the pore velocity.

Solution: From Equation (20), the discharge per unit width is $q_x = 1.75(10^2 - 8^2)/(2*200) = 0.1575$ m^2/d. The Darcy velocity or specific discharge at the observation well with a 10-m depth is $q = q_x/h_1 = 0.1575/10 = 0.01575$ m/d. The pore velocity is $v = 0.01575/0.3 = 0.0525$ m/d.

2.7.2 Seepage from Open Channels

The Dupuit–Forchheimer assumptions can be used to analyze the seepage from an open channel embedded in a homogeneous soil underlain by a material of much lower hydraulic conductivity, assumed impervious in the analysis. (Figure 2.11). The average slope of the water table is $D_w/(L - 0.5 \ W_s)$. The specific discharge is given by Darcy's law (5) as $q = KD_w/(L - 0.5W_s)$. The average flow depth is $D_i + H_w - 0.5D_w$. The total seepage per unit length of channel (on both sides) is thus

$$Q = 2KD_w \frac{D_i + H_w - 0.5D_w}{L - 0.5W_s} \qquad (23)$$
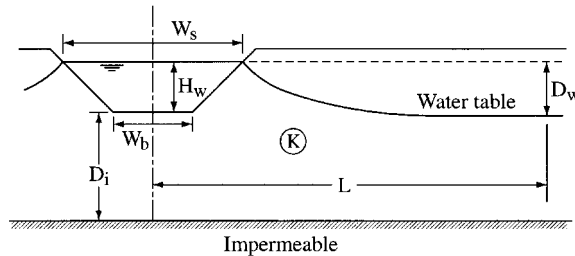
FIGURE 2.11 Seepage from a channel in a soil underlain by impervious material. (From Bouwer, H. 1978. *Groundwater Hydrology*. McGraw-Hill Inc., New York. With permission.)

Bouwer (1969) states that this type of analysis gives reasonable results for $D_i < 3W_s$. General solutions of the canal seepage problem have been given by Bouwer (1969, 1978), Harr (1962), Polubarinova-Kochina (1962), and Yussuff et al. (1994). The case of partially lined channels has been examined by Subramanya et al. (1973). Integrated groundwater – surface water models have been developed and applied to the Imperial County in southern California by Saquib et al. (1995) and Taghavi et al. (1995).

Example 7. Estimate the seepage from a canal with a depth $H_w = 1$ m, dug in a soil with a hydraulic conductivity K = 2 m/d so that the distance from the bottom of the channel to the impervious stratum is $D_i = 10$ m, given a drop in the water table $D_w = 0.5$ m is observed at a distance $L - 0.5 \ W_s = 400$ m from the shore. Is this seepage a significant portion of the flow?

Solution: Replacement in Equation (23) yields Q = 0.05375 m^3/d per m of channel length or 53.75 m^3/d per km. Assume the channel had a bottom $W_b = 4$ m, side slopes of 45° and a longitudinal slope of $S_o = .0004$, and a Manning roughness coefficient n = 0.022 (straight channel excavated in clean earth, after weathering (Chow, 1959). The discharge capacity can be calculated by Manning's equation $Q = 1/n \ AR_h^{2/3}S^{1/2}$ (see Equation [12], Chapter 18 on karst hydrology). The cross-sectional area is A = 5 m^2. The wet perimeter is P = 6.828 m. The hydraulic radius is $R_h = 5/6.828 = 0.732$ m. The flow rate is Q = 5*0.732$^{2/3}$*0.0004$^{1/2}$ /0.022 = 3.692 m^3/s or 318,988 m^3/d. In a distance of 40 km, the seepage loss would amount to 40*53.75*100/318,988 = 0.674 % or less than 1%. This does not appear to be significant.
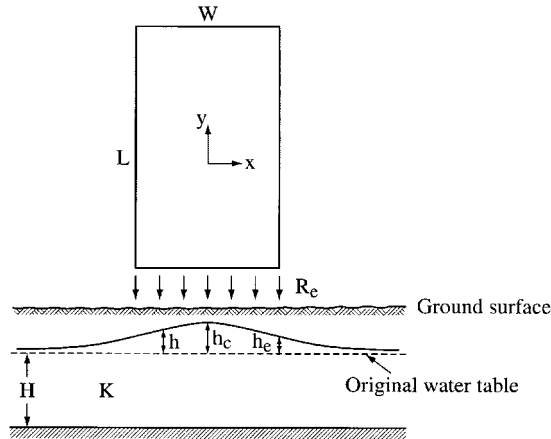
FIGURE 2.12 Seepage from a recharge basin.

2.7.3   Recharge Basins

A third application of the Dupuit-Forchheimer assumptions is the analysis of the recharge of an uncon-fined aquifer from a recharge basin. A long rectangular basin of width W is considered with the assump-tion that the flow is horizontal and steady (Figure 2.12). Thus the depth of the aquifer, H, should not be large compared to the width, W, of the recharge basin, i.e., H ≤ W, otherwise a more detailed analysis, such as a digital model, should be used. Letting R_e be the infiltration rate from the recharge area, x the horizontal distance from the centerline of the recharge area, h the height of the groundwater mound above the static water table, and approximating the average transmissivity T = K (H + h_c/2) by KH, (Figure 2.12), the Dupuit–Forchheimer assumptions yield

$$R_e x = -T\frac{dh}{dx} \tag{24}$$

Integration of (24) yields $h_c - h_e = R_e W^2/(8T)$, where h_c and h_e are the height of the groundwater mound at the center and at the edge of the recharge basin, respectively. If h_c and h_e are measured in the field, then the aquifer transmissivity can be calculated. Further discussion can be found in Bouwer (1970, 1978). An analytical solution has been given for rectangular recharging areas by Mariño (1975).

Example 8: A long recharge basin has a width W = 70 m and the recharge rate is R_e = 0.6 m/d. The observed heights of the mound are h_c = 1.2 m, and h_e = 0.7 m. Find the aquifer transmissivity.

Solution: From the integration of Equation (24) T = $0.6(70)^2/[8(1.2 - 0.7)]$ = 735 m²/d. Note that if H > 70 m, the tranmissivity obtained is an "effective" transmissivity, as the lower layers of the aquifer are not fully contributing to the flow.

2.7.4   Steady Flow Toward a Well in a Confined Aquifer

Perhaps the most important application of the Dupuit assumptions is the calculation of a steady flow toward a well as done by Forchheimer. Consider a well that fully penetrates an isotropic confined aquifer of hydraulic conductivity *K*. The initial piezometric level is assumed to be horizontal so that originally there is no motion of the groundwater. As water is being pumped, it flows from the aquifer toward the well lowering the piezometric surface and creating a drawdown. (Figure 2.13). Consider two imaginary cylinders around the well with radii r_1 and r_2. The flows through each of these cylinders of height *b* are
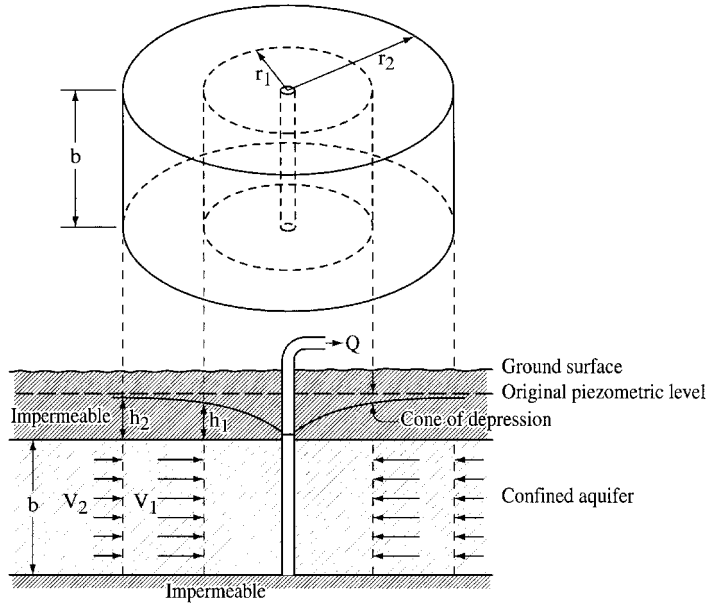
FIGURE 2.13 Flow toward a fully penetrating well in a confined aquifer.

horizontal, thus satisfying the Dupuit–Forchheimer assumptions. Furthermore the flows must be equal to the discharge at the well. Thus, from Darcy's law

$$Q = 2\pi r_1 bKi_1 = 2\pi r_2 bKi_2 \tag{25}$$

where i_1 and i_2 are the values of the hydraulic gradient at radii r_1 and r_2, respectively. Since $2\pi bK$ is constant, and since $r_1 < r_2$ then $i_1 > i_2$. Thus the hydraulic gradient becomes steeper as the water approaches the well, creating the *cone of depression*. As the same flow occurs through the two cylinders, the gross velocity, or specific discharge, increases as the well is approached. Writing Equation (25) as $Q = 2\pi rbK \, dh/dr$ and integrating yields the *Thiem equation:*

$$Q = \frac{2\pi Kb\left(h_2 - h_1\right)}{\ln\left(r_2/r_1\right)} = \frac{2\pi T\left(h_2 - h_1\right)}{\ln\left(r_2/r_1\right)} \tag{26a,b}$$

where $T = Kb$ is the transmissivity of the aquifer. In the case of an unconfined aquifer the saturated depth b is not constant and decreases toward the well. The vertical component of the flow comes into effect and the Dupuit–Forchheimer assumptions are not fully satisfied. However, Equation (26b) can still be used with reasonable results if T is interpreted as the average transmissivity $K(h_1 + h_2)/2$. Equation (26) can also be used with one observation well and considering the pumping well as the other observation point.

Example 9: A well with a radius of 50 cm completely penetrates an unconfined aquifer. It has been pumped for a long time at the rate of 15,000 m³/d. The drawdown in the well is 10 m. Find the hydraulic conductivity of the formation given that the well essentially does not affect the water table at a distance of 600 m where the depth of the water table is 50 m. Well losses are neglected.

Solution: Using equation (26b) with $T = K(h_1 + h_2)/2$ and with $h_1 = 50 - 10 = 40$ m and solving for K, one obtains

$$K = \frac{Q\ln(r_2/r_1)}{\pi(h_2^2 - h_1^2)} = \frac{15000\ln(600/0.50)}{\pi(50^2 - 40^2)} = 37.6\,\text{m/d}$$

Solving the well equation for the aquifer properties is called the *inverse problem.* The method employed in Example 9 using the Thiem equation only gives a rough approximation because in practice steady state rarely exists and the transient flow formulas must be used instead. The transient flow methods (discussed in Chapter 8 on well hydraulics and aquifer tests) also yield the storage coefficient S in addition to the hydraulic conductivity K.

2.8   Velocity Potential, Force Potential, and Flow Nets

The piezometric head $h = z + p/\gamma + C$, where C is an arbitrary constant, is interpreted as an energy or potential per unit weight. The quantity

$$\Phi = Kh = K\left(z + \frac{p}{\gamma}\right) + C \tag{27}$$

is defined as the *velocity potential,* Φ. By virtue of Darcy's law, the negative of its derivative in the flow direction, for constant K, is the Darcy velocity, q. In a more general term, the negative of the gradient of the velocity potential is a velocity. This is the definition of the velocity potential in classical hydrodynamics. (Note: in Chapter 4 a minus sign is introduced in the right-hand side of [27]).

Hubbert (1953, 1987) introduced the concept of the *force potential* Ψ,

$$\Psi = gh = \frac{g}{K}\Phi \tag{28}$$

(Note: in Chapter 4 the symbol ψ is used for the "stream function," Equation [16]). The force potential at a point is the work that is required to move a unit mass of fluid from a reference elevation and pressure to the elevation and pressure at the given point. In vector form, for K constant,

$$q = -\text{grad}\,\Phi \qquad \text{or} \qquad q = -\frac{K}{g}\text{grad}\,\Psi \tag{29a,b}$$

The gradient of the force potential is the force per unit of mass acting upon the water at a given point. The right formula (29b) is a generalized Darcy's law in three dimensions (Hubbert 1953, 1987).

From Equation (29b) q is seen to have the same direction as $-\text{grad}\,\Psi$ as long as K is constant. Thus the *streamlines*, which are lines everywhere tangent to the velocity vector, are perpendicular to lines of $\Psi = $ constant or *equipotential* lines. The streamlines and the equipotential lines are orthogonal. A network of streamlines and equipotential lines form the *flow net* which is a useful tool in the analysis of two-dimensional flows. When the hydraulic conductivity is not constant, then Equation (29a) must be used. With some practice, good flow nets can be drawn by hand. The equipotential lines are drawn so that the drop of head Δh or potential drop $\Delta\Phi$ between adjacent lines is the same. The streamlines are drawn so that the same fraction of the total flow ΔQ takes place between adjacent streamlines. They are normal to the equipotential lines (Figure 2.14) forming "square" shapes. The flow between two streamlines, ΔQ, is obtained from Darcy's equation as

$$\Delta Q = Kbw\frac{\Delta h}{\Delta L} \tag{30}$$
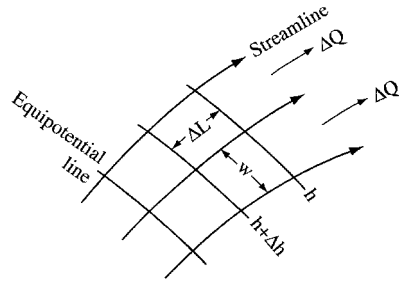
FIGURE 2.14 Streamlines and equipotential lines in a flow net.
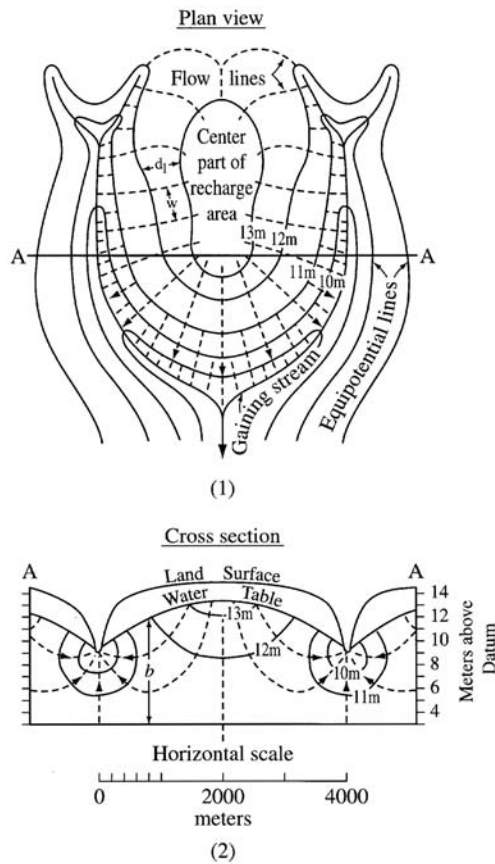
FIGURE 2.15 Flow net of a recharge area and gaining stream. (Adapted from Heath, R. C. 1995. *Basic Ground-Water Hydrology*. U.S. Geological Survey, Water-Supply paper 2200, seventh printing, Denver, CO.)

where b is the average depth of flow, Δh and ΔL are the difference in head and the distance between adjacent equipotential lines, and w is the width between adjacent streamlines for the square considered. The total flow through a group of n flow paths is $Q = n\Delta Q$. Figure 2.15 shows a flow net for a recharge area in an unconfined aquifer.

The mathematics, construction and application of flow nets are discussed in Chapter 4 on groundwater and seepage.

2.9   Laplace's Equation

Laplace's equation is fundamental to the analysis of many groundwater flow problems. It arises by the combination of Darcy's law and the equation of continuity or conservation of mass for a homogeneous isotropic aquifer. Consider a steady flow of an incompressible fluid through an elementary cube (Figure 2.16) of a porous medium of porosity n. Let u, v, w be the velocity components in the x, y, z directions. The inflow through the vertical face near the origin is n u dy dz. The outflow through the vertical face away from the origin is

$$\left[nu + \frac{\partial}{\partial x}(nu)\,dx \right] dy\,dz$$

and the net change of volume in the x-direction is

$$\frac{\partial}{\partial x}(nu)\,dx\,dy\,dz$$

The sum of the net changes of volume in the x,y,z directions must be equal to zero. Thus

$$\frac{\partial}{\partial x}(nu) + \frac{\partial}{\partial y}(nv) + \frac{\partial}{\partial z}(nw) = 0$$

For an incompressible isotropic homogeneous porous medium, the equation of continuity for steady incompressible flow is thus

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0 \tag{31}$$

Using Darcy's law for an isotropic homogeneous medium

$$u = -K\frac{\partial h}{\partial x}, \quad v = -K\frac{\partial h}{\partial y}, \quad w = -K\frac{\partial h}{\partial z}$$

Laplace's equation is obtained in terms of the head h:

$$\nabla^2 h = \frac{\partial^2 h}{\partial x^2} + \frac{\partial^2 h}{\partial y^2} + \frac{\partial^2 h}{\partial z^2} = 0 \tag{32}$$
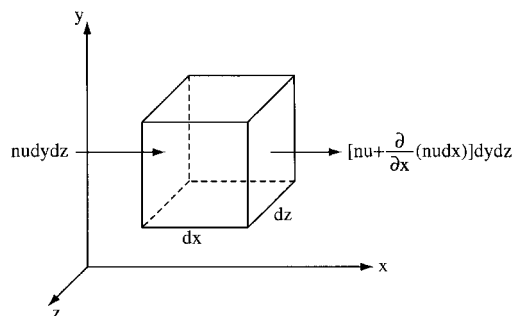
FIGURE 2.16 Elementary cube of porous media.

If in (31) the velocity components are expressed in terms of the velocity potential (see 29a)

$$u = -\frac{\partial \Phi}{\partial x}, \quad v = -\frac{\partial \Phi}{\partial y}, \quad w = -\frac{\partial \Phi}{\partial z}$$

then Laplace's equation is expressed in terms of the velocity potential Φ

$$\nabla^2 \Phi = \frac{\partial^2 \Phi}{\partial x^2} + \frac{\partial^2 \Phi}{\partial y^2} + \frac{\partial^2 \Phi}{\partial z^2} = 0 \tag{33}$$

Equation (33) is fundamental in the analysis of flow nets as developed in Chapter 4.

If the porous medium of a confined aquifer is assumed compressible with a compressibility α (reciprocal of the modulus of elasticity, for quartz 2×10^{-11} Pa) and the water has a compressibility β (5×10^{-10} Pa^{-1}), then Equation (32) is replaced by a more elaborate expression which is discussed in Chapter 3. A good approximation is

$$\nabla^2 h = \frac{1}{K}\left[\beta \rho g \left(n + \frac{\alpha}{\beta}\right)\right]\frac{\partial h}{\partial t} \tag{34}$$

The term in square brackets is the specific storage or storage coefficient, S, namely the volume of storage release per unit drop of the piezometric surface per unit horizontal area, (see Chapter 1). It is dimensionless and for confined aquifers it is of the order of 5×10^{-2} to 10^{-5}. The term $\rho g \alpha$ represents the water yield from storage due to the compression of the porous medium, and $\beta \rho g n$ is the water yield resulting from the expansion of water storage. For an unconfined aquifer, S is the drainage porosity or specific yield, that is, the volume of water release per unit drop of the water table per unit horizontal area. Since, for the elementary cube of Figure 2.16, the difference between the inflow and the outflow is now equal to the rate of change of storage, the flow is now unsteady, thus requiring the time derivative in the right-hand-side of Equation (34). For two-dimensional flow in a horizontal confined aquifer Equation (34) becomes

$$\frac{\partial^2 h}{\partial x^2} + \frac{\partial^2 h}{\partial y^2} = \frac{S}{T}\frac{\partial h}{\partial t} \tag{35}$$

where $T = Kb$ is the transmissivity of the aquifer, and b is the depth for a confined aquifer and the average depth for an unconfined aquifer. If there is a leakage or inflow rate q in the aquifer per unit area then Equation (35) becomes

$$\frac{\partial^2 h}{\partial x^2} + \frac{\partial^2 h}{\partial y^2} + \frac{q}{T} = \frac{S}{T}\frac{\partial h}{\partial t} \tag{36}$$

The leakage rate can be calculated by Darcy's law (see Equation 18).

The Dupuit–Forchheimer assumption that was applied to one-dimensional unconfined flow (see Equations 18 and 19) can be generalized to the two-dimensional case. In this case, the inflows and outflows into a small parallelepiped of unconfined aquifer can be calculated by Darcy's law, and their sum must be equal to zero for a steady incompressible flow. This yields

$$\frac{\partial^2 \left(h^2 \right)}{\partial x^2} + \frac{\partial^2 \left(h^2 \right)}{\partial y^2} = 0 \tag{37}$$

It is seen that for the Dupuit–Forchheimer assumptions the square of the head must satisfy Laplace's equation (Harr, 1962). If there is recharge at the rate R_e, then Equation (37) is replaced by (Fetter, 1994)

$$- K \left(\frac{\partial^2 \left(h^2 \right)}{\partial x^2} + \frac{\partial^2 \left(h^2 \right)}{\partial y^2} \right) = 2 R_e \tag{38}$$
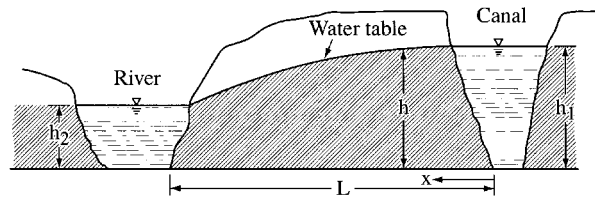
FIGURE 2.17 Example 10.

Example 10: A river and a canal run parallel to each other L = 500 m apart (Figure 2.17). They fully penetrate an unconfined aquifer with a hydraulic conductivity of 0.3 m/d. The elevation of the water surface in the river is 1.25 m lower than in the canal where the depth is 5 m. Assuming no recharge, find the water table elevation midway between the river and the canal and find the discharge into the river.

Solution: The one-dimensional form of Equation (38) is $\dfrac{d^2 (h^2)}{dx^2} = 0,$ which integrates as $h^2 = c_1 x + c_2$. Applying the boundary conditions: at x = 0, h = h_1 and at x = L, h = h_2, one obtains $h^2 = h_1^2 - (h_1^2 - h_2^2)\, x/L$. Thus with h_1 = 5 m, h_2 = 3.75 m, and L = 500 m, h = $[5^2 - (5^2 - 3.75^2)* 0.5]^{1/2}$ = 4.42 m. The seepage into the river is given by q = –Kh(dh/dx) = $K(h_1^2 - h_2^2)/2L$ = 0.3 $(5^2 - 3.75^2)/1000$ = 0.0033 m^3/(m d) or 3.3 m^3/(km d).

2.10 Land Subsidence

Pumping large volumes of water from a confined aquifer at rates substantially larger than the natural recharge causes a contraction of the aquifer which can result in a downward movement of the land surface. In unconfined aquifers, groundwater pumping causes a downward movement of the water table which likewise can lead to a downward movement of the land surface. This downward movement is called subsidence or consolidation. This movement can be a few centimeters to several meters. If the subsidence is not uniform, the differential settlement can produce severe damage to structures. Important subsidence has occurred in the San Joaquin Valley in California, in Mexico City, in Venice, around Shanghai, and in southern Taiwan. These large subsidences tend to occur in thick deposits containing fine sands, silt, and clays and often are the result of excessive pumping. Decrease of the groundwater pumpage can reduce and sometimes reverse the subsidence as is currently the case in the Harris-Galveston area in south-central Texas (Hibbs, 1997). UNESCO (United Nations Educational, Scientific, Cultural Organization) has published a guidebook to studies of land subsidence due to groundwater withdrawals (Poland, 1984).

2.10.1 Calculation of Subsidence

Consider a unit area of a horizontal plane at a depth Z below the ground surface. The total downward pressure P_t due to the weight of the overburden on the plane is resisted partly by the upward hydrostatic pressure P_h and partly by the *intergranular pressure* P_i exerted between the grains of the material: $P_t = P_h + P_i$ or $P_i = P_t - P_h$. (See also Chapter 1, Section 1.4.6.)

A lowering of the water table results in a decrease of the hydrostatic pressure and a corresponding increase of the intergranular pressure. If P_{i1} and P_{i2} denote the intergranular pressures before and after a drop in the water table or piezometric surface, the vertical subsidence can be calculated as

$$S_u = Z \frac{P_{i2} - P_{i1}}{E} \tag{39}$$

where Z is the thickness of the soil layer and E is the modulus of elasticity of the soil. Typical ranges of values of E are given in Table 2.4. In general, the modulus of elasticity increases nonlinearly with the intergranular pressure. If there are layers of different soil types, the subsidences are calculated separately for each layer and added to obtain the total subsidence. As the modulus of elasticity of clayey materials is much less than that of sand or gravel, most of the settlement occurs in the clayey layers.

TABLE 2.4 Modulus of Elasticity of Soils and Rocks

Material	Modulus of Elasticity N/cm²
Peat	10–50
Clay, plastic	50–390
Clay, stiff	390–790
Clay, medium hard	790–1470
Sand, loose	980–1960
Sand, dense	4910–7850
Gravel, sandy gravel, dense	9806–19620
Rock, fissured, jointed	14710–294200
Rock, sound	294200–∞

Source: Jumikis, A. R. 1984. *Soil Mechanics.* Robert E. Krieger Co., Malabar, FL. Reproduced with permission.

The previous equation can also be used to calculate the rebound when the intergranular pressure decreases. Caution must be exercised because the modulus of elasticity is not the same for decompression as for compression. This is particularly the case for clays. For Boston blue clay the rebound modulus of elasticity is only about 50% of that for compression (Bouwer, 1978, p. 323). If subsidence has occurred for a long time, complete rebound is unlikely

Example 11: Consider a 60-m-thick sand layer. The water table is located at a depth of 10 m below the ground surface. Calculate the total and the intergranular pressures at 10 m depth and at the bottom of the sand layer, given that the porosity of the sand is n = 0.35, its volumetric water content above the water table is θ = 0.08, the specific weight of the solids is γ_s = 25.5 kN/m³, and the specific weight of the water is γ = 9.81 kN/m³.

Solution: At the water table the intergranular pressure, which is also the total pressure, is P_t = 10[(1 – 0.35)25.5 + 0.08*9.81] = 173.6 kPa. The total pressure at the bottom of the sand layer is P_t = 173.6 + 50[(1 – 0.35)25.5 + 0.35*9.81] = 1174.0 kPa. The hydrostatic pressure at the bottom of the sand layer is P_h = 9.81*50 = 490.5 kPa. The intergranular pressure is thus 1174.0 – 490.5 = 683.5 kPa.

Example 12: If in the previous problem the water table drops 40 m, what is the change in intergranular pressure at the bottom of the sand layer? See Figure 2.18.
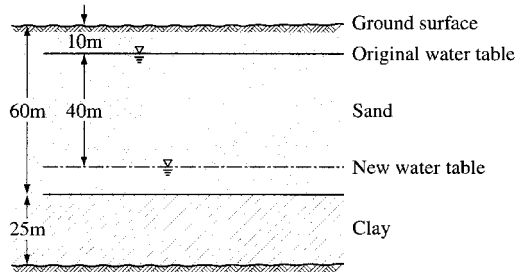
FIGURE 2.18 Subsidence problems.

Solution: The depth to the water table would then be 50 m. The total pressure at the bottom of the sand layer is $P_t = 50[(1 - 0.35)25.5 + 0.08*9.81] + 10[(1 - 0.35)25.5 + 0.35*9.81] = 1068.1$ kPa. The hydrostatic pressure is $P_h = 9.81*10 = 98.1$ kPa. The intergranular pressure is $1068.1 - 98.1 = 970.0$ kPa. The increase in intergranular pressure due to the 40 m drop in the water table is $970.0 - 683.5 = 286.5$ kPa.

Example 13: Calculate the subsidence for the situation depicted in the previous problem if the modulus of elasticity of the sand is 10,000 N/cm^2 or 10^5 kN/m^2.

Solution: The drop in the water table produces a linear increase in the intergranular pressure varying from 0.0 at the 10-m depth to 286.5 kPa at the 50-m depth. The average increment in intergranular pressure is $(0 + 286.5)/2 = 143.25$ kPa, and the settlement in the layer from 10 to 50 m is $S_{u1} = 40*143.25/10^5 = 0.0573$ m. The subsidence in the layer from 50 to 60 m is $S_{u2} = 10*286.5/10^5 = 0.0287$ m. The total subsidence is thus $0.0573 + 0.0287 = 0.086$ m.

Example 14. If the sand layer of the previous problem is underlain by a 25-m-thick layer of clay with a modulus of elasticity of 10^3 N/cm^2 or 10^4 kN/m^2, what is the total subsidence?

Solution: The increment of intergranular pressure of 286.5 kPa carries through the clay layer. The subsidence is thus $Z_3 = 25*286.5/10^4 = 0.716$ m The total subsidence of the sand and clay layers is thus $0.086 + 0.716 = 0.802$ m. It is seen that 89% of the subsidence occurs in the clay layer because its modulus of elasticity is 10 times smaller.

2.10.2 Seepage Force

When water flows horizontally through an aquifer, the flow undergoes a reduction of pressure head because of friction. Thus the pressure on the upstream side of a small element is larger than on the downstream side. The water then exerts a net force on the aquifer element. The net force in the flow direction is the *seepage force*. This force can cause lateral displacements. If the drop of the water table in a length L is Δh, the horizontal movement is $S_h = \gamma_w \Delta h L/E_h$, where E_h is the modulus of elasticity in the horizontal direction. Letting i be the slope of the water table and $\Delta h = iL$, the horizontal movement, S_h, is calculated as

$$S_h = \gamma_w i \frac{L^2}{E_h} \tag{40}$$

If there is an upward vertical flow, the head loss due to friction as the water flows into the pores results in an increase in the hydrostatic pressure. This in turn results in a decrease of the intergranular pressure. A point can be reached when the upward seepage force is large enough to carry the weight of the sand grains so that the sand or silt behaves like a liquid. It has no strength to support any weight on it. This condition is known as quicksand. It is reached when the intergranular pressure vanishes and sand loses

its bearing capacity. It can be shown that the upward hydraulic gradient necessary to produce quicksand is very close to one. (Harr, 1962; Bouwer, 1978).

2.11 Salt Water Interfaces

The fresh and saline groundwaters have densities ρ_f, and ρ_s. In coastal aquifers under natural conditions, the lighter fresh water lies over the heavier saline water and the flow is usually from the aquifer to the sea. Mixing of fresh water and salt water occurs only by molecular diffusion. Turbulent diffusion, the most effective mixing mechanism, is absent in aquifers. As a result, the mixing zone between salt water and fresh water is small compared to the thickness of the aquifer, and an abrupt, well-defined interface is usually assumed. At a point on the interface between the fresh and saline waters, the pressure of the fresh water, $\rho_f g h_f$, usually exceeds the pressure of the saline water, $\rho_s g h_s$, causing the flow from land to sea (Figure 2.19). But when pumping takes place in excess of replenishment, the drawdown of the water table creates a piezometric head in the fresh water that becomes less than in the adjacent salt water wedge. Then the saline water moves inland causing a *salt water intrusion.* The salt water may reach the well which becomes contaminated. Salt water intrusions have occurred in many coastal aquifers, for example, in the coastal aquifers of Florida, California, the Netherlands, Israel, and the South coast of England, mostly because of excessive pumping.
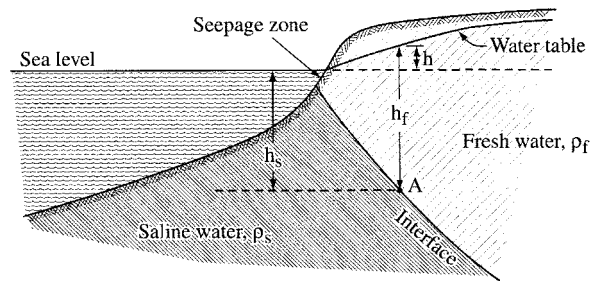
FIGURE 2.19 Coastal aquifer under natural conditions.

Assuming static conditions, the seepage zone is reduced to a point. The pressure at point A on the interface must be the same on the salt water side and on the fresh water side. Thus the depth h_s of the interface below sea level is (Figure 2.19)

$$h_s = \frac{\rho_f}{\rho_s - \rho_f} h \tag{41}$$

where h is the height of the water table above sea level. With $\rho_s/\rho_f = 1.025$ the fraction in Equation (41) is equal to 40. For a confined aquifer, h is the fresh water piezometric head. The wedge of fresh water is known as the Ghyben-Hertzberg lens after the Dutch and German scientists who first obtained Equation (41).

When fresh water is pumped from an aquifer overlying a body of salt water, the drawdown of the fresh water table around the well causes a pressure reduction on the interface. This in turn causes the interface to rise below the well. This is called *upconing.* If the salt water cone reaches the well, it will discharge a mixture of salt and fresh water. For a water table well, assuming hydrostatic conditions, a rough approximation of the height δ of the cone is (Figure 2.20)
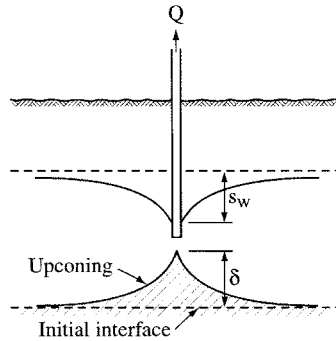
FIGURE 2.20 Upconing.

$$\delta = \frac{\rho_f}{\rho_s - \rho_f} s_w \qquad (42)$$

where s_w is the drawdawn at the well. More exact relationships are given, among others, by Bear and Dagan (1964), Dagan and Bear (1968), and Bear (1979).

2.12 Groundwater Quality

The quality of groundwater is determined by the dissolved elements and gases and by the presence of suspended solids, bacteria, and viruses. The quality of groundwater depends upon its natural and physical state and on the changes due to human activity. In its natural state, the dissolved elements and their concentrations depend on the chemical composition of the aquifer and on the travel time of the water through the rock formation. If the rock minerals are relatively soluble, slow water velocity and the ensuing long travel time result in a chemical equilibrium between the water and the rock medium. Because of the large range of flow velocities and of chemical compositions of the aquifers, there is a very wide range of compositions of the groundwater. Figure 2.21 shows the comparison of the range of concentrations of several constituents in groundwater to the concentrations of a 2.7 gm sugar cube dissolved in a 2.7 liter bottle, in a gasoline truck, and in an oil tanker.

If the groundwater is no longer fit for a specific use, such as drinking, the water is said to be *contaminated.* If the water becomes heavily contaminated it is said to be *polluted.* Chapter 13 discusses in detail the types of groundwater contaminants, and a table of drinking water standards can be found in Chapter 22 on laws and regulations in the U.S. Groundwater monitoring for water quality is discussed in Chapter 24.

Figure 2.22 illustrates the principal groundwater contaminant sources. These sources can be classified as *point*, *line*, or *nonpoint* (aereal) sources. The geometry of the source affects the geometry of the contaminant plume. Contaminants can reach the groundwater in several manners. The contaminant may be *miscible*, i.e., it is dissolved in water, *immiscible*, i.e., the water and the contaminant are in separate phases: lighter or heavier than water and the contaminant can be *adsorbed* on fine particles that are transported in suspension by the water. The miscible liquid and the water form a single phase. The movement of this single phase flow is governed by Darcy's law, discussed earlier in this chapter. Solutes are further classified as *conservative* and *nonconservative.* Conservative solutes remain stable in the groundwater: they do not react with the rock medium. Chloride solutions fall into this category. *Tracers* are conservative constituents that do not affect the viscosity and density of the water, such as fluorescine in low concentration, for example. Conservative and nonconservative transport processes are discussed in detail in Chapters 14 and 15, respectively. Multiphase flow occurs primarily in the unsaturated zone,
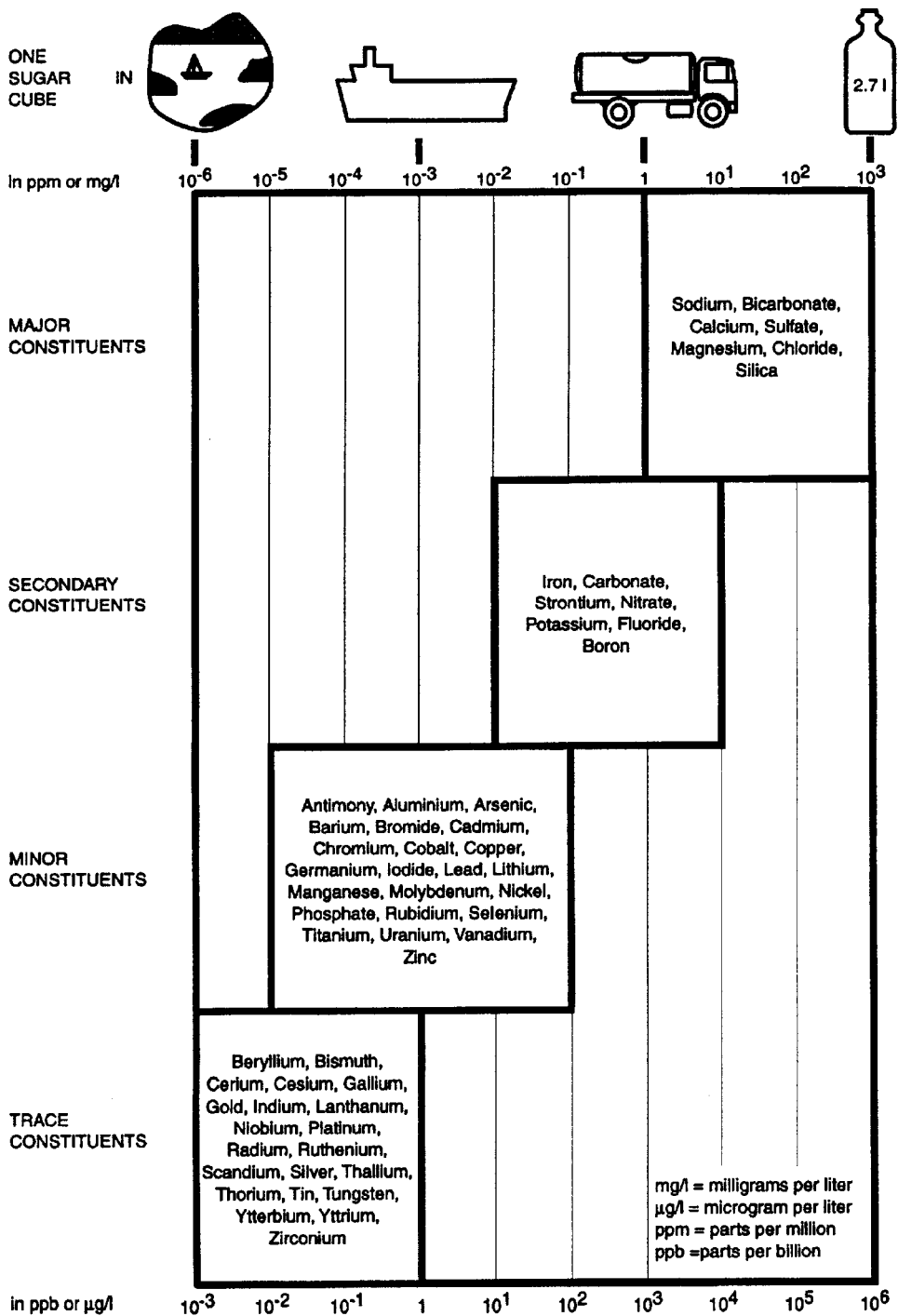
FIGURE 2.21 Examples of dissolved constituents in groundwater. (From Spitz, K. and Moreno, J. 1966. *Practical Guide to Groundwater and Solute Transport Modeling.* John Wiley & Sons, New York. With permission.)
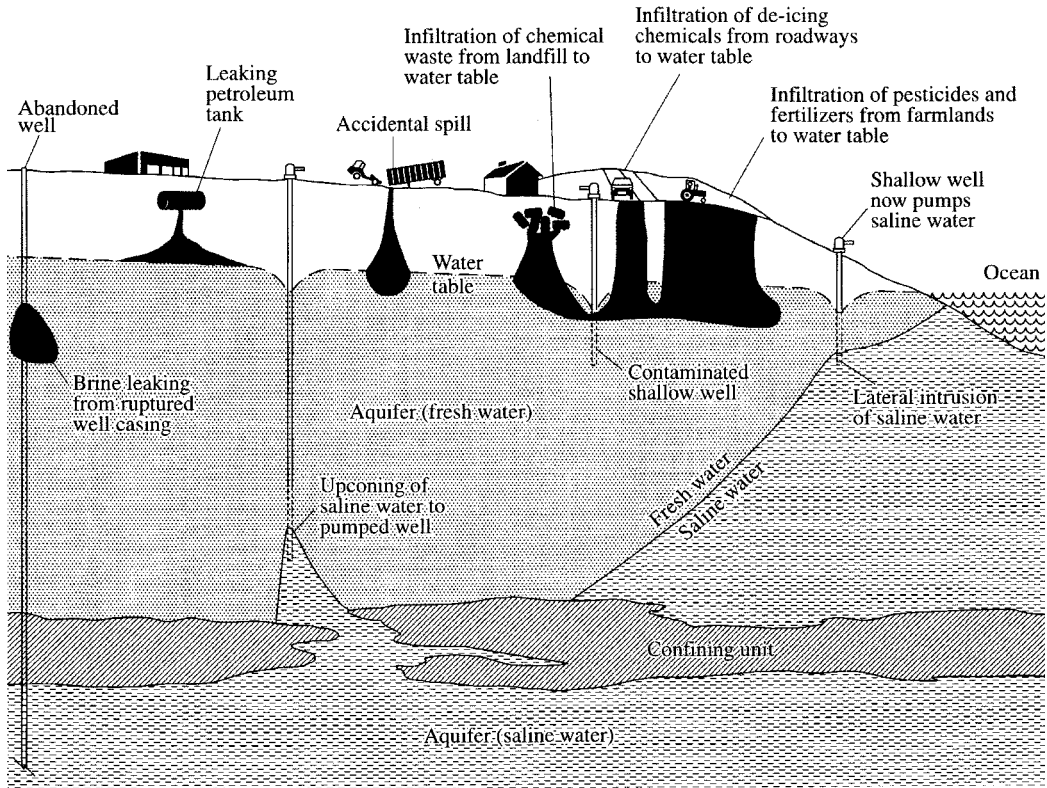
FIGURE 2.22 Sources of groundwater contamination. (From Johnston, R. H. 1986. *Water Quality Issues — Factors Affecting Ground-Water Quality.* U.S. Geological Survey, Water-Supply paper 2325. Washington, D.C.)

for example in the case of spills of hydrocarbons resulting in hydrocarbon and water phases. Multiphase flow is discussed in detail in Chapter 16 entitled *Contaminant Transport in the Unsaturated Zone.*

2.13 Transport Mechanisms of Dissolved Contaminants

The principal transport mechanisms are advection, diffusion, dispersion, sorption, and decay. These processes are discussed in their simplest one-dimensional form in this section and in more detail in Chapters 14 and 15. Other processes such as hydrolysis, volatilization, and biotransformation are not discussed in this section. Biotransformation is discussed in Chapter 26. The combined effects of advection, dispersion, and biodegradation on the transport of contaminants was visualized in the Cape Cod experiments described by LeBlanc et al. (1991). (See also Chapter 14, Section 14.5.2.)

2.13.1 Advection

Advection is the transport of solute by the bulk groundwater flow. The average pore velocity, v, is obtained by dividing the Darcy flux q (see Equation [4]) by the effective porosity n_e

$$v = \frac{q}{n_e} \tag{43}$$

The one-dimensional mass flux due to advection, F, is the product of the quantity of water flowing and the concentration of dissolved solids

$$F = v n_e C \qquad (44)$$

The change of mass of contaminant over time in a control volume

$$\left[n_e \frac{\partial C}{\partial t} \, dx \, dy \, dz \right]$$

is equal to the balance between the mass inflow and outflow of contaminant

$$\left[F \, dz \, dy - \left(F + \frac{\partial F}{\partial x} \, dx \right) dz \, dy = -\frac{\partial F}{\partial x} \, dx \, dy \, dz = -n_e \frac{\partial}{\partial x} \left(v C \right) dx \, dy \, dz \right]$$

Thus for a conservative solute in a homogeneous aquifer, the one-dimensional advective transport equation is

$$\frac{\partial C}{\partial t} = -v \frac{\partial C}{\partial x} \qquad (45)$$

where x is distance in the flow direction [L], t is time [T], C the concentration [M/L^3], and v is the advective transport velocity [L/T]. It is observed that a precise estimation of the flow velocity is needed for an accurate estimation of the transport. In sand/gravel aquifers with significant groundwater, the plume movement is dominated by advection. However, it must be recalled that Equation (45) is not valid in karstic aquifers for which Darcy's law is not applicable.

2.13.2 Diffusion

Diffusion is the flux of solute from a zone of higher concentration to one of lower concentration due to the Brownian motion of ionic and molecular species. Under steady-state conditions the diffusion flux F is described by Fick's law

$$F = -D \frac{\partial C}{\partial x} \qquad (46)$$

where D is the *diffusion coefficient* [L^2/T]. For diffusion in water, D ranges from 1×10^{-9} to 2×10^{-9} m^2/s. For diffusion in porous media, Freeze and Cherry (1979) suggest taking an effective diffusion coefficient $D^* = \omega D$, with ω ranging from 0.5 to 0.01, to account for the tortuosity of the flow paths. The change of concentration over time inside a control volume subject to diffusion flux is given by Fick's second law

$$\frac{\partial C}{\partial t} = D^* \frac{\partial^2 C}{\partial x^2} \qquad (47)$$

Consider the case of two adjacent saturated strata. The first is initially with zero concentration throughout its length $C(x,0) = 0$, and the second stratum maintains constant concentration at the interface $C(0,t) = C_0$. Because of the concentration step at the interface, diffusion takes place from the interface into the

first stratum. Far away in the first stratum, where the effect of diffusion has not yet been reached, the concentration is still zero, namely $C(\infty, t) = 0$. Crank (1956) gives the following solution of Equation (47) subject to the above boundary and initial conditions for the evolution of the concentration as a function of time and distance from the interface:

$$C(x,t) = C_0 \, erfc\left(\frac{x}{2\sqrt{D^*t}}\right) \tag{48}$$

where *erfc* is the complementary error function

$$erfc(u) = 1 - erf(u) \tag{49}$$

and *erf* is the error function

$$erf(u) = \frac{2}{\pi} \int_0^u e^{-v^2} \, dv \tag{50}$$

Tables of the error function and of the complementary error function can be found in Appendix 2.1.

Example 15: Consider a substance diffusing at the boundary of a clay layer with an effective diffusion coefficient of 10^{-9} m^2/s. Find the time it will take to obtain a concentration of 15% of the initial concentration at a distance of 20 m.

Solution: Equation (48) yields $erfc\,[x/(2\sqrt{D^*t})] = 0.15$. Interpolating from a table of complementary error functions, one finds $[x/(2\sqrt{D^*t})] = 1.02$. Thus

$$t = \frac{400}{4 \times 10^{-9}(1.02)^2} = 9.61 \times 10^{10} \text{ secs, } \approx 3047 \text{ years}$$

This shows that diffusion is a very slow mechanism, but over geological time it can become important.

The importance of diffusion increases as flow velocities decrease. Thus diffusion may be the governing transport mechanism in unfractured clays with low hydraulic conductivities. Diffusion can generally be neglected in gravel aquifers with high flow velocities. It can also be significant in fractured porous aquifers. Flow and transport processes in fractured rocks are discussed in Chapter 17.

2.13.3 Dispersion

Dispersion is the spreading of the plume that occurs along and across the main flow direction due to aquifer heterogeneities at both the small scale (pore scale) and at the macroscale (regional scale). Dispersion tends to increase the plume uniformity as it travels downstream. Factors that contribute to dispersion include: faster flow at the center of the pores than at the edges, some pathways are longer than others, the flow velocity is larger in smaller pores than in larger ones. This is known as *mechanical dispersion*. The spreading due to both mechanical dispersion and molecular diffusion is known as *hydrodynamic dispersion.*

As a conceptual example, consider an aquifer with an abrupt concentration front at $t = 0$. At that time there is a *tracer,* such as sodium chloride, with a concentration C = 1 to the left of the front, and there is no tracer to the right of the front, that is C = 0, as shown in Figure 2.23a. At time t later, the center

of the front has moved through a distance L = vt, where v is the pore flow velocity. But, due to dispersion, the tracer has spread around the center as shown in the lower part of Figure 2.23a. The plot of the S-shaped curve of the concentration at time t, C(t), is called a *breakthrough curve*. In a second experiment, a quantity of tracer is injected at point x = 0 at time t = 0. At time t_1 later, the center of the plume has moved a distance $L_1 = vt_1$ but, due to dispersion, the tracer has spread around the center with elliptical concentration contours as shown in Figure 2.23b. At time $t_2 > t_1$ the spreading has extended farther as shown in the figure.
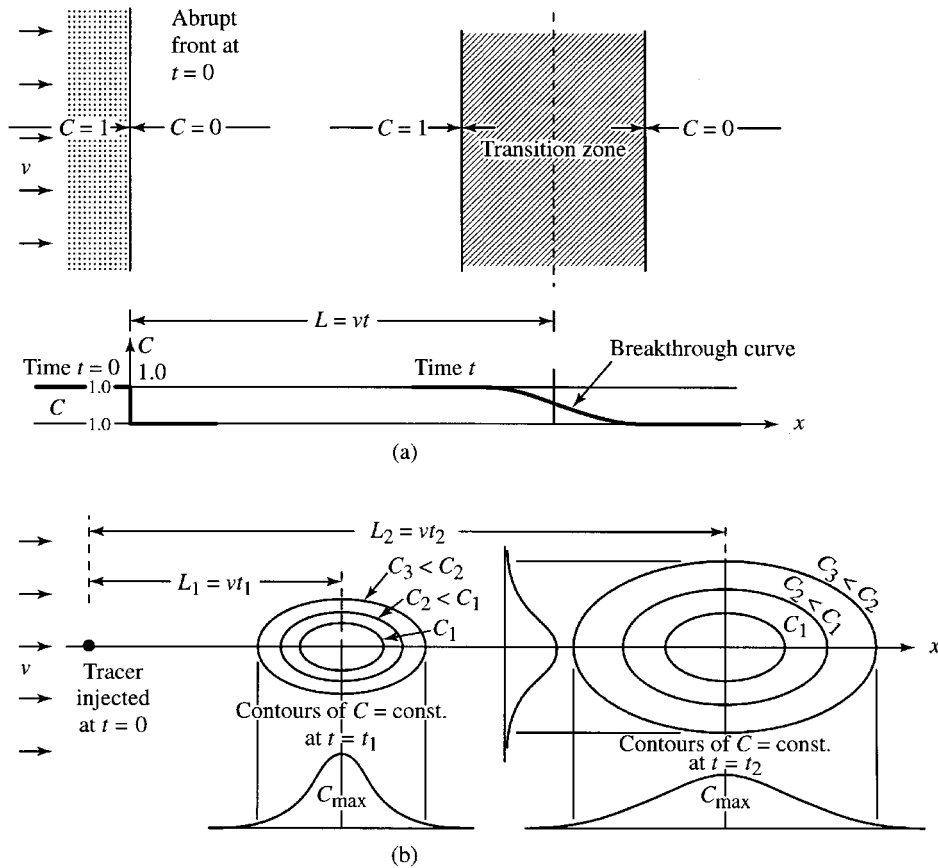
FIGURE 2.23 Longitudinal and transverse spreading due to mechanical dispersion. (From Bear, J. and Verruijt, A. 1987. *Modeling Groundwater Flow and Pollution.* Reidel, Dordrecht, The Netherlands. With permission.)

If one considers a *representative elementary volume* (REV) (see Chapter 3, section on macroscopic approach), dispersion can be described by Fick's law (46). The *dispersion coefficient*, D_L, replaces the diffusion coefficient D^* and becomes a phenomenological coefficient that combines the effects of diffusion and dispersion. As mechanical dispersion is more pronounced in the longitudinal direction than in the transverse direction, a *longitudinal* dispersion coefficient D_L and a *transverse* dispersion coefficient, D_T, are introduced. The longitudinal and the transverse dispersion-diffusion coefficients are defined as

$$D_L = \alpha_L v + D^* \qquad \text{and} \qquad D_T = \alpha_T v + D^* \qquad (51a,b)$$

where α_L is the longitudinal dispersivity [L], α_T is the transverse dispersivity [L], and v is the pore velocity [L/T].

The use of the Fickian theory to describe dispersion requires that the dispersion coefficients be travel-distance dependent or time dependent. A rough approximation based on averaging published data is (Gelhar et al., 1992)

$$\alpha_L \approx 0.1L \tag{52}$$

where L is the length of the flow path. Another estimate for flow lengths less than 3500 m was given by Neuman (1990) as

$$\alpha_L \approx 0.0175 L^{1.46} \tag{53}$$

where L is the length of the flow path (m). The transverse dispersivity α_T is typically 1/10 to 1/100 of the longitudinal dispersivity α_L.

The combined advection-dispersion equation in one dimension is thus

$$\frac{\partial C}{\partial t} = -v\frac{\partial C}{\partial x} + D_L \frac{\partial^2 C}{\partial x^2} \tag{54}$$

The solution of Equation (54) subject to the fixed step conditions:

$$\text{Initial condition} \qquad C(x,0) = 0 \qquad x \geq 0$$

$$\text{Boundary condition} \qquad C(0,t) = C_0 \qquad t \geq 0$$

$$\text{Boundary condition} \qquad C(\infty,t) = 0 \qquad t \geq 0$$

has been given by Ogata and Banks (1961) as

$$C(x,t) = \frac{C_0}{2}\left[erfc\left(\frac{x - vt}{2\sqrt{D_L t}}\right) + \exp\left(\frac{vx}{D_L}\right) erfc\left(\frac{x + vt}{2\sqrt{D_L t}}\right) \right] \tag{55}$$

where x is the distance from the injection point. The argument of the exponential is the *Peclet number* [$P_e = vx/D_L$]. The Peclet number is a measure of the ratio of the rate of transport by advection to the rate of transport by diffusion. For large Peclet numbers ($P_e > 100$), the advection dominates and the second term in the right-hand side becomes negligible. Figure 2.24 shows the behavior of the solution (55).

If instead of a fixed step function, a line source with continuous injection into the aquifer is considered as, for example, the leakage from a canal, Sauty (1980) found that in the solution (55) the + sign in front of the exponential in the right-hand side is replaced by a – sign. For large Peclet numbers, the fixed step solution and the line source solution are essentially identical as the second term in the right-hand side of (55) is negligible. Further discussion and other boundary conditions are considered in Fetter (1993).

Example 16: An aquifer has a hydraulic conductivity of 2×10^{-5} m/s, a hydraulic gradient of 0.003 m/m and effective porosity $n_e = 0.2$, and an effective diffusion coefficient $D^* = 0.5 \times 10^{-9}$ m^2/s. A chloride solution with a concentration of 500 mg/l penetrates in the aquifer along a line source. Find the chloride concentration at a distance of 20 m from the point of entry, after a period of two years.

Solution: From Darcy's law, the pore flow velocity is
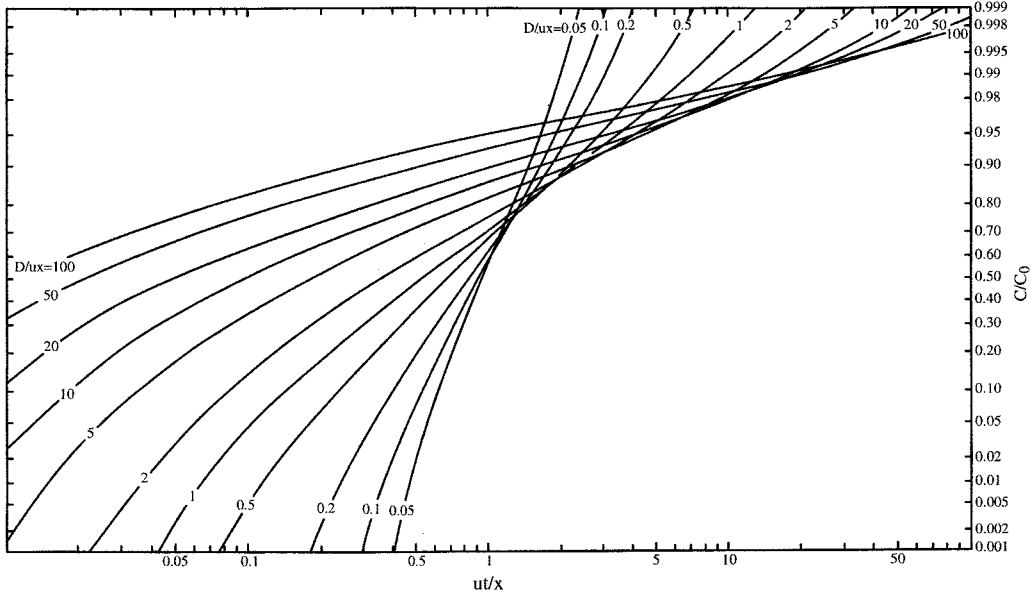
FIGURE 2.24 Solution of the one-dimensional advection-dispersion equation. (From Ogata, A. and Banks, R. B. 1961. *A Solution of the Differential Equation of Longitudinal Dispersion in Porous Media*, U.S. Geological Survey, Professional Paper 411A. U.S. Government Printing Office, Washington, DC.)

$$v = \frac{K}{n_e}\, dh/dx = \frac{2 \times 10^{-5} * 0.003}{0.2} = 3 \times 10^{-7} \text{ m/s}$$

Using Neuman's approximation (53), the longitudinal dispersivity is $\alpha_L = 0.0175 * 20^{1.46} = 1.388$ m. The coefficient of longitudinal dispersion-diffusion, from (51a) is $D_L = 1.388 * 3.0 \times 10^{-7} + 0.5 \times 10^{-9} = 4.170 \times 10^{-7}$ m²/s. The Peclet number is $P_e = vL/D_L = 3 \times 10^{-7} * 20/4.170 \times 10^{-7} = 14.387$. As $P_e < 100$, both terms in the right-hand side of (55) need to be considered. Thus after 2 yrs = 2*365*24*60*60 = 6.3072×10^7 s and at distance of 20 m the concentration is obtained from (55) as

$$C = \frac{500}{2}\left[erfc\left(\frac{20 - 3 \times 10^{-7} * 6.3072 \times 10^7}{2\sqrt{4.170 \times 10^{-7} * 6.3072 \times 10^7}}\right) + \exp 14.384 * erfc\left(\frac{20 + 3 \times 10^{-7} * 6.3072 \times 10^7}{2\sqrt{4.170 \times 10^{-7} * 6.3072 \times 10^7}}\right)\right]$$

$$= 250\left(erfc\, 0.105 + 14.384 * erfc\, 3.795\right) \approx 250 * 0.881984 = \textbf{220.5 mg/l}$$

as the last *erfc* (.) is negligible as seen in Appendix 2.1.

2.13.4 Sorption

Sorption refers to the exchange of molecules and ions between the solid phase and the liquid phase. It includes adsorption and desorption. *Adsorption* is the attachment of molecules and ions from the solute to the rock material. Adsorption produces a decrease of the concentration of the solute or, equivalently, causes a *retardation* of the contaminant transport compared to the water movement. *Desorption* is the release of molecules and ions from the solid phase to the solute.

The relationship between the solute concentration in the adsorbed phase and in the water phase is called a *sorption isotherm*. The simplest expression is the *linear isotherm*

$$C_a = K_d C \tag{56}$$

where C_a is the sorbed concentration as mass of contaminant per mass of dry rock matrix [dimensionless], C is the dissolved concentration in mass of contaminant per volume of water [M/L^3] and K_d is the *distribution coefficient* [L^3/M]. This expression implies that there is an *equilibrium* between the adsorbed concentration and the dissolved concentration. This can be assumed when the adsorption process is fast compared to the advection of contaminant.

The adsorption causes a retardation in the migration of contaminants compared to the advection. The contaminant transport gets more retarded as the fraction adsorbed increases. This effect can be described by a *retardation factor*, R_a, which for a linear isotherm, is

$$R_a = 1 + \frac{(1-n)\rho_s}{n} K_d \tag{57}$$

where n is the porosity and ρ_s is the density of the solids. The retardation coefficient may take values from 1 to 10,000. The velocity of the solute front v_c (where the concentration is $1/2$ that of the original concentration) is given by

$$v_c = \frac{v}{R_a} \tag{58}$$

The one-dimensional advection-dispersion equation then becomes

$$\frac{\partial C}{\partial t} = -\frac{v}{R_a} \frac{\partial C}{\partial x} + \frac{D_L}{R_a} \frac{\partial^2 C}{\partial x^2} \tag{59}$$

where the term on the left side represents the change in storage of contaminant in the control volume, the first term on the right-hand side represents the retarded advective inflow-outflow, and the last term represents the retarded diffusion and dispersion. Solutions of the advection/dispersion/adsorption equation have been given by van Genuchten (1981).

The linear isotherm also applies to the case of organic compounds dissolved in groundwater. If the aquifer or soil contains at least 1% of organic carbon, then

$$K_d = K_{oc} f_{oc} \tag{60}$$

where K_{oc} is the organic carbon partition coefficient for the organic solute, and f_{oc} is the fraction of organic carbon by weight in the soil. Methods for estimating K_{oc} from solubility data and from the octanol-water partition coefficient can be found, for example, in Fetter (1993) and in Spitz and Moreno (1996). (See Table 13.9, Chapter 13 on groundwater contaminants; see also Section 2.15).

2.13.5 Radioactive Decay and Degradation

In the previous section it was assumed that the adsorption was fast compared to the advection of the contaminant. If, instead, the reaction is slow compared to the travel time and chemical equilibrium is not attained, then it is necessary to describe the kinetics of the reaction. The simplest model is the irreversible (i.e., the solute cannot be desorbed) first order model:

$$\frac{\partial C}{\partial t} = -\lambda C \tag{61}$$

where λ is the first order decay rate constant $[T^{-1}]$. This relation also applies to radioactive decay and degradation processes. Equation (61) integrates as

$$C = C_0 e^{-\lambda t} \tag{62}$$

where C_0 is the concentration at time $t = 0$ and

$$\lambda = \frac{\ln 2}{T_{1/2}} \tag{63}$$

where $T_{1/2}$ is the half-life $[T]$ of the radioactive isotope or of the degraded contaminant. Some values of radioactive and organic half-lives can be found in Spitz and Moreno (1996). If decay or degradation occurs, the quantity $-\lambda C$ is appended to the right-hand side of (59).

2.14 Monitoring, Site Remediation, and Landfills

Monitoring wells are necessary to measure the elevations of the water table or of the piezometric level, to collect samples of water for chemical analyses, to collect samples of nonaqueous-phase liquids, to provide access for geophysical instruments, etc. The techniques used in groundwater monitoring and soil sampling are discussed in detail in Chapter 24 and the geophysical exploration techniques requiring wells and other types are presented in Chapters 10 and 11.

Site remediation must consider at least source control and treatment of contaminated water and/or soil. Source control is necessary to prevent continuing discharge of contaminants to the subsurface or to the groundwater. Treatment may be necessary to remove or substantially decrease the concentration of contaminants. Techniques for the proper design of landfills have been developed to eliminate or minimize the leakage to the vadose zone or to the groundwater. These activities are regulated by legislations such as the Resource Conservation and Recovery Act (RCRA) and the Comprehensive Environmental Response, Compensation and Liability Act (CERCLA).

These topics of regulations, landfills, and site remediation are the subject of Chapters 22, 23, and 25, respectively.

2.15 Parameter Values

In the previous sections a number of equations have been presented that describe the flow of groundwater and the transport of contaminants. The application of these equations and of the models which are based upon them requires the use of a number of parameters. In general the hydrogeological parameters and the pollutant characteristics exhibit great variability and hence uncertainty. In many applications, data may not be readily available. It may then be useful to have access to a database that brings together most of the previous experience, at least until field experiments can be conducted.

Spitz and Moreno (1996) present an extensive compilation of hydrogeological and pollutant transport parameters and list the literature sources of these data: porosity, p. 342-344; specific yield, p. 345; horizontal hydraulic conductivity, p. 346-350; vertical hydraulic conductivity, p. 351-352; specific storage, p. 353; unsaturated hydraulic conductivity relationships, p. 353-354; moisture content relationships, p. 354; molecular diffusion coefficients, p. 368-369; mechanical dispersion coefficients, p. 370-371; field-scale dispersivities, p. 372-379; K_d values for metals, p. 380-389; K_d values for organics, p. 390-392; organic carbon content of sediments, f_{oc}, p. 392; empirical K_{oc} relationships, p.393-394; radioactive half-lives, p. 395; half-lives for organics, p. 396-409.

For Further Information

This chapter deals with elementary groundwater flow and transport problems. More advanced aspects are defered to subsequent chapters. The following books give an introductory treatment of groundwater problems (see References for complete citations):

Heath, R. C. (1995) provides an excellent, practical, and well-illustrated introduction to groundwater flow, well hydraulics, well tests and groundwater pollution.

Lowman, S. W. (1972) gives a good introductory treatment of hydrologic properties of water bearing materials, flowing wells, aquifer tests by well methods and by areal methods.

National Research Council (1984) provides a well-documented, non-mathematical introduction to groundwater contamination, including case studies.

Palmer, C. M. (1996) gives a non-mathematical introduction to contaminant hydrogeology. A practical approach to completing investigations and the basics of collecting data are presented.

Price, M. (1996) gives an excellent physical, non-mathematical, and exceptionally well written description of the phenomena associated with groundwater occurrence, movement, pumping, and pollution.

There is a number of excellent textbooks on the subject of this chapter. Here we select a few:

Bedient et al. (1994) give a good introduction on groundwater hydraulics and then concentrate on groundwater quality, pollution of groundwater, contaminant transport, and groundwater management and remediation.

Fetter C. W. (1994) provides a general treatment of groundwater occurrence, movement, contamination, development, and management and models.

Fetter C. W. (1993) focuses on the transport of contaminants in the saturated and unsaturated zones and gives practical information on monitoring and site remediation.

Freeze and Cherry (1979) have written a classical text that covers the fundamentals of groundwater geology, the flow and chemical evolution of groundwater, and groundwater contamination.

Todd, D. K. (1980) has written a good practical text on groundwater hydrology.

At a more advanced level, there are also excellent textbooks among which we have selected the following:

Bear J. (1979) gives an excellent mathematical treatment of the laws and equations that describe the flow and management of groundwater.

Bear J. and Verruijt A. (1987) are concerned with the movement and accumulation of groundwater and pollutants in aquifers, and the construction of conceptual and mathematical models. A number of computer programs written in BASIC are included.

De Marsily G. (1986) presents a very lucid and classical treatment of the mathematics of groundwater flow and contaminant transport.

Domenico P. A. and Schwartz F. W. (1990) give an advanced treatment of physical and chemical hydrogeology.

If one needs a handy bibliographical reference of publications in the groundwater field, there is van der Leeden (1991). He lists approximately 5600 references on general bibliographies, journals, texts, handbooks and dictionaries, and references by subjects such as history, environment, geophysical exploration, well logging, hydraulics of soils and aquifers, theory of groundwater flow, pumping, well maintenance, pumping equipment, tracers, water quality, contamination, salt-water intrusion, models, laws and regulations, management, etc.

References

Bear, J. 1972. *Dynamics of Fluids in Porous Media*. American Elsevier, New York.
Bear, J. 1979. *Hydraulics of Groundwater*. McGraw-Hill Inc., New York.

Bear, J. and Dagan, G. 1964. Some exact solutions of interface problems by means of the hodograph method. *Jour. Geophysical Research.* 69, 2, 1563-1572.

Bear, J. and Verruijt, A. 1987. *Modeling Groundwater Flow and Pollution.* Reidel, Dordrecht, The Netherlands.

Bedient, P. B., Rifai, H. S., and Newell, C. J. 1994. *Ground Water Contamination: Transport and Remediation*, Prentice Hall, Englewood Cliffs, NJ.

Bögli, A. 1980. *Karst Hydrology and Physical Speleology.* Springer-Verlag, Berlin.

Bouwer, H. 1969. Theory of seepage from open channels, in *Advances in Hydroscience.* Vol. 5, 121-172, Ven Te Chow, Ed., Academic Press, New York.

Bouwer, H. 1970. Groundwater recharge design for renovating waste water, *Jour. of the Sanitary Engineering Division, Proceedings of the ASCE.* Vol. 96, No. SA1, 59-74.

Bouwer, H. 1978. *Groundwater Hydrology.* McGraw-Hill Inc., New York.

Chow, Ven Te. 1959. *Open-Channel Hydraulics.* McGraw-Hill Inc., New York

Crank, J. 1956. *The Mathematics of Diffusion.* Oxford University Press, Oxford.

Darcy, H. 1856. *Les Fontaines Publiques de la Ville de Dijon.* V. Dalmont, Paris.

Dagan, G. and Bear, J. 1968. Solving the problem of interface upconing in coastal aquifers by the method of small perturbations. *Jour. of Hydraulic Research.* 6,1,15-44.

Daugherty, L., Franzini, J. B., and Finnemore, E. J. 1985. *Fluid Mechanics with Engineering Applications.* 8th ed. McGraw-Hill Inc., New York.

Davis, S. N. and De Wiest, R. J. N. 1966. *Hydrogeology.* Wiley, New York.

de Marsily, G. 1986. *Quantitative Hydrogeology — Groundwater Hydrology for Engineers.* Academic Press, Orlando, FL.

Domenico, P. A. and Schwartz, F. W. 1990. *Physical and Chemical Hydrogeology.* John Wiley & Sons, New York.

Dupuit, J. 1863. *Études Théoriques et Pratiques sur le Mouvement des Eaux dans les Canaux Découverts et à Travers les Terrains Perméables.* Dunod, Paris.

Fetter, C. W. 1994. *Applied Hydrogeology.* Prentice Hall, Englewood Cliffs, NJ.

Fetter, C. W. 1993. *Contaminant Hydrogeology.* Macmillan, New York.

Forchheimer, P. 1930. *Hydraulik.* Teubner Verlagsgesellschaft, Stuttgart.

Franzini, J. B. and Finemore, E. J. 1997. *Fluid Mechanics with Engineering Applications.* McGraw-Hill Inc., New York.

Freeze, R. A. and Cherry, J. A. 1979. *Groundwater.* Prentice-Hall, Englewood Cliffs, NJ.

Gelhar, L. W. 1986. Stochastic subsurface hydrology from theory to application. *Water Resources Research.* 22, 9, 1358-1458.

Gelhar, L. W., Welty, C., and Rehfeldt, K. R. 1992. A critical review of data on field-scale dispersion in aquifers. *Water Resources Research.* 28, 7, 1955-1974.

Harr, M. E. 1962. *Ground Water and Seepage.* McGraw-Hill Inc., New York.

Heath, R. C. 1995. *Basic Ground-Water Hydrology.* U.S. Geological Survey, Water-Supply paper 2200, seventh printing, Denver, CO.

Hibbs, B. J. 1997. Ground-water modeling and land subsidence prediction in the Harris-Galveston coastal subsidence district, Texas, in *Proceedings of the* 24th *Annual Water Resources Planning and Management Conference, ASCE.* D. H. Merritt, Ed.

Hubbert, M. K. 1953. Entrapment of petroleum under hydrodynamic conditions. *Bulletin of American Association of Petroleum Geologists.* Vol. 37, 1954-2026.

Hubbert, M. K. 1987. Darcy's law: its physical theory and application to entrapment of oil and gas, in *History of Geophysics.* Vol. 3, 1-26, C. S. Gillmor, Series Editor, AGU, Washington, DC.

Jumikis, A. R. 1984. *Soil Mechanics.* Robert E. Krieger Co., Malabar, FL.

Lowman, S. W. 1972. *Ground-Water Hydraulics.* Geological Survey Professional Paper 708, U.S. Government Printing Office, Washington, DC.

Karplus, W. J. 1958. *Analog Simulation.* McGraw-Hill Inc., New York.

LeBlanc, D. R., Garabedian, S. P., Hess, K. M., Gelhar, L. W., Quadri, R. D., Stollenwerk, K. G., and Wood, W. W. 1991. Large scale natural gradient tracer test in sand and gravel, Cape Cod, Massachusetts, 1. Experimental design and observed tracer movement, *Water Resources Research.* 27, 5, 895-910.

Mariño, M. A. 1975. Artificial groundwater recharge, II rectangular recharging area, *Jour. of Hydrology.* 26, 29-37.

National Research Council. 1984. *Groundwater Contamination.* National Academy Press, Washington, DC.

Neuman, S. P. 1990. Universal scaling of hydraulic conductivities and dispersivities in geologic media. *Water Resources Research.* 26, 8, 1749-1758.

Ogata, A. and Banks, R. B. 1961. *A Solution of the Differential Equation of Longitudinal Dispersion in Porous Media.* U.S. Geological Survey, Professional Paper 411A. U.S. Government Printing Office, Washington, DC.

Palmer, C. M. 1996. *Contaminant Hydrogeology.* Lewis Publishers, Boca Raton, FL.

Poland, J. F. 1984. Ed. *Guidebook to Studies of Land Subsidence Due to Ground-Water Withdrawal.* UNESCO, International Hydrological Programme, Working group 8.4, Paris.

Polubarinova-Kochina, P. YA. 1962. *Theory of Ground Water Movement.* Translated from the Russian by J.M. Roger De Wiest, Princeton University Press, Princeton, NJ.

Price, M. 1996. *Introducing Groundwater.* Chapman Hall, London.

Sauty, J.-P. 1978. Identification des paramètres du transport hydrodispersif dans les aquifères par interprétation de traçage en écoulement cylindrique convergent ou divergent, *Journal of Hydrology.* 39, $^3/_4$:69,-103.

Saquib, M. N., Davis, L. C., Khan, A. Q., and Taghavi, S. A. 1995. Imperial County Groundwater Evaluation: Model Applications, *Proceedings* 22*nd. Annual Conference. Integrated Water Resources for the* 21*st. Century.* Cambridge, MA. 929-932.

Schneebeli, G. 1955. Expériences sur la limite de validité de la loi de Darcy et l'apparition de la turbulence dans un écoulement de filtration. *La Houille Blanche.* Vol. 10, 141-149.

Spitz, K. and Moreno, J. 1996. *A Practical Guide to Groundwater and Solute Transport Modeling.* John Wiley & Sons, New York.

Subramanya, K., Madhav, M. R., and Mishra, G. C. 1973. Studies on seepage from canals with partial lining. *Jour. Hydraulics Division, Proceedings of the ASCE.* Vol. 99, No. HY12, 2333-2351.

Taghavi, S. A., Saquib, M. N., Khan, A. Q., and Davis, L. C. 1995. Imperial County Groundwater Evaluation: Model Development and Calibration, *Proceedings* 22*nd. Annual Conference. Integrated Water Resources for the* 21*st. Century.* Cambridge, MA. 933-936.

Todd, D. K. 1980. *Ground Water Hydrology.* 2nd ed. John Wiley & Sons, Inc., New York.

Van der Leeden, F. 1991. *Geraghty & Miller's Groundwater Bibliography.* fifth ed. Water Information Center, Plainview, NY.

van Genuchten, M. T. 1981. Analytical solutions for chemical transport with simultaneous adsorption, zero-order production and first order decay. *Journal of Hydrology.* 49, 213-233.

White, W. B. 1988. *Geomorphology and Hydrology of Karst Terrains.* Oxford University Press, New York.

Yussuff, S. M., Chauhan, H. S., Kumar, M., and Srivastava, V. K. 1994. Transient canal seepage to sloping aquifer, *Jour. Irrigation and Drainage Engineering. ASCE,* Vol. 120, No.1, 97-109.

Glossary

Adsorption   Adhesion of solute molecules and ions to rock or soil material.

Advection   Mass transport of solute by the gross movement of groundwater.

Conservative Solute   Solute that remains stable, does not react with rock or soil material.

Contaminant   Substance that causes contamination.

Contamination   Degradation of groundwater quality that renders it unfit for domestic consumption.

Darcy's Law   An equation that relates the gross flow velocity (i.e., the discharge divided by the gross cross-section of an aquifer segment) to the product of the hydraulic conductivity and the gradient of the total head.

Datum   An arbitrary reference elevation from which hydraulic heads are measured.

Desorption   Removal of molecules or ions from rock or soil.

Diffusion   The flux of solute from areas of higher concentration to areas of lower concentration due to random molecular motion.

Dispersion   The spread of solute due to heterogeneities of the pore sizes and shapes (mechanical dispersion) heterogeneities in the aquifer (macrodispersion).

Dispersion Coefficient   A coefficient in Fick's law that relates mass flux to concentration gradient.

Dispersivity   A constant of dispersion which, when multiplied by the pore flow velocity yields the dispersion coefficient.

Dupuit-Forchheimer Assumption   The assumption of primarily horizontal flow, neglecting unimportant vertical flow.

Effective Porosity   The part of the porosity that is available for the fluid flow.

Elevation Head   Difference in elevation between a point in a flow field and an arbitrary reference datum, the latter often being taken as mean sea level.

Equipotential Line   Line of equal potential, used in the flow net.

Fick's Law   Equation that describes the diffusion or dispersion of solutes.

Force Potential   The product of the acceleration of gravity and the total head, it represents the total energy per unit of mass.

Flow Net   A network of streamlines and equipotential lines that intersect at right angles.

Half-life   The time required for the concentration of a solute to be reduced to half its initial value by radioactive decay or biodegradation.

Head   See Elevation Head, Pressure Head, Velocity Head, and Total Head

Hydraulic Gradient   The rate of change of the piezometric head with displacement in a given direction.

Karst   Geological formation characterized by features associated with dissolution and collapse of carbonate rocks such as underground drainage, caves, sinkholes, and deep gullies. Named after the Karst plateau, a barren limestone plateau in Western Slovenia,

Leakage   Seepage of water through a semipermeable layer called an aquitard.

Leaky Aquifer   An aquifer into which there is seepage from an overlying formation.

Linear Isotherm   A chemical equilibrium relationship in which the concentration of the adsorbed solute in the solid phase is assumed to be proportional to the concentration in the water phase.

Longitudinal Dispersion Coefficient   Dispersion coefficient in the flow direction.

Monitoring Well   A non-pumping well used to measure water levels or to obtain water samples for chemical analysis.

Organic   material containing a carbon compound often associated with hydrogen, not necessarily derived from a living organism.

Peclet Number   A dimensionless quantity that expresses the relative importance of advection and dispersion of solutes.

Piezometric Head   The sum of the elevation head and the pressure head.

Pollutant   Substance that causes pollution.

Pollution   Excessive contamination of the environment; in this chapter specifically contamination of groundwater that renders it unfit for human consumption as a result of human or natural activities.

Pressure Head   The ratio of the fluid pressure intensity to the fluid specific weight; has the dimension of length.

Retardation Factor   Ratio of transport velocity of nonreacting solute to transport velocity of solute reacting with the solid phase.

Reynolds Number   A dimensionless quantity that expresses the relative importance of inertia forces and viscous forces in a flow system. A small Reynolds number is associated with laminar flow; a large Reynolds number is associated with turbulent flow.

Salt Water Interface   A surface forming a common boundary between adjacent salt water and fresh water.

Sorption   includes adsorption and desorption.

Specific Discharge   The flow rate through a cross-section of an aquifer divided by the area of that cross-section.

Streamline   A line everywhere tangent to the flow velocity vector, also called a flow line. It is used in the construction of flow nets.

Suction   A negative pressure head found in the unsaturated zone.

Total Head   The sum of the elevation head, pressure head, and velocity head. In flow through porous media the velocity head is small and often neglected in the calculation of the total head.

Tracer   A conservative solute that is used to track the path of groundwater movement.

Transverse Dispersion Coefficient   Dispersion coefficient in the direction perpendicular to the flow direction.

Velocity Head   The kinetic energy of the flow per unit weight of fluid; has the dimension of length.

Velocity Potential   The product of the hydraulic conductivity and the total head. By virtue of Darcy's law, the negative of the gradient of the velocity potential is a flow velocity (specific discharge or Darcy velocity).

Water Balance   An accounting of the inflows and outflows in a fluid control volume.

Appendix 2.1

Values of the Error Function and Complementary Error Function

x	erf x	erfc x	x	erf x	erfc x
0	0	1.0	1.0	0.842701	0.157299
0.05	0.056372	0.943628	1.1	0.880205	0.119795
0.1	0.112643	0.887537	1.2	0.910314	0.089686
0.15	0.167996	0.823004	1.3	0.934008	0.065992
0.2	0.222703	0.777297	1.4	0.952285	0.047715
0.25	0.276326	0.723674	1.5	0.996105	0.033895
0.3	0.328627	0.671373	1.6	0.976348	0.023652
0.35	0.379382	0.620618	1.7	0.983790	0.016210
0.4	0.428392	0.571608	1.8	0.989091	0.010909
0.45	0.475482	0.524518	1.9	0.992790	0.007210
0.5	0.520500	0.479500	2.0	0.995322	0.004678
0.55	0.563323	0.436677	2.1	0.997021	0.002979
0.6	0.603856	0.396144	2.2	0.998137	0.001863
0.65	0.642029	0.357971	2.3	0.998857	0.001143
0.7	0.677801	0.322199	2.4	0.999311	0.000689
0.75	0.711156	0.288844	2.5	0.999593	0.000407
0.8	0.742101	0.257899	2.6	0.999764	0.000236
0.85	0.770668	0.229332	2.7	0.999866	0.000134
0.9	0.796908	0.203092	2.8	0.999925	0.000075
0.95	0.820891	0.179109	2.9	0.999959	0.000041
			3.0	0.999978	0.000022

$$erf\, x = \left(\frac{4}{\pi}\right)^{1/2} \int_0^x \exp\left(-z\right)^2 dz$$

$$erf\,(0) = 0; \quad erf\,(\infty) = 1; \quad erf\,(-x) = -erf\,(x); \quad 1 - erf\,(x) = cerf\,(x)$$

3

Two- and Three-Dimensional Flow of Groundwater

F. De Smedt
Free University, Brussels

3.1 Fundamentals

3.1.1 Introduction

The objective of this chapter is to show the logical development of equations that explain and predict the movement of groundwater in two and three dimensions as observed in nature. Flow of groundwater is a special case of fluid flow in porous media and is governed by the laws of physics, in particular the laws of fluid mechanics. Fluid mechanics deals with the motion of fluids and with the forces exerted on solid bodies in contact with fluids. The fundamental principles are *conservation of mass* and *Newton's second law of motion*. It will be shown that these fundamental principles are sufficient to explain groundwater flow.

The purpose of this chapter is to obtain flow equations that enable us to calculate the state variables of groundwater in each point of the flow domain and, if necessary, also in time. Therefore, groundwater flow equations are expressed in the form of partial differential equations, with the spatial coordinates and time as independent variables. Hence, the variables that describe the state of groundwater are explicit functions of position and time, as for instance the groundwater potential

$$h = h(x, y, z, t) \tag{1}$$

where h represents the groundwater potential, the space position is denoted by Cartesian coordinates x,y,z, and the time is represented by the symbol t; x,y,z and t are independent variables, while other variables, such as h, are considered dependent, i.e., with each set of values (x,y,z,t) variables can be associated, as $h(x,y,z,t)$.

Variables have dimensions. Whenever a new variable is introduced, the dimensions will be indicated between brackets []. The symbols used are: [L] for length, [T] for time, [M] for mass, and [$F = ML/T^2$] for force. For practical applications the dimensions can be substituted by any consistent set of units. A logical choice would be the International Systems of Units, i.e., respectively meters, seconds, kilograms, and Newtons. However, because groundwater flow is usually extremely slow, often days are used instead of seconds for the time dimension.

3.1.2 Continuity Equation

The first fundamental law governing groundwater flow is the continuity equation, which expresses the principle of mass conservation. Consider an *elementary control volume* of soil centered around a point with Cartesian coordinates (x,y,z) as shown in Figure 3.1. It is customary and convenient to choose the z-axis vertical and pointing upward in the positive direction. The size of the elementary volume is Δx, in the x-direction, Δy in the y-direction, and Δz in the z-direction. At a certain time instant, t, the mass of groundwater, M, present in the elementary control volume is given by

$$M = \rho\theta\,\Delta x\,\Delta y\,\Delta z \tag{2}$$

where θ is the volumetric water or moisture content of the porous medium, with dimensions [L^3/L^3], and ρ the density of water with dimensions [M/L^3]. This quantity of water can change when groundwater enters or leaves the control volume through the sides. The principle of mass conservation implies that the net result of inflow minus outflow is balanced by the change in storage versus time, or
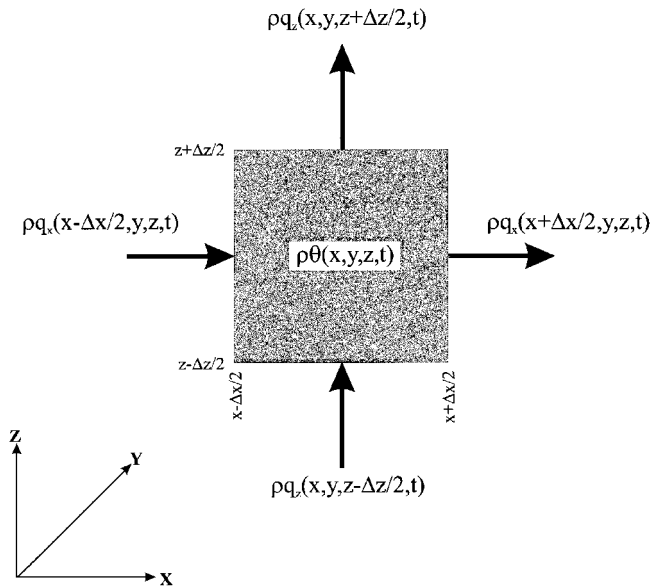
FIGURE 3.1 Mass conservation in a reference elementary volume; the actual volume is three dimensional, but for clarity it is depicted only in two dimensions, *x* and z.

$$\frac{\partial M}{\partial t} = \text{inflow} - \text{outflow} \tag{3}$$

Hence, it is necessary to calculate the groundwater flows through the sides of the elementary control volume in order to evaluate the net result of inflow minus outflow. The amount of groundwater flow is denoted by means of the flux, $\boldsymbol{q}$, which is the volumetric discharge or flow rate per cross-sectional area. The flux is a vector quantity with components along the x, y and z-directions: $\boldsymbol{q} = (q_x, q_y, q_z)$; the dimensions are $[L^3/T/L^2 = L/T]$. For instance, the mass inflow of groundwater along the left side, situated at $x - \Delta x/2$ of the control volume, is given by

$$\rho q_x \left(x - \Delta x/2, y, z, t \right) \Delta y \Delta z \tag{4}$$

Because Δx is small, q_x at position $x - \Delta x/2$ can be approximated by a *Taylor series expansion*, where only zero and first-order terms are maintained, such that this inflow can be calculated as

$$\left(\rho q_x - \frac{\Delta x}{2} \frac{\partial \rho q_x}{\partial x} \right) \Delta y \Delta z \tag{5}$$

In this expression ρq_x and its derivative versus x are evaluated at the center of the control volume. Similar expressions can be established for the other sides; for instance at the right side, at position $x + \Delta x/2$ of the control volume, the groundwater outflow is given by

$$\rho q_x \left(x + \Delta x/2, y, z, t \right) \Delta y \Delta z \approx \left(\rho q_x + \frac{\Delta x}{2} \frac{\partial \rho q_x}{\partial x} \right) \Delta y \Delta z \tag{6}$$

Hence, the total inflow minus outflow can be calculated as

$$\left(\rho q_x - \frac{\Delta x}{2} \frac{\partial \rho q_x}{\partial x} \right) \Delta y \Delta z + \left(\rho q_y - \frac{\Delta y}{2} \frac{\partial \rho q_y}{\partial y} \right) \Delta x \Delta z + \left(\rho q_z - \frac{\Delta z}{2} \frac{\partial \rho q_z}{\partial z} \right) \Delta x \Delta y$$
$$- \left(\rho q_x + \frac{\Delta x}{2} \frac{\partial \rho q_x}{\partial x} \right) \Delta y \Delta z - \left(\rho q_y + \frac{\Delta y}{2} \frac{\partial \rho q_y}{\partial y} \right) \Delta x \Delta z - \left(\rho q_z + \frac{\Delta z}{2} \frac{\partial \rho q_z}{\partial z} \right) \Delta x \Delta y \tag{7}$$

Working out term-by-term and combining, yields the following expression

$$- \left(\frac{\partial \rho q_x}{\partial x} + \frac{\partial \rho q_y}{\partial y} + \frac{\partial \rho q_z}{\partial z} \right) \Delta x \Delta y \Delta z \tag{8}$$

Using the *del operator* $\nabla = (\partial/\partial x, \partial/\partial y, \partial/\partial z)$, this can be written as

$$- \nabla \cdot \left(\rho \boldsymbol{q} \right) \Delta x \Delta y \Delta z \tag{9}$$

where the dot represents the scalar vector product operation.

The principle of mass balance states that inflow minus outflow is equal to the change in storage; using Equation (2) the change in storage is given by

$$\frac{\partial M}{\partial t} = \frac{\partial}{\partial t}\left(\rho\theta\,\Delta x\,\Delta y\,\Delta z\right) \tag{10}$$

In this expression, the variables that can really change with time are the water content, θ, because pores can be emptied or filled with water, the density of the water, ρ, because water is compressible, and the size of the control volume, $\Delta x\Delta y\Delta z$, because the porous medium can be compressible. However, for the latter it is assumed that under natural conditions only vertical deformation needs to be considered, such that only Δz depends upon time, while $\Delta x\Delta y$ remains constant. Hence, the storage term can be worked out by using the rules of differentiation as

$$\frac{\partial M}{\partial t} = \frac{\partial\rho}{\partial t}\theta\,\Delta x\,\Delta y\,\Delta z + \rho\frac{\partial\theta}{\partial t}\Delta x\,\Delta y\,\Delta z + \rho\theta\,\Delta x\,\Delta y\frac{\partial\Delta z}{\partial t} = \rho\left(\frac{\theta}{\rho}\frac{\partial\rho}{\partial t} + \frac{\partial\theta}{\partial t} + \frac{\theta}{\Delta z}\frac{\partial\Delta z}{\partial t}\right)\Delta x\,\Delta y\,\Delta z \tag{11}$$

The compression of a porous formation can be expressed in function of the water pressure (Bear, 1972)

$$\frac{1}{\Delta z}\frac{\partial\Delta z}{\partial t} = \alpha\frac{\partial p}{\partial t} \tag{12}$$

where α is the elastic compressibility coefficient of the porous formation, with dimensions $[L^2/F]$, and p represents the groundwater pressure, $[F/L^2]$. The compressibility of the water can be expressed by a similar law

$$\frac{1}{\rho}\frac{\partial\rho}{\partial t} = \beta\frac{\partial p}{\partial t} \tag{13}$$

with β the compressibility coefficient of water, $[L^2/F]$. It is assumed that other effects on the density, as solutes or temperature, are of minor importance and can be ignored. Substitution of these relations in the storage term gives

$$\frac{\partial M}{\partial t} = \rho\left[\theta(\alpha+\beta)\frac{\partial p}{\partial t} + \frac{\partial\theta}{\partial t}\right]\Delta x\,\Delta y\,\Delta z \tag{14}$$

The continuity equation is obtained by putting the change in storage, Equation (14), equal to the net inflow given by Equation (9), and dividing by $\Delta x\Delta y\Delta z$ to express the mass balance per unit volume of porous medium. This gives the following result

$$\rho\left[\theta(\alpha+\beta)\frac{\partial p}{\partial t} + \frac{\partial\theta}{\partial t}\right] = -\nabla\cdot\left(\rho\boldsymbol{q}\right) \tag{15}$$

This relationship states the principle of mass conservation of groundwater in its most general form. However, often density changes of water are of limited importance, and a simplified continuity equation can be obtained by dividing Equation (15) by ρ, and neglecting spatial density differences

$$\left[\theta(\alpha+\beta)\frac{\partial p}{\partial t} + \frac{\partial\theta}{\partial t}\right] = -\frac{1}{\rho}\nabla\cdot\left(\rho\boldsymbol{q}\right) \approx -\nabla\cdot\boldsymbol{q} \tag{16}$$

In this simplified form, the continuity equation expresses the groundwater balance on a volumetric basis. The left-hand side of the equation gives the change in volume of groundwater present in the porous

medium; this change in storage can be due to compression of the medium and the water, or due to changes in water content. The right-hand side of the equation gives the convergence or divergence of the volumetric flow rate of groundwater. When the flow rate is converging, the storage increases and, vice versa, when it diverges the storage decreases.

3.1.3   Macroscopic Approach

The obtained continuity equation is elementary yet fundamental, but not without ambiguity. The problem arises in whether Δx, Δy and Δz can be made small enough to justify the truncation of the Taylor series to zero and first-order terms in the derivation of the inflows and outflows, as in Equation (6). Indeed, when the limit is taken of Δx, Δy and Δz going to zero, the elementary control volume reduces to a point, such that the concept of a porous medium becomes illusive, because a point is either situated in the solid phase, the air phase, or the water phase, and variables such as water content and groundwater flux lose their meaning.

Hence, it seems that groundwater flow problems cannot be formulated and solved correctly at the microscopic level. Of course, such a microscopic approach is not really of any interest in practice. Nevertheless, this ambiguity needs to be cleared up. This is achieved by abandoning the microscopic pore-scale level and by moving to a coarser macroscopic level. At this level, microscopic features such as solid grains and in between pore spaces are ignored, and the medium is conceived as a continuous space with average properties, such that porosity exists in any point of the medium, regardless of whether this point is situated in the solid phase or in the pore space. Also the variables describing the state and movement of water through the medium become macroscopic, such that the average behavior is described, and not the fate of individual fluid particles moving in the pores of the porous medium.

The question remains how macroscopic variables need to be defined and interpreted. Different approaches are in use. The first and still most popular technique is the method of the representative elementary volume (REV), developed by Bear (1972), where macroscopic variables are defined as mean values over a REV, and attributed to the center of the REV. The exact size of the REV is not determined, but it is assumed that the size is much larger than the pore scale and much smaller than the scale of the porous medium. The results obtained with this approach yield macroscopic values for the groundwater flow, representing the average behavior of the fluid over the REV. The main advantage of the method is that macroscopic quantities have a clear physical meaning as they are measurable in the field by taking and analyzing soil samples of an adequate size. The main disadvantages are the loss of detailed information on the microscopic pore-scale level, and the uncertainty due to the assumption that the values of the macroscopic variables are independent of the size of the REV. A detailed discussion can be found in Pinder (1983).

A second, more recent technique is a statistical approach taking into account the uncertainty of the spatial distribution on a microscopic scale (Dagan, 1989). The microscopic arrangement of the porous medium is considered to be random, such that a set of media can be imagined with similar characteristics. The macroscopic variables are considered to be the averages of the variables of the media in the set. The main conceptual difficulty stems from the fact that statistical averaging should be carried out over an ensemble of realizations, whereas in practice usually only one particular porous formation is available from which the statistical information needs to be determined. This is feasible under the assumption of ergodicity, i.e., the characteristics of one sample are assumed to be representative of the whole set. Hence, with the theory of random functions, macroscopic laws can be derived. As long as only relationships between averaged quantities are derived, with no special concern about their fluctuations, the results obtained are essentially the same as with the REV approach. However, the statistical technique is more powerful when dealing with fluctuations and deviations, as in the case of particle or pollutant transport in groundwater.

3.1.4 Motion Equation

The second fundamental law is the momentum equation, based on Newton's second law of motion, i.e., forces induce motion or a change in motion. Consider the elementary control volume as used previously when deriving the continuity equation. An inventory of all forces acting on the water present in the control volume can be made in order to obtain a momentum equation (Figure 3.2). Under natural conditions, the forces to consider are pressure forces, gravity forces, and reaction forces of the solid matrix exerted on the fluid. Because forces are vectors, it is necessary to consider different components along the different directions. For instance, along the left side of the control volume, pressure is acting on the water phase, which yields the following contribution to the force balance in the x-direction

$$\theta p \left(x - \Delta x/2, y, z, t \right) \Delta y \, \Delta z \tag{17}$$

where θ appears in the expression because the water occupies only a θ-fraction of the boundary. A similar force is acting on the right side, but in the opposite direction

$$-\theta p \left(x + \Delta x/2, y, z, t \right) \Delta y \, \Delta z \tag{18}$$

Using truncated Taylor series expansions to relate the θp-terms to the center of the control volume, the resulting pressure force component in the x-direction becomes

$$\left[\left(\theta p - \frac{\Delta x}{2} \frac{\partial \theta p}{\partial x} \right) - \left(\theta p + \frac{\Delta x}{2} \frac{\partial \theta p}{\partial x} \right) \right] \Delta y \, \Delta z = -\frac{\partial \theta p}{\partial x} \Delta x \, \Delta y \, \Delta z \tag{19}$$

Similar expressions can be obtained for the pressure force components in the y- and z-directions, acting on the other sides of the elementary control volume.
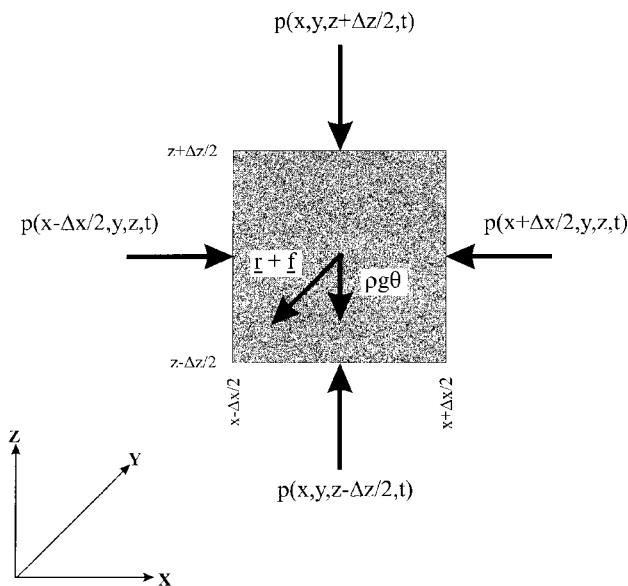

FIGURE 3.2 Forces on a reference elementary volume; the actual volume is three dimensional, but for clarity it is depicted only in two dimensions, x and z.

The gravity force only acts in the z-direction downward and is equal to the total weight of the water in the control volume

$$-\rho g \theta \, \Delta x \, \Delta y \, \Delta z \tag{20}$$

where g is the gravity constant, with dimension $[L/T^2]$.

The evaluation of the reaction force of the solid material on the water is more complicated. It consists of forces acting against the water pressure and friction forces due to the groundwater movement. These forces are extremely difficult to evaluate on a pore-scale level because the shape of the contact surface between the solid phase and the water phase is very complex from a geometrical point of view. Because the exact contact surface is generally not known and would be very difficult to express in mathematical terms anyway, it is impossible to describe these forces on a microscopic scale, and one is forced to adapt a less precise macroscopic approach. Therefore, the reaction forces are defined as average body forces per water volume; the reaction force against the water pressure is denoted as $\boldsymbol{r} = (r_x, r_y, r_z)$, $[F/L^3]$, and the friction force against water movement, as $\boldsymbol{f} = (f_x, f_y, f_z)$, $[F/L^3]$. Hence, the effect of these forces in the x-direction can be written as

$$\left(r_x + f_x\right)\theta \, \Delta x \, \Delta y \, \Delta z \tag{21}$$

with similar expressions in the y- and z-directions.

Now, the force balance in the three directions can be calculated as, respectively

$$\left[-\frac{\partial \theta p}{\partial x} + \left(r_x + f_x\right)\theta\right]\Delta x \, \Delta y \, \Delta z \tag{22}$$

$$\left[-\frac{\partial \theta p}{\partial y} + \left(r_y + f_y\right)\theta\right]\Delta x \, \Delta y \, \Delta z \tag{23}$$

$$\left[-\frac{\partial \theta p}{\partial z} - \rho g \theta + \left(r_z + f_z\right)\theta\right]\Delta x \, \Delta y \, \Delta z \tag{24}$$

Using the *del operator*, ∇, these can be combined in one vector equation

$$\left[-\nabla\left(\theta p\right) - \rho g \theta \nabla z + \left(\boldsymbol{r} + \boldsymbol{f}\right)\theta\right]\Delta x \, \Delta y \, \Delta z \tag{25}$$

This equation can be worked out further and simplified by dividing by the total volume of water present in the control volume, $\theta \Delta x \Delta y \Delta z$, yielding

$$-\nabla p - \frac{p}{\theta}\nabla\theta - \rho g \nabla z + \boldsymbol{r} + \boldsymbol{f} \tag{26}$$

Next, note that when the fluid is at rest, the sum of all forces should be zero. Also, there would be no friction, $\boldsymbol{f} = 0$, and the pressure should be hydrostatic, $\nabla p = -\rho g \nabla z$, from which the overall reaction force, $\boldsymbol{r}$, of the grains opposing the water pressure can be evaluated (Dagan, 1989, p. 73)

$$\boldsymbol{r} = \frac{p}{\theta}\nabla\theta \tag{27}$$

Substituting this result in the force balance, simplifies Equation (26) to

$$-\nabla p - \rho g \nabla z + f \tag{28}$$

In case of motion, the sum of forces is not zero, but equals the change in momentum of the fluid. Furthermore, the friction along the solid–water interface is non-zero and should be specified in function of the motion and friction properties. Several additional assumptions and considerations are necessary to arrive at a useful result.

From field observations it is known that groundwater flow under natural conditions is generally very slow, which leads to a series of important simplifications. First, changes in momentum are also very small and can be neglected compared to other forces acting on the fluid. Hence, although the fluid is in motion, the forces acting on the fluid are approximately in equilibrium

$$-\nabla p - \rho g \nabla z + f \approx 0 \tag{29}$$

This type of flow is known in fluid dynamics as creeping motion. The active forces yielding motion, such as pressure and gravity, are immediately balanced by resisting friction forces of equal strength. Of course, in porous media this is due to the large contact area between fluid and solid material which causes extensive friction, such that significant movement of the fluid is prevented.

Second, because water is a viscous fluid, the friction force results from viscous momentum transfer between the stagnant pore walls and the moving fluid. In principle, the *Navier-Stokes equation* should enable us to calculate the overall resistance force, but due to the complex geometry of the pore walls and pore sizes, this is not feasible in practice. However, in the case of creeping viscous flow, it is known that for an isotropic medium the overall resistance force is opposite in direction to the fluid flow, proportional to the viscosity of the fluid and the magnitude of the flow, and dependent on the size of the obstacles in the flow field. Hence, it is justified to express the friction force as

$$f = -\frac{\mu}{k} q \tag{30}$$

where μ is the dynamic viscosity of the fluid, [FT/L^2], q is the groundwater flux as defined in previous paragraph, here representing the amount and direction of the water movement, and k is a proportionality factor representing the geometry of the pore space; this coefficient has dimensions [L^2] and is denoted as intrinsic permeability or sometimes permeability in short, for reasons that will become evident hereafter.

Substitution of Equation (30) into the force balance Equation (29) and rearranging, results in the following motion equation

$$q = -\frac{k}{\mu}\left(\nabla p + \rho g \nabla z\right) \tag{31}$$

This expression clearly resembles Darcy's law as introduced in the previous chapter. In fact, the obtained expression is a generalization of Darcy's law, describing the flow of a fluid in a porous medium, in case the fluid has a variable density (and viscosity). Such conditions can be present in, for instance, coastal aquifers where salt and fresh water intermix, or in geothermal reservoirs where the density (and viscosity) of the fluid changes with temperature.

However, in case density gradients are not significant, or when the density only changes due to compressibility of the fluid, the motion equation can be simplified as

$$q = -\frac{k\rho g}{\mu}\nabla\!\left(\phi + z\right) = -K\,\nabla h \tag{32}$$

where $K = k\rho g/\mu$ is the hydraulic conductivity $[L/T]$, $\phi = \int dp/\rho g$ is the pressure potential $[L]$, and $h = \phi + z$ the groundwater potential $[L]$; all of these variables have been defined in previous chapters. Hence, Darcy's law, which originally was based on experimental evidence, is nothing else but Newton's second law of motion reduced to a form suitable for describing flow of fluids in porous media. The derivation presented above sheds more light on the underlying principles and assumptions that result in Darcy's law, and enables us to appreciate its applicability in field conditions.

The basic assumption leading to Darcy's law is that movement of a fluid through a porous medium is very restricted, due to large friction forces that balance the driving forces for motion. Hence, Darcy's law is applicable in cases such as water flow in soils or other types of granular porous media, or flow in fractured rocks, but not in cases involving caves or other large openings, like cracks, fissures, etc. These latter problems should be analyzed with viscous flow theories. Second, all ambiguity as discussed in the case of the mass balance equation is also present here. Darcy's law is a macroscopic approach; on a microscopic level there are no such things as hydraulic conductivity or permeability. Hence, one might wonder about the errors involved in using a macroscopic approach such as Darcy's law. However, this is not really an important issue, because the uncertainty in obtaining representative and accurate values for the hydraulic conductivity or permeability has a much larger effect on the overall accuracy. Indeed, field investigations show that natural porous media such as ground layers, exhibit a large variability and heterogeneity in conductive properties, which are difficult to quantify accurately by experimental or deterministic means.

3.1.5   Extensions of Darcy's Law

Darcy's law, like the momentum equation, is a vector relationship. When the flow is three-dimensional, a Darcy law can be written for each of the directions. For instance, in the case of three-dimensional Cartesian coordinates, there are three Darcy equations

$$q_x = -K\frac{\partial h}{\partial x} \tag{33}$$

$$q_y = -K\frac{\partial h}{\partial y} \tag{34}$$

$$q_z = -K\frac{\partial h}{\partial z} \tag{35}$$

In the case of other coordinate systems, the rules of nabla-calculus should be applied in order to find the different expressions along the coordinate axes. For instance, in a cylindrical coordinate system (r,φ,z), this becomes

$$q_r = -K\frac{\partial h}{\partial r} \tag{36}$$

$$q_\varphi = -\frac{K}{r}\frac{\partial h}{\partial \varphi} \tag{37}$$

$$q_z = -K \frac{\partial h}{\partial z} \tag{38}$$

In addition to heterogeneity, i.e., porosity and conductivity variations from point to point, a dependence on direction is also possible. This is the case for so-called anisotropic porous media, where, due to some direction-related properties such as preferential lining of fractures, stratifications or layering, the conductivity changes depending upon direction. Such situations can be described by an extension of Darcy's law, where the conductivity becomes a second order symmetrical tensor, K, with following components

$$K = \begin{bmatrix} K_{xx} K_{xy} K_{xz} \\ K_{xy} K_{yy} K_{yz} \\ K_{xz} K_{yz} K_{zz} \end{bmatrix} \tag{39}$$

In Cartesian coordinates, Darcy's law becomes

$$q_x = -K_{xx} \frac{\partial h}{\partial x} - K_{xy} \frac{\partial h}{\partial y} - K_{xz} \frac{\partial h}{\partial z} \tag{40}$$

$$q_y = -K_{xy} \frac{\partial h}{\partial x} - K_{yy} \frac{\partial h}{\partial y} - K_{yz} \frac{\partial h}{\partial z} \tag{41}$$

$$q_z = -K_{xz} \frac{\partial h}{\partial x} - K_{yz} \frac{\partial h}{\partial y} - K_{zz} \frac{\partial h}{\partial z} \tag{42}$$

Notice, that potential gradients in one direction can yield flows in other directions. However, such situations are rather of academical interest, while in practice these equations are not much used, because it is not feasible to assess all conductivity components accurately, even regardless of heterogeneity. One notable exception is the case of layered formations, usually of sedimentary origin. For horizontal layering only two conductivity components exist: a horizontal conductivity, K_h, and a vertical conductivity, K_v. In such a case, Darcy's law becomes

$$q_x = -K_h \frac{\partial h}{\partial x} \tag{43}$$

$$q_y = -K_h \frac{\partial h}{\partial y} \tag{44}$$

$$q_z = -K_v \frac{\partial h}{\partial z} \tag{45}$$

These equations are also useful in practice, in view of the fact that effects of horizontal layering on the conductivity are the rule rather than the exception in ground layers. In the case of folded or dipping layered formations also off-diagonal terms are important; all components of the conductivity tensor can be calculated from the dip angle and the conductivity values normal and parallel to the layering, following the rules of tensor calculus (Bear, 1972).

3.2 Groundwater Flow Equations

3.2.1 General

The groundwater flow equation is obtained by combining the continuity equation with Darcy's law. The most general form is obtained when the groundwater balance, Equation (15), is combined with the general Darcy Equation (31), yielding

$$\theta(\alpha+\beta)\frac{\partial p}{\partial t}+\frac{\partial \theta}{\partial t}=\nabla\cdot\left[\frac{k}{\mu}\left(\nabla p+\rho g\nabla z\right)\right] \tag{46}$$

Next, note that the water content, θ, is related to the water pressure, either in the case of saturated conditions because the water content equals the porosity, which depends upon the water pressure due to elastic deformation of the porous medium, or in the case of unsaturated conditions, because the water content and the water pressure (or suction) are related through the so-called water retention relationship. (We will not go into details of unsaturated flow; more information can be found in the pertinent chapters.) It is sufficient to see that the groundwater flow equation can be expressed as a function of the water pressure only

$$S_p\frac{\partial p}{\partial t}=\nabla\cdot\left[\frac{k}{\mu}\left(\nabla p+\rho g\nabla z\right)\right] \tag{47}$$

where S_p is the storage coefficient of the porous medium related to water pressure changes, with dimensions $[L^2/F]$, and given by

$$S_p=\theta(\alpha+\beta)+\frac{d\theta}{dp} \tag{48}$$

This storage coefficient depends upon different parameters, related to soil and water properties, and saturated or unsaturated conditions. Hence, it seems rather complicated, but fortunately it has a simple physical interpretation, i.e., the storage coefficient gives the volume of water released per unit volume of porous medium and per unit decline of the water pressure.

It is instructive to write the obtained groundwater flow equation in Cartesian coordinates, taking into consideration different horizontal and vertical permeabilities

$$S_p\frac{\partial p}{\partial t}=\frac{\partial}{\partial x}\left(\frac{k_h}{\mu}\frac{\partial p}{\partial x}\right)+\frac{\partial}{\partial y}\left(\frac{k_h}{\mu}\frac{\partial p}{\partial y}\right)+\frac{\partial}{\partial z}\left[\frac{k_v}{\mu}\left(\frac{\partial p}{\partial z}+\rho g\right)\right] \tag{49}$$

This equation describes three-dimensional movement of groundwater in its most general form. Notice that as a consequence, the flow equation is expressed as a function of the water pressure, and not as a function of the groundwater potential.

3.2.2 Saturated Groundwater Flow Equation

In practice, the general groundwater flow equation is not much used because simplifications are usually introduced. When density effects can be ignored, a groundwater flow equation can be obtained by combining the continuity equation with the simple form of Darcy's law, given by Equation (32), which yields

$$\theta(\alpha + \beta)\frac{\partial p}{\partial t} + \frac{\partial \theta}{\partial t} = \nabla \cdot (K \nabla h) \tag{50}$$

Making use of the fact that when the fluid density is constant the water pressure differences in time can be related to the temporal variation of the groundwater potential

$$\frac{\partial p}{\partial t} = \rho g \frac{\partial h}{\partial t} \tag{51}$$

the resulting flow equation can be written as

$$\rho g \theta(\alpha + \beta)\frac{\partial h}{\partial t} + \frac{\partial \theta}{\partial t} = \nabla \cdot (K \nabla h) \tag{52}$$

This expression can be considered as a basic groundwater flow equation, because its relates the amount of groundwater present, θ, and the groundwater potential, h, to the characteristics of the porous medium and the fluid in the space and time continuum. However, generally both the water content, θ, and the potential, h, are unknown, and it becomes necessary to have more information on the type of flow before the equation can be solved.

For instance, groundwater flow can be considered to occur under saturated conditions. In such a case, the water content equals the porosity, and under the assumption of incompressible solid grains the change in porosity can be related to the compression of the porous medium, which depends upon the water pressure or groundwater potential, as given by Equation (12), i.e.,

$$\frac{\partial \theta}{\partial t} = \frac{\partial n}{\partial t} = \frac{(1-n)}{\Delta z}\frac{\partial \Delta z}{\partial t} = (1-n)\alpha\frac{\partial p}{\partial t} = (1-n)\alpha\rho g\frac{\partial h}{\partial t} \tag{53}$$

Substituting this expression in Equation (52) gives

$$S_0 \frac{\partial h}{\partial t} = \nabla \cdot (K \nabla h) \tag{54}$$

where S_0 is the specific storage coefficient, depending only upon compressibility of the porous medium and the fluid

$$S_0 = \rho g(\alpha + n\beta) \tag{55}$$

The specific storage can be interpreted physically as the volume of water released per unit saturated porous medium and per unit decline of the groundwater potential. The units are $[L^{-1}]$. Common values are usually very small because elastic deformation of ground layers or water is limited and certainly not noticeable in the field. Of course, plastic deformation of ground layers, consisting of materials like clay or peat, is excluded. Such deformations mainly occur under the influence of external loads or extensive groundwater pumping, and are noticeable in the field as *land subsidence*. Elastic deformations and storage in aquifers are described by Verruijt (1969). Theories on porous medium deformability under various soil conditions, with special attention to the subject of land subsidence, are discussed in part 2 of Bear and Corapcioglu (1984).

Written in Cartesian coordinates and using vertical and horizontal conductivities, the saturated groundwater flow equation becomes

$$S_0 \frac{\partial h}{\partial t} = \frac{\partial}{\partial x}\left(K_h \frac{\partial h}{\partial x}\right) + \frac{\partial}{\partial y}\left(K_h \frac{\partial h}{\partial y}\right) + \frac{\partial}{\partial z}\left(K_v \frac{\partial h}{\partial z}\right) \tag{56}$$

This is the most-used groundwater flow equation in practice. The equation is of the diffusive type, indicating that potential differences will be dissipated through the medium with a speed depending upon the value of K/S_0. Hence, because S_0 is usually small, this process can be rather fast, unless K is also small, as for instance in the case of clay layers.

It is also interesting to write the flow equation in cylindrical coordinates (r, φ, z)

$$S_0 \frac{\partial h}{\partial t} = \frac{1}{r}\frac{\partial}{\partial r}\left(rK_h \frac{\partial h}{\partial r}\right) + \frac{1}{r^2}\frac{\partial}{\partial \varphi}\left(K_h \frac{\partial h}{\partial \varphi}\right) + \frac{\partial}{\partial z}\left(K_v \frac{\partial h}{\partial z}\right) \tag{57}$$

This equation is used for the prediction of groundwater flow toward wells, as described by Hantush (1964), Walton (1970), Huisman (1972), and Kruseman and de Ridder (1991).

When groundwater flow is stationary, the variables become independent of time, and the flow equation is reduced to

$$\nabla \cdot \left(K \nabla h\right) = 0 \tag{58}$$

Written in Cartesian coordinates, and generalizing by using vertical and horizontal conductivities, we obtain

$$\frac{\partial}{\partial x}\left(K_h \frac{\partial h}{\partial x}\right) + \frac{\partial}{\partial y}\left(K_h \frac{\partial h}{\partial y}\right) + \frac{\partial}{\partial z}\left(K_v \frac{\partial h}{\partial z}\right) = 0 \tag{59}$$

This equation clearly shows that movement of groundwater is a *potential flow problem.* The driving forces for the movement are differences in groundwater potential, and the resulting fluxes depend upon the conductive properties of the medium. In this way groundwater flow resembles other types of potential flow, like heat flow in heat conducting media under the influence of temperature gradients, or electric flow in electric conductive materials under the influence of electric potential differences. Many solutions have been obtained for groundwater flow problems based on the theory of potential functions, especially in two dimensions (Polubarinova-Kochina, 1962; Harr, 1962; Verruijt, 1970; Halek and Svec, 1979; Strack, 1989).

Another possibility is to assume that the groundwater flow occurs under unsaturated conditions, but we will not go into detail here because these derivations can be found in the pertinent chapters.

3.2.3 Boundary Conditions

Groundwater flow equations are partial differential equations that relate the dependent variable, as the groundwater potential, to the independent variables, the coordinates of the flow domain, and the time. However, flow equations are only valid inside the flow domain, and not on the boundaries, because there the porous medium ends and other phenomena occur which can influence the groundwater flow inside the domain. These interactions at the boundaries are governed by other physical laws and have to be described separately by mathematical expressions, the so-called *boundary conditions.*

First, consider the time dimension. Because transient groundwater flow equations contain first-order partial derivatives versus time, it can be shown mathematically that only conditions at the start, i.e., at time zero, are needed. Hence, knowing the distribution of the dependent variable in the flow domain at the start is sufficient to describe the further evolution in time. For instance, when the groundwater potential is the dependent variable, the so-called initial condition is

$$h(x, y, z, 0) = h_0(x, y, z) \tag{60}$$

where h_0 represents a known function of x, y, and z.

Boundary conditions for the physical domain are somewhat more complicated. Mathematically, it can be shown that because groundwater flow equations contain second-order partial derivatives versus the space coordinates, conditions are required at every point of the boundary of the physical flow domain, even when the boundaries tend to infinity. Because the interaction between the inside groundwater flow and the outside world can be complex, and our ability to conceive such phenomena accurately is rather limited, boundary conditions are usually described in a simplified way.

Generally, three types of boundary conditions are considered. First-type boundary conditions apply when the value of the dependent variable at the boundary is known. For instance, in the case of groundwater potentials, such a boundary condition would be

$$h(x_b, y_b, z_b, t) = h_b(t) \tag{61}$$

where (x_b, y_b, z_b) represents a point on the boundary, and h_b is a known function of time. Such an expression is called a potential boundary condition and is used when the groundwater is in contact with a water body having a known potential h_b, as for instance a river, lake, or reservoir, etc. The key assumption is that whatever groundwater flow occurs inside the flow domain and at the boundary, it will have no influence on the potential of the outside water body, such that this potential remains fixed as stated by the boundary condition.

Second-type boundary conditions are so-called flux boundary conditions, where it is assumed that the amount of groundwater exchange through the boundary is known. The amount of water exchange is given by the groundwater flux component perpendicular to the boundary, such that the boundary condition can be expressed as

$$q_n(x_b, y_b, z_b, t) = q_b(t) \tag{62}$$

where q_n represents the flux component normal to the boundary, and $q_b(t)$ is a known function of time. A flux-type condition can be expressed in function of the dependent variable, the groundwater potential, by using Darcy's law

$$q_n = -K \frac{\partial h}{\partial n} = q_b(t) \tag{63}$$

where $\partial h / \partial n$ represents the derivative of the potential perpendicular to the boundary. From a mathematical point of view, the choice of the positive direction of the normal to the boundary determines whether q_n values are positive or negative. However, in practice, often a simpler convention is used by considering fluxes entering the flow domain as positive and outgoing fluxes as negative, regardless of the sense of the normal on the boundary.

A flux boundary condition implies that whatever the state and flow of the groundwater inside the flow domain and at the boundary, the normal flux is fixed by external conditions and remains as stated by the boundary condition. Typical examples of such conditions are pumping wells, groundwater recharge, infiltration, etc., in general any situation where the flux is supposedly fixed and known. An obvious example is an impervious boundary because the flux component perpendicular to the boundary is strictly zero.

A third-type boundary condition is a mixture of the two previous types. It applies when potential and normal flux component at the boundary are related to each other. Such a condition applies in the case of a flow domain in contact with a water body with a fixed potential h_b, but where the exchange between

the groundwater reservoir and the water body is restricted, due to the presence of some resistance. Mathematically, this condition is expressed as

$$q_n(x_b, y_b, z_b, t) = C_b \left[h_b(t) - h(x_b, y_b, z_b, t) \right] \tag{64}$$

where we have used the above-mentioned sign convention for the normal boundary flux, and C_b represents the boundary conductance, with dimensions $[T^{-1}]$, between the groundwater reservoir and the outside water source or sink. This parameter can be explained physically as the conductive capacity of a permeable boundary layer present between the groundwater flow domain and the water body, such that

$$C_b = K_b / D_b \tag{65}$$

where K_b is the hydraulic conductivity of the boundary layer, and D_b its thickness. For instance, this situation would apply when a river is in contact with the groundwater, but the interaction is restricted due to the presence of a mud layer in the river bed. Another example is a well with a clogged filter.

One of the most complicated boundary conditions is the water table. Actually, the water table is not a true boundary, because groundwater flow can cross the water table when transfer of water occurs between the unsaturated and saturated zones. However, because unsaturated flow is usually difficult to solve, it is often neglected and the water table is considered as an upper boundary of the groundwater flow domain. If such a simplification is used, it becomes necessary to specify the water table as a boundary condition. When the position of the water table is known, this is not so complicated, because the pressure at the water table is atmospheric such that the relative pressure is zero, and the following potential boundary condition applies

$$h(x_w, y_w, z_w, t) = z_w \tag{66}$$

where (x_w, y_w, z_w) represents a point on the water table. However, in most cases the position of the water table is not known exactly, and in addition to Equation (66) another boundary condition is needed, in order to solve the groundwater flow problem, which now also includes the determination of the shape and position of the water table. Obviously, this additional condition should state something about the amount of flow crossing the water table. Several possibilities exist, the most simple being the case where a steady water table acts as an impervious upper boundary, i.e., no flow crosses this surface. This can mathematically be expressed as follows. An implicit equation describing the position of the water table is $h - z = 0$. The gradient of this expression is a vector perpendicular to the water table, and hence, the scalar product with the groundwater flux is zero if no water is crossing the water table, hence

$$\mathbf{q} \cdot \nabla(h - z) = 0 \tag{67}$$

When this is worked out further and Darcy's law is used to express the flux, the following boundary condition is obtained

$$K_h \left(\frac{\partial h}{\partial x} \right)^2 + K_h \left(\frac{\partial h}{\partial y} \right)^2 + K_v \left(\frac{\partial h}{\partial z} \right)^2 - K_v \frac{\partial h}{\partial z} = 0 \tag{68}$$

This equation applies to every point (x_w, y_w, z_w) of the water table. Notice that it is nonlinear, which is the price to pay for neglecting the unsaturated flow. (There is no such thing as a free lunch.) More complicated equations are needed when the water table, is moving in time, or when certain flows are crossing the water table, as for instance a groundwater recharge flux. It is not possible to go into more

detail here. For the interested reader, these cases are discussed by Bear (1972). However, in practice, these types of boundary conditions are not used very much, because they are too complicated, so that it becomes impossible to solve the groundwater flow problem exactly. Therefore, approximate techniques are used. For instance in the case of numerical simulation models, the position of the water table is determined in an iterative way. For steady-state problems, first a water table position is chosen above the expected position, then the groundwater flow equation is solved with a flux-type boundary condition as Equation (62), and afterwards a new water table position is determined by identifying the surface where $h = z$, after which the procedure can be restarted. For transient conditions, the position of the water table is tracked in time by equating the inputs and outputs of groundwater at the water table and calculating the resulting changes in water table position.

3.3 Hydraulic Approach to Groundwater Flow

3.3.1 Concept

In general, groundwater flow is three-dimensional, but due to the geometry of ground layers and differences in hydraulic conductivities, actual groundwater flows tend to be concentrated in certain directions, as illustrated in Figure 3.3. The reason for this is that ground layers usually extend horizontally over large distances, while the vertical dimensions are rather restricted. Also, due to the large variety in conductive properties, ground layers can be grouped into three classes: previous formations or aquifers, semipervious formations or aquitards, and impervious formations. Because water flows through a porous medium along the path of least resistance, groundwater moves predominantly longitudinally in the previous formations, and transversely in the semipervious layers. Hence, due to the horizontal layering, groundwater flow is essential horizontal in aquifers, and vertical in aquitards, as depicted schematically in Figure 3.3B. Therefore, instead of considering the flow as three dimensional, a simplified description of groundwater flow is possible when horizontal flow components are considered to be dominant in aquifers and vertical components in aquitards, as shown in Figure 3.3C. This concept is termed the hydraulic approach to groundwater flow, because the movement is considered to be directed as water flow in pipes or channels (Polubarinova-Kochina, 1962; Bear, 1972; Bear, 1979). In the case of a phreatic aquifer, this is also-called the Dupuit–Forchheimer theory, after the two scientists who pioneered the approach in the second half of the 19th century.

3.3.2 Motion Equations

In order to have a clear understanding of the basic concepts of the hydraulic approach, it is convenient to start with the equation of motion before discussing the continuity equation. Consider a confined aquifer, bound by a bottom surface, $b(x,y)$, and an upper surface, $s(x,y)$, such that its thickness is given by $D = s - b$, as shown in Figure 3.4A. When the groundwater potential is measured by a *piezometer*, the actual depth of the piezometer is unimportant because vertical potential differences are very small. Hence, h is almost independent of z, and the total horizontal volumetric flow rate of the groundwater over the entire thickness of the aquifer, in the x direction, can be calculated as

$$Q_x = \int_b^s q_x \, dz = \int_b^s -K_h \frac{\partial h}{\partial x} dz \approx -\int_b^s K_h \, dz \frac{\partial h}{\partial z} = -T \frac{\partial h}{\partial x} \tag{69}$$

where Q_x is the x-component of the hydraulic groundwater flux, $\boldsymbol{Q} = (Q_x, Q_y)$, with dimensions volume per time per length $[L^3/TL = L^2/T]$, and T is the transmissivity $[L^2/T]$, defined as

$$T = \int_b^s K_h \, dz \tag{70}$$
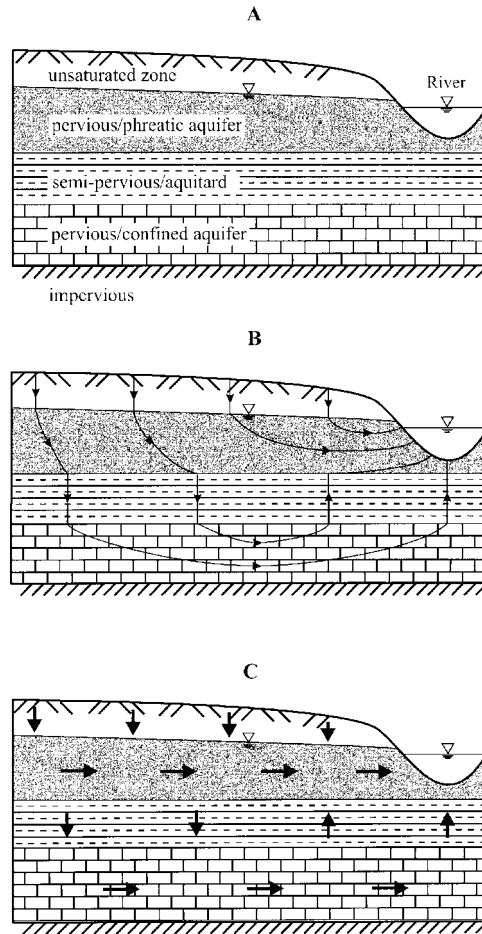
FIGURE 3.3 Occurrence and movement of groundwater: (A) types of groundwater layers, (B) actual groundwater flow paths, and (C) predominant groundwater flow directions.

In a similar way, the y-component of the hydraulic flux is given as

$$Q_y = -T\frac{\partial h}{\partial y} \tag{71}$$

The transmissivity is a parameter expressing the overall horizontal conductance of a confined aquifer. Equations (70) and (71) are transformations of Darcy's law, describing overall horizontal movement of groundwater in a confined aquifer.

In the case of an aquitard bounded by two aquifers, vertical flow can pass through the aquitard from one aquifer to the other, as depicted in Figure 3.4B. This vertical groundwater flux can be calculated by Darcy's law as

$$q_z = -K_v \frac{h_2 - h_1}{D} = C(h_1 - h_2) \tag{72}$$
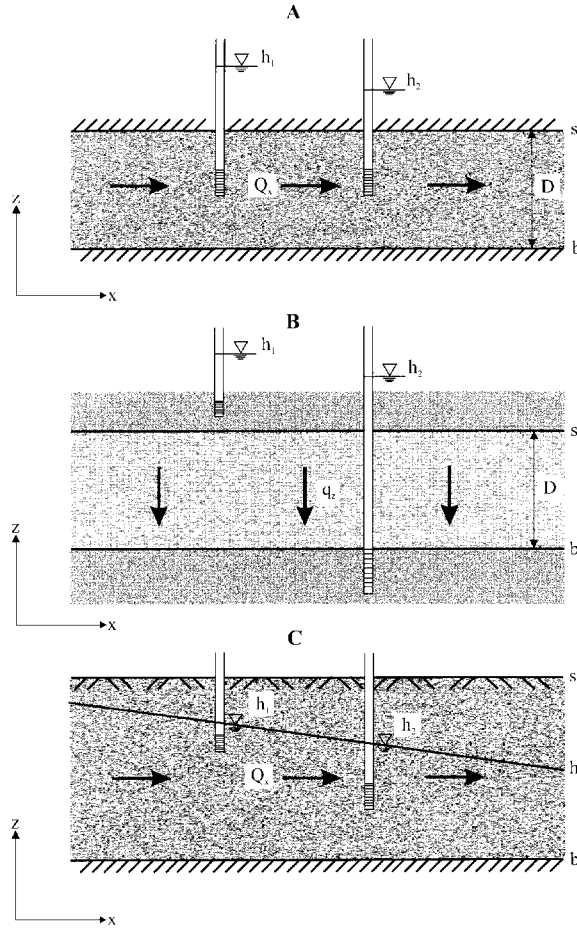
FIGURE 3.4 Hydraulic approach to groundwater flow: (A) horizontal flow in a confined aquifer, (B) vertical flow in an aquitard, and (C) horizontal flow in a phreatic aquifer.

where K_v is the vertical conductivity of the aquitard, D is its thickness, and the potential gradient is obtained from the difference in potential between the two aquifers bounding the aquitard. Parameter C is the conductance, commonly called the leakage coefficient, and is defined as

$$C = K_v / D \tag{73}$$

with dimensions $[T^{-1}]$; this coefficient expresses the leaking capacity of an aquitard, allowing vertical groundwater transfer between two aquifers bounding the aquitard.

The case of a phreatic aquifer is shown schematically in Figure 3.4C. The aquifer is bound from below by its base, $b(x,y)$, and from above by the soil surface, $s(x,y)$. Actually, the true upper boundary is the water table. If a piezometer were to be installed at a certain depth below the water table, the water level would reveal a potential very close to the water table elevation, because vertical potential differences are insignificant as vertical groundwater flow is very small. Hence, h is nearly independent of z and almost coincides with the position of the water table, which is the upper surface of the saturated zone. Consequently, the total flow of groundwater passing horizontally through the aquifer in the x-direction is given by

$$Q_x = \int_b^s q_x\, dx = \int_b^h q_x\, dz = \int_b^h -K_h \frac{\partial h}{\partial x}\, dz \approx -\int_b^h K_h\, dz \frac{\partial h}{\partial z} = -K_e (h-b) \frac{\partial h}{\partial x} \qquad (74)$$

where horizontal flows in the unsaturated zone have been neglected, and K_e is an effective hydraulic conductivity given by

$$K_e = \frac{1}{h-b} \int_b^h K_h\, dz \qquad (75)$$

In case of a homogeneous medium, K_e is equal to K_h. However, ground layers are generally heterogeneous, such that K_e needs to be determined by averaging; this is called upscaling or homogenization and requires careful consideration of the hydrogeological and stratigraphical characteristics of the subsurface and of the averaging procedure. Theoretical and mathematical aspects are discussed by Zijl and Nawalany (1993).

An expression similar to Equation (74) can be obtained for the flow component in the y-direction

$$Q_y = -K_e (h-b) \frac{\partial h}{\partial y} \qquad (76)$$

This is basically the approach of Dupuit, who postulated that groundwater flow in a phreatic aquifer is proportional to the slope of the groundwater table. Also, a transmissivity can be defined similar to that for a confined aquifer, but this makes little sense because the thickness of the phreatic aquifer changes with water table position, such that the transmissivity is not constant.

A simplification is possible when the base of the phreatic aquifer is horizontal. If the water table position is measured from this base, $H = h - b$, the hydraulic flow approximation can be written as

$$Q_x = -K_e H \frac{\partial H}{\partial x} = -K_e \frac{\partial}{\partial x}\left(\frac{H^2}{2} \right) \qquad (77)$$

$$Q_y = -K_e H \frac{\partial H}{\partial y} = -K_e \frac{\partial}{\partial y}\left(\frac{H^2}{2} \right) \qquad (78)$$

Dupuit's equation is mostly known in this form. It replaces Darcy's law for describing overall horizontal groundwater flow in a phreatic aquifer.

3.3.3   Flow Equation for a Confined Aquifer

To obtain a flow equation, the equation of motion has to be combined with the continuity equation. The latter results from the principle of mass conservation applied to a control volume. Because in the hydraulic approach the flow is integrated over the total depth of an aquifer, the control volume should be of a similar concept, i.e., encompass the total vertical extent of the aquifer, as shown in Figure 3.5. Hence, the control box centered around a point (x,y) has sizes Δx and Δy in the horizontal plane and extends from the bottom plane, $b(x,y)$, to the upper surface, $s(x,y)$, along the vertical such that its vertical size is the aquifer thickness, D. The dimensions of the control volume reflect the hydraulic approach concept, i.e., elementary sizes in the horizontal direction are of the same order as the thickness of the aquifer; hence, the aquifer is considered to be essentially a two-dimensional object.

The groundwater mass inflow minus outflow in the control volume can be calculated as

FIGURE 3.5 Mass conservation in a reference elementary volume in a confined aquifer; the actual volume is three dimensional, but for clarity it is depicted only in two dimensions, x and z.

$$
\begin{aligned}
&\left[\rho Q_x\left(x - \Delta x/2, y, t\right) - \rho Q_y\left(x + \Delta x/2, y, t\right)\right]\Delta y \\
&+ \left[\rho Q_y\left(x, y - \Delta y/2, t\right) - \rho Q_y\left(x, y + \Delta y/2, t\right)\right]\Delta x + \rho R \Delta x\, \Delta y
\end{aligned}
\tag{79}
$$

where R is a recharge or net exchange of groundwater passing through the base or through the upper surface of the aquifer. Actually, we consider here a semiconfined aquifer; in the case of a truly confined aquifer, R would be zero. The sign convention for R is that inputs are considered positive, irrespective of the direction along the z-axis.

Using Taylor series expansions, the total net inflow in the control box, becomes

$$
\left(-\frac{\partial \rho Q_x}{\partial x} - \frac{\partial \rho Q_y}{\partial y} + \rho R\right)\Delta x\, \Delta y
\tag{80}
$$

This should be balanced by a change in storage. The amount of water mass, M, present in the control volume is given by

$$
M = \int_b^s \rho n\, dz\, \Delta x\, \Delta y = \rho n D \Delta x\, \Delta y
\tag{81}
$$

where in the last term we have considered average porosity and density values over the thickness of the aquifer. The change in storage can be calculated as

$$
\frac{\partial M}{\partial t} = \left(\frac{\partial \rho}{\partial t} nD + n\frac{\partial n}{\partial t}D + \rho n\frac{\partial D}{\partial t}\right)\Delta x\, \Delta y
\tag{82}
$$

where compression of the water and aquifer is considered, and vertical deformation of the aquifer. Using the relationships for expressing the compressibility of water and porous medium and assuming that the solid grains are incompressible, this can be worked out in a way similar to what was done for the general groundwater flow equation, yielding

$$\frac{\partial M}{\partial t} = \rho D (\alpha + n\beta) \frac{\partial p}{\partial t} \Delta x \Delta y = \rho^2 g D (\alpha + n\beta) \frac{\partial h}{\partial t} \Delta x \Delta y \tag{83}$$

Making use of the definition of the specific storage coefficient, Equation (55), this can also be written as

$$\frac{\partial M}{\partial t} = \rho S_0 D \frac{\partial h}{\partial t} \Delta x \Delta y = \rho S \frac{\partial h}{\partial t} \Delta x \Delta y \tag{84}$$

where S is the aquifer storage coefficient or storativity, defined as

$$S = S_0 D = \rho g (\alpha + n\beta) D \tag{85}$$

and which is a measure of the overall aquifer storage properties. The storativity is dimensionless. It can be defined as the volume of groundwater released by a confined aquifer per unit horizontal surface of the aquifer and per unit decline of groundwater potential, all in accordance with the hydraulic approach, where variables are considered averaged or integrated over the vertical dimensions of the aquifer.

Combining Equations (80) and (84), dividing by $\rho \Delta x \Delta y$, and neglecting density gradients, results in

$$S \frac{\partial h}{\partial t} = -\frac{\partial Q_x}{\partial x} - \frac{\partial Q_y}{\partial y} + R \tag{86}$$

Using the expression for horizontal hydraulic flow given by Equations (69) and (71) results in the following hydraulic groundwater flow equation for a (semi-) confined aquifer

$$S \frac{\partial h}{\partial t} = \frac{\partial}{\partial x} \left(T \frac{\partial h}{\partial x} \right) + \frac{\partial}{\partial y} \left(T \frac{\partial h}{\partial y} \right) + R \tag{87}$$

Compared to saturated three-dimensional groundwater flow, Equation (56), it can be noted that this equation contains no z-dimension, which is the main simplification resulting from the hydraulic approach. Also, as a consequence, the needed porous medium properties are transmissivity and storativity, which are aquifer properties integrated over the vertical dimension. Therefore, Equation (87) is very useful in practice because it is simpler and needs less detailed knowledge about the medium properties. Care should be taken when applying this equation to practical situations, because boundary conditions are required in accordance with the hydraulic approach. This implies that boundary conditions should be independent of the elevation, i.e., they should apply to the total aquifer thickness. For instance, pumping by a fully penetrating well can be considered with the hydraulic approach, but not a partially penetrating well because this induces vertical potential gradients in the aquifer.

3.3.4 Flow Equation for a Phreatic Aquifer

A flow equation for a phreatic aquifer based on the hydraulic approach can be obtained as for a confined aquifer, but some complications occur. The control volume is similar; it is centered around the point (x,y), with sizes Δx and Δy horizontally, and extends from the base to the soil surface, as shown in
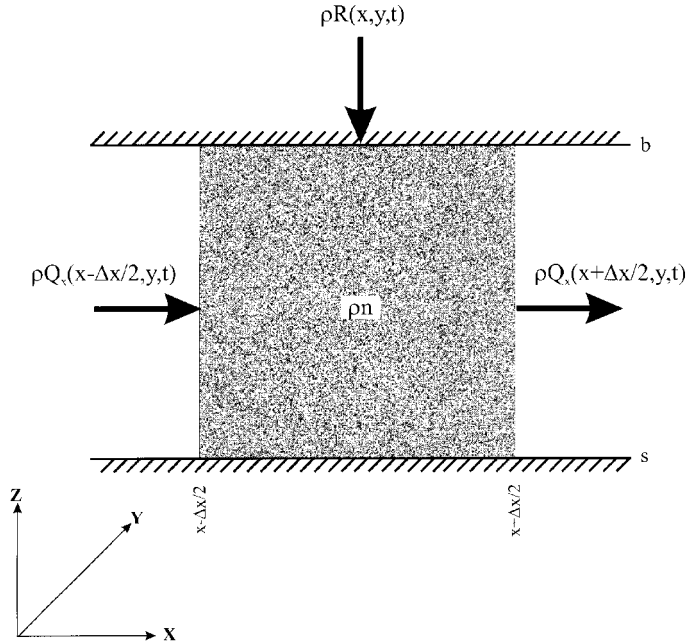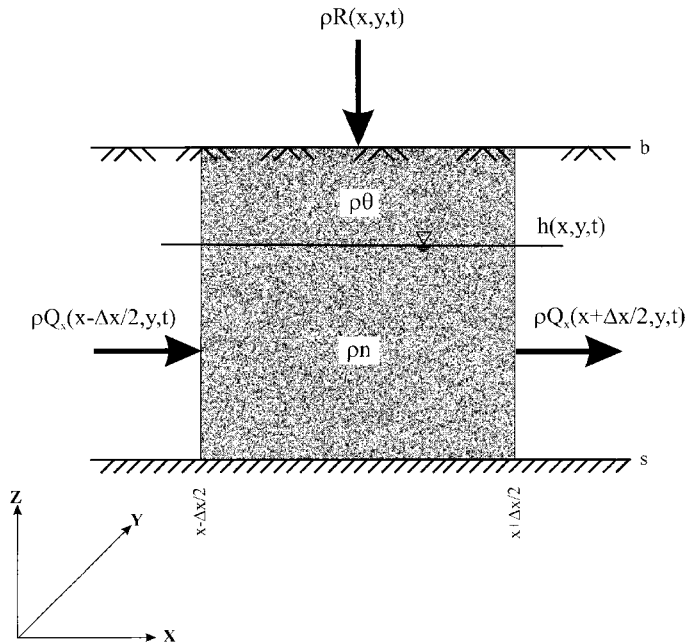
FIGURE 3.6 Mass conservation in a reference elementary volume in a phreatic aquifer; the actual volume is three dimensional, but for clarity it is depicted only in two dimensions, *x* and *z*.

Figure 3.6. Notice, that the control volume includes the unsaturated zone because the groundwater flow and in particular the resulting changes in the water table elevation will also have effects on the water balance in the unsaturated zone, which have to be taken into account.

The inflow minus the outflow depends upon the hydraulic horizontal flow components in the saturated zone and possible vertical inputs or outputs through the upper or lower boundary. For the latter, water exchanges at the soil surface are especially important. These can be inputs as infiltration or outputs as evapotranspiration. Because groundwater flow is a slow process, it is reasonable to consider an average flux at the soil surface, a net precipitation or recharge, which is the result of all hydrological processes occurring at the soil surface. Hence, the input minus output becomes similar to Equation (79), and when developed further by Taylor series expansions the same expression as Equation (80) is obtained. Of course, in addition to recharge, transfer could also take place through the base of the phreatic aquifer if an aquitard is present connected to an underlying aquifer, in which case another R term should be added to the equation.

Somewhat more complicated is the calculation of the change in storage in the control volume because the unsaturated zone has to be included. The total mass of water present can be calculated by integration along the vertical

$$M = \Delta x \, \Delta y \int_{b}^{s} \rho \theta \, dz \tag{88}$$

The change in storage is given by

$$\frac{\partial M}{\partial t} = \Delta x \, \Delta y \frac{\partial}{\partial t} \int_{b}^{s} \rho \theta \, dz \tag{89}$$

Now it becomes clear why the control volume extends to the soil surface, because changes in storage can be significant when differences in water content are possible and, for instance, part of the saturated zone can become unsaturated, or vice versa. In the presence of such changes, storage effects due to compressibility of the water or porous medium become insignificant and can be neglected. It follows that the change in storage can be evaluated as

$$\frac{\partial M}{\partial t} = \rho \, \Delta x \, \Delta y \int_b^s \frac{\partial \theta}{\partial t} \, dz \qquad (90)$$

The exact evaluation of the integral is not simple. The situation is shown schematically in Figure 3.7. If the water table drops by an amount Δh, the change in storage depends upon the moisture distribution above the water table, before and after the drop in water table position. Actually, time will also have an influence because the water in the unsaturated zone needs time to percolate downward to the water table. Hence, the exact change in storage is difficult to calculate unless unsaturated flow is taken into consideration, which is exactly what one would like to avoid by applying the hydraulic approach. Hence, the storage needs to be approximated in a simpler way. This can be achieved as follows. Consider the situation depicted in Figure 3.7A, where the soil surface is high above the water table. Under normal average conditions, the distribution of the water content will vary with elevation above the water table. It will equal the porosity at the water table, and generally will decrease with height above the water table, gradually approaching a residual water content, which is characteristic for the unsaturated zone far above the water table. This residual water content is denoted by θ_r. Now, if the water table is lowered by an amount Δh, given sufficient time, the moisture distribution will eventually be similar to what it was before, only shifted downward over a distance Δh. Hence, the volumetric difference in storage can be calculated as

$$\int \frac{\partial \theta}{\partial t} \, dz = \frac{1}{\Delta t} \left[\int \theta(z, t + \Delta t) - \theta(z, t) \right] dz = \frac{1}{\Delta t} \int \left[\theta(z + \Delta h, t) - \theta(z, t) \right] dz = (n - \theta_r) \frac{\Delta h}{\Delta t} = n_e \frac{\partial h}{\partial t} \qquad (91)$$

where $n_e = n - \theta_r$ is defined as the effective porosity or specific yield, i.e., the amount of water released per volume of porous medium when changing from the saturated state to the unsaturated state high above the water table. In reality, things are not that simple. First, it is possible that the soil surface is not situated far above the water table, as shown in Figure 3.7B, so the mathematical derivation given above will not be valid. Second, retardation effects will occur, because conductivities in the unsaturated zone are generally much smaller than in the saturated zone and, consequently, water movement in the unsaturated zone lags behind saturated groundwater flow. Third, soil heterogeneity complicates the concept of residual water content and equilibrium moisture distributions in the unsaturated zone. Hence, in practice the change in storage will usually be smaller. Therefore, Equation (91) remains useful only when the concept of effective porosity is extended to match a particular situation. This gross simplification is justified in view of the fact that with the hydraulic approach groundwater flow is described in a simplified way. However, the concept of effective porosity has a clear physical interpretation, similar to storativity, i.e., the volume of water released by a phreatic aquifer per unit horizontal aquifer surface and per unit decline of the water table. Of course, the position of the water table, the speed of water table decline, and the soil heterogeneity play important roles in this.

When we combine the obtained storage equation with the net inflow, divide by $\rho \Delta x \Delta y$, and use the hydraulic flow approximation, given by Equations (74) and (76), the following hydraulic flow equation for a phreatic aquifer is obtained

$$n_e \frac{\partial h}{\partial t} = \frac{\partial}{\partial x} \left[K_e(h - b) \frac{\partial h}{\partial x} \right] + \frac{\partial}{\partial y} \left[K_e(h - b) \frac{\partial h}{\partial y} \right] + R \qquad (92)$$

FIGURE 3.7 Storage and release of groundwater in a phreatic aquifer due to changes in water table position: (A) deep water table, and (B) shallow water table.

Also, in this equation there are no terms with respect to the z-dimension, which is again the main advantage of the hydraulic approach. However, the equation is nonlinear, which complicates practical applications and makes finding exact analytical solutions more difficult. Nevertheless, the equation is extensively used, especially in numerical groundwater flow simulation models.

A simplification is possible for phreatic aquifers with a horizontal base, using the elevation of the water table above the base, $H = h - b$, as a dependent variable; the flow equation becomes

$$n_e \frac{\partial H}{\partial t} = \frac{\partial}{\partial x}\left(K_e H \frac{\partial H}{\partial x} \right) + \frac{\partial}{\partial y}\left(K_e H \frac{\partial H}{\partial y} \right) + R \tag{93}$$

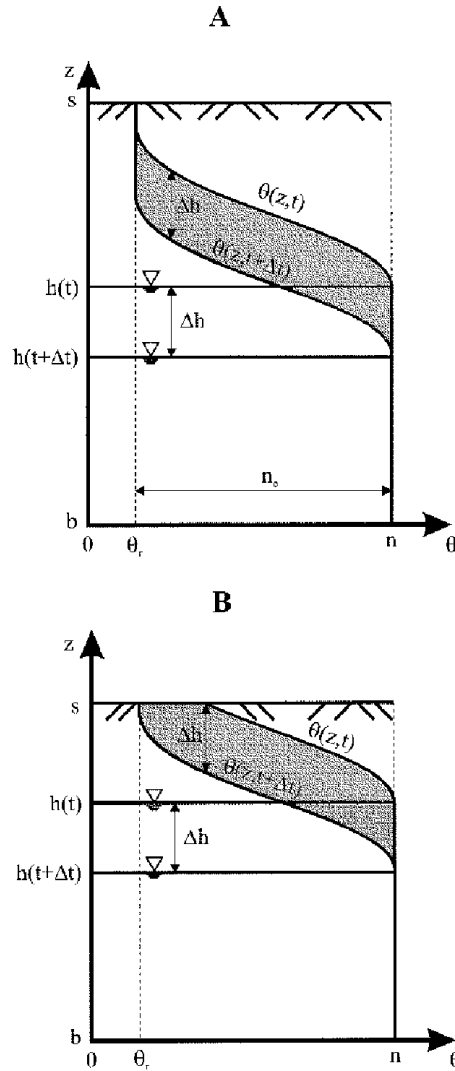This is called the *Boussinesq equation*. The equation is still nonlinear, but in the case of steady-state conditions, becomes linear in $H^2/2$

$$\frac{\partial}{\partial x}\left[K_e \frac{\partial}{\partial x}\left(\frac{H^2}{2}\right)\right] + \frac{\partial}{\partial y}\left[K_e \frac{\partial}{\partial y}\left(\frac{H^2}{2}\right)\right] + R = 0 \qquad (94)$$

Several useful solutions have been derived with this equation (Polubarinova-Kochina, 1962; Bear, 1972; Bear, 1979). However, care should be taken with respect to boundary conditions because these have to respect the hydraulic approach and consequently be independent of elevation. Consider for instance the very common boundary condition: a river in contact with a phreatic aquifer. The water level in the river can only be used as a potential boundary condition, when the river is wide and deep such that the groundwater beneath the river effectively has the same potential. When the surface water body is narrow and shallow, such as with drains or ditches, there can be significant vertical flow components resulting in a potential difference between the surface water and the groundwater below. In practice, such a situation can be treated with a third-type boundary condition, where a boundary resistance is introduced that takes into account the effect of converging groundwater flow toward the drain or ditch.

3.3.5 Flow Equation for an Aquitard

For obtaining a flow equation for an aquitard, it is sufficient to note that the flow is considered to be strictly vertical, such that the flow equation can be derived from the three-dimensional saturated groundwater flow Equation (56) by only considering the z-component of the flow

$$S_0 \frac{\partial h}{\partial t} = \frac{\partial}{\partial z}\left(K_v \frac{\partial h}{\partial z}\right) \qquad (95)$$

This equation describes flow through an aquitard, while the potentials of the aquifers bounding the aquitard can be used as boundary conditions. However, in practice this equation is not used very much. Usually, an additional simplifying assumption is made, namely that the specific storage coefficient is negligibly small, such that

$$\frac{\partial}{\partial z}\left(K_v \frac{\partial h}{\partial z}\right) = 0 \qquad (96)$$

This means that the flow is considered to be in a quasi-steady state, and the solution is given by Equation (72), i.e., the flow adjusts immediately to changes in the aquifer potentials above and below the aquitard. This simplification makes vertical flows through aquitards easy to evaluate, such that these can be incorporated as inputs or outputs in overlying or underlying aquifers. However, the approach is not very accurate, because conductivities of aquitards are much smaller than for aquifers, so that in general the flow through an aquitard lags behind the flow in the bounding aquifers, and time and storage effects can be important (Frind, 1983). However, lack of knowledge of aquitard properties often forces us to use a simplified approach.

For Further Information

The standard reference in the field of mathematical aspects related to groundwater hydrology is Bear (1972). This work gives a comprehensive coverage of the dynamics of fluid flow through porous media. This book has been reprinted (Bear, 1988) and is highly recommended for scientists and engineers. Also, parts have been revised or updated and republished (Bear, 1979) dealing with principles of regional groundwater management, with special emphasis on the hydraulic approach to groundwater flow. Bear and Verruijt (1987) discuss mathematical and numerical modeling of groundwater flow and pollution transport.

Another classic work is Polubarinova-Kochina (1962), which gives an extensive overview of mathematical aspects of groundwater flow, with many analytical solutions to practical problems. Other notable works are McWhorter and Sunada (1977), Todd (1980), and de Marsily (1986). Recent advances in the field of stochastic groundwater flow modeling are discussed in Dagan (1989). A recent and mathematical in-depth discussion of groundwater flow is Zijl and Nawalany (1993).

New developments can be found in journals. The most renowned is *Water Resources Research*, published by the American Geophysical Union. This journal covers all scientific hydrology subjects. Very popular is the *Journal of Hydrology*, published by Elsevier and recognized as one of the official journals of the European Geophysical Society. This journal also covers all aspects of hydrology. The most specialized journal in the field of mathematical groundwater hydrology is Transport in Porous Media, edited by J. Bear and published by Kluwer Academic Publishers.

References

Bear, J. 1972. *Dynamics of Fluids in Porous Media*. American Elsevier, New York.

Bear, J. 1979. *Hydraulics of Groundwater*. McGraw-Hill Inc., New York.

Bear, J. 1988. *Dynamics of Fluids in Porous Media*. Dover Publ., Inc., New York.

Bear, J. and Corapcioglu, M.Y. (Eds.) 1984. *Fundamentals of Transport Phenomena in Porous Media*. Martinus Nijhoff Publishers, Dordrecht, The Netherlands.

Bear, J. and Verruijt, A. 1987. *Modeling Groundwater Flow and Pollution*. D. Reidel Publ. Co., Dordrecht, The Netherlands.

Dagan, G., 1989. *Flow and Transport in Porous Formations*. Springer-Verlag, Berlin.

Frind, E. O. 1983. Exact aquitard response functions for multiple aquifer mechanics., in *Flow through Porous Media*, Ed. G.F. Pinder, CML U.K. Publications, Southampton, U.K., 86-92.

Halek, V. and Svec, J. 1979. *Groundwater Hydraulics*. Elsevier Scientific Publ. Co., Amsterdam.

Hantush, M. S. 1964. Hydraulics of wells, in *Advances in Hydrosciences*, vol. 1, Ed. V.T. Chow, Academic Press, New York, 281-432.

Harr, M. E. 1962. *Groundwater and Seepage*. McGraw-Hill Inc., New York (reprinted in 1990 by Dover Publ. Inc., New York).

Huisman, L. 1972. *Groundwater Recovery*. Macmillan Press Ltd., London, U.K.

Kruseman, G. P. and de Ridder, N. A. 1991. *Analysis and Evaluation of Pumping Test Data*. International Institute for Land Reclamation and Improvement/ILRI. Wageningen, The Netherlands.

de Marsily, G. 1986. *Quantitative Hydrogeology*, Academic Press, New York.

McWhorter, P. B. and Sunada, D. K. 1977. *Groundwater Hydrology and Hydraulics*. Water Resources Publications, Colorado.

Pinder, G. P. (Ed.) 1983. *Flow through Porous Media*. CML Publications, Southampton, U.K.

Polubarinova-Kochina, P. Ya. 1962. *Theory of Ground Water Movement*. Princeton University Press, Princeton, New Jersey.

Strack, O. D. L. 1989. *Groundwater Mechanics*. Prentice Hall, Inc., Englewood Cliffs, New Jersey.

Todd, D. K. 1980. *Groundwater Hydrology*, 2nd ed. Wiley, New York.

Verruijt, A. 1969. Elastic storage of aquifers, in *Flow through Porous Media*, Ed. J.M. De Wiest, Academic Press, New York, 331-376.

Verruijt, A. 1970. *Theory of Groundwater Flow*. Macmillian, London, U.K.

Walton, W. A. 1970. *Groundwater Resource Evaluation*. McGraw-Hill, Inc., New York.

Zijl, W. and Nawalany, M. 1993. *Natural Groundwater Flow*. Lewis Publishers, Boca Raton, Florida.

Glossary

Boundary Conditions   All partial differential equations have an infinite number of possible solutions, each of which corresponds to a particular case; to obtain from this multitude of possible solutions one particular solution corresponding to a certain specific case, it is necessary to provide supplementary information that is not contained in the partial differential equation; the supplementary information is boundary conditions and should include the initial state of the considered situation and how the considered domain interacts with its surroundings.

Conservation of Mass   Also termed continuity equation; a mathematical statement of the mass balance, i.e., in the absence of mass-producing or mass-inhaling processes, the mass present in an arbitrary portion of a moving fluid remains constant.

Del Operator   Symbolic notation of a differential vector operator; several important operations such as gradient, divergence, and curve can be described with this operator.

Elementary Control Volume   A volume of infinitesimal size, which is used to equate some physical property or process; when the size of the volume is taken in the limit going to zero, a mathematical equation is obtained that describes the property or process in every point of the continuum.

Land Subsidence   A phenomenon that involves the lowering or settling of the land surface due to various factors, mostly under the impact of man's activities; often the extensive exploitation of groundwater resources has been accompanied in many places by significant land subsidence; in each case, the ground layers consist of a sequence of sand and gravel intermixed with poorly consolidated and highly compressible clays; the pumping of water from the conductive layers induces a drop in water potentials and gradual drainage of water from the clays, which leads to compaction of the formations and sinking of the land surface.

Navier-Stokes Equation   Equation describing the motion of a Newtonian fluid, i.e., a viscous fluid for which the viscous stress components are linearly related to the deformation; when the Navier-Stokes equation is put in dimensionless form the Reynolds number appears, which expresses the ratio of inertia to viscous forces; for very small Reynolds numbers the inertia terms can be neglected and the Navier-Stokes equation reduces to the Stokes equation, describing a situation called Stokes flow, or creeping motion; the movement of groundwater under natural conditions is always considered to be creeping because of the very small groundwater flow velocities.

Newton's Second Law of Motion   Also termed momentum or impulse equation; mathematical statement of the physical fact that forces are needed to induce motion or to change motion; the law states that the time rate of change of an arbitrary portion of a moving fluid is equal to the resulting forces acting upon the considered portion.

Piezometer   Observation well tapping a saturated porous formation at a certain position and depth, indicating the groundwater potential present in the formation at that point.

Potential Flow Problem   Potential flow is used to denote movement of fluids, without rotations, which is a sufficient condition for the existence of a potential function, such that the gradient of that function gives the direction and magnitude of the flow; in a homogeneous medium, potential functions are harmonic functions as they are solutions of the Laplace equation; there are many mathematical techniques for solving these types of problems. Especially two-dimensional flow problems can be solved with the theory of complex functions.

Taylor Series Expansion   Let $f(x)$ be a differentiable function of x, then the following series expansion holds

$$f(x) = \sum_{n=0}^{\infty} \frac{x - x_0}{n!} \frac{d^n f(x_0)}{dx^n}$$

$$= f_0 + \Delta x \frac{df_0}{dx} + \frac{(\Delta x)^2}{2} \frac{d^2 f_0}{dx^2} + \frac{(\Delta x)^3}{6} \frac{d^3 f_0}{dx^3} + \dots$$

where $\Delta x = x - x_0$, and $f_0 = f(_0)$.

4

Groundwater and Seepage: Accounting for Variability

Milton E. Harr
Purdue University

4.1   Introduction

Figure 4.1 shows the pore space available for flow in two highly idealized soil models: *regular cubic* and *rhombohedral*. It is seen that even for these special cases, the pore space is not regular, but consists of cavernous cells interconnected by narrower channels. Pore spaces in real soils can range in size from molecular interstices to tunnel-like caverns. They can be spherical (as in concrete) or flat (as in clays), or display irregular patterns which defy description. Add to this the fact that pores may be *isolated* (inaccessible) or *interconnected* (accessible from both ends) or may be *dead-ended* (accessible through one end only).

In spite of the apparent irregularities and complexities of the available pores, there is hardly an industrial or scientific endeavor that does not concern itself with the passage of matter, solid, liquid, or gaseous, into, out of, or through porous media. Contributions to the literature can be found among such diverse fields (to name only a few) as soil mechanics, groundwater hydrology, petroleum, chemical and metallurgical engineering, water purification, materials of construction (ceramics, concrete, timber, paper), chemical industry (absorbents, varieties of contact catalysts, and filters), pharmaceutical industry, traffic flow, and agriculture.
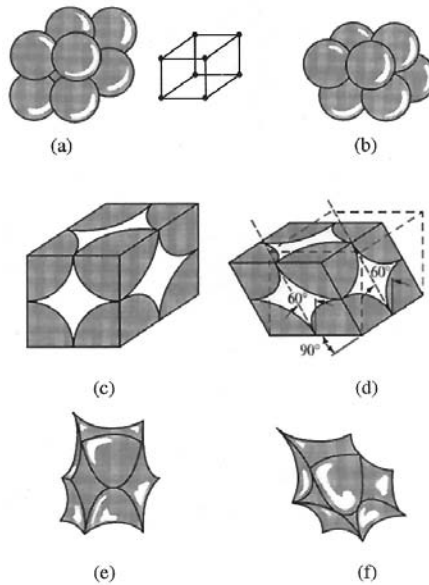
FIGURE 4.1 Idealized void space.

In no field of engineering are practitioners faced with more complex and uncertain sets of conditions than those concerned with groundwater systems. Unlike their colleagues in other engineering disciplines, who have the advantage of observing the performance of many prototypes under relatively known states of input, groundwater engineers generally deal with systems that are custom-built and tailored to special demands and specific locations. Add to these the very size of projects and the expense of underground sampling and data acquisition.

This chapter was written with the following main objectives.

1. To provide the groundwater engineer with an organized, logical, and systematic body of knowledge for the solution of groundwater and seepage problems.
2. To provide the groundwater engineer with probabilistic concepts and techniques to account for uncertainty in designs and to illustrate their relevance with application to practical problem situations.

4.2   Some Groundwater Fundamentals

The literature is replete with derivations and analytical excursions of the basic equations of steady-state groundwater flow (Polubarinova-Kochina 1952; Harr, 1962; Cedergren, 1967; Bear, 1972; Domenico and Schwartz, 1990; to name only a few). A summary and brief discussion of these will be presented below for the sake of completeness.

4.2.1   Bernoulli's Equation

Underlying the analytical approach to groundwater flow is the representation of the actual physical system by a tractable mathematical model. In spite of their inherent shortcomings, many such analytical models have demonstrated considerable success in simulating the action of their prototypes.

As is well known from fluid mechanics, for steady flow of nonviscous incompressible fluids, Bernoulli's equation (Lamb, 1945)

$$\frac{p}{\gamma_w} + z + \frac{\bar{v}^2}{2g} = \text{constant} = h \tag{1}$$

where

h = total head, ft

p = pressure, lb/ft²

γ_w = unit weight of fluid, lb/ft³

g = gravitational constant, 32.2 ft/sec²

$\bar{v}$ = seepage velocity, ft/sec

demonstrates that the sum of the *pressure head*, p/γ_w, *elevation head*, z, and *velocity head*, $\bar{v}^2/2g$, at any point within the region of flow is a constant.

To account for the loss of energy due to the viscous resistance within the individual pores, Bernoulli's equation is taken as

$$\frac{p_A}{\gamma_w} + z_A + \frac{\bar{v}_A^2}{2g} = \frac{p_B}{\gamma_w} + z_B + \frac{\bar{v}_B^2}{2g} + \Delta h \tag{2}$$

where Δh represents the total head loss (energy loss per unit weight of fluid) of the fluid over the distance Δs. The ratio

$$i = -\lim_{\Delta s \to 0} \frac{\Delta h}{\Delta s} = -\frac{dh}{ds} \tag{3}$$

is called the *hydraulic gradient* and represents the space rate of energy dissipation per unit weight of fluid (a pure number).

In most problems of interest the velocity heads (the kinetic energy) are so small they can be neglected. For example, a velocity of 1 ft/sec, which is large compared to typical seepage velocities through soils, produces a velocity head of only 0.015 ft. Hence, Equation (2) can be simplified to

$$\frac{p_A}{\gamma_w} + z_A = \frac{p_B}{\gamma_w} + z_B + \Delta h$$

and the total head at any point in the flow domain is simply

$$h = \frac{p}{\gamma_w} + z \tag{4}$$

4.2.2 Darcy's Law

The flow of groundwater is taken to be governed by *Darcy's law*, which states that the velocity of the flow is directly proportional to the hydraulic gradient. A similar statement in an electrical system is *Ohm's law* and in a thermal system, *Fourier's law*. The grandfather of all such relations is *Newton's law of motion*. Table 4.1 presents some other points of similarity.

Prior to 1856, the formidable nature of the flow through porous media defied rational analysis. In that year, Henry Darcy published a simple relation based on his experiments on the flow of water in vertical sand filters in *Les Fontaines Publiques de la Ville de Dijon*, namely,

TABLE 4.1 Some Similarities of Flow Models

Form of Energy	Name of Law	Quantity	Storage	Resistance
Electrical	Ohm's law	Current (voltage)	Capacitor	Resistor
Mechanical	Newton's law	Force (velocity)	Mass	Damper
Thermal	Fourier's law	Heat flow (temperature)	Heat capacity	Heat resistance
Fluid	Darcy's law	Flow rate (pressure)	Liquid storage	Permeability

$$v = ki = -k\frac{dh}{ds} \tag{5}$$

Equation (5), commonly called Darcy's law, demonstrates a linear dependency between the hydraulic gradient and the *discharge velocity* v. The discharge velocity, $v = n\bar{v}$, is the product of the porosity n and the seepage velocity, $\bar{v}$. The coefficient of proportionality k in Equation (5) is called by many names depending on its use; among these are the *coefficient of permeability*, *hydraulic conductivity*, and *permeability constant*. As shown in Equation (5), k has the dimensions of a velocity. It should be carefully noted in this equation that flow is a consequence of differences in total head and not of pressure gradients. This is demonstrated in Figure 4.2 where the flow is directed from A to B, even though the pressure at point B is *greater* than that at point A.
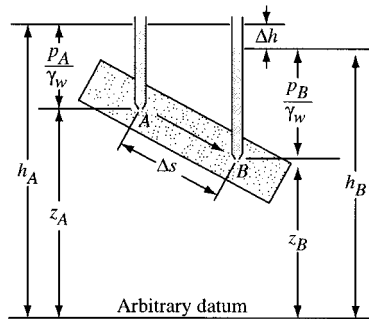
FIGURE 4.2 Heads in Bernoulli's equation.

Defining Q as the total volume of flow per unit time through a cross-sectional area *A*, Darcy's law takes the form

$$Q = Av = Aki = -Ak\frac{dh}{ds} \tag{6}$$

Darcy's law offers the single parameter *k* to account for both the characteristics of the medium and the fluid. It has been found that *k* is a function of γ_w, the unit weight of the fluid, μ, the coefficient of viscosity, and *n*, the porosity, as given by

$$k = C\frac{\gamma_w n}{\mu} \tag{7}$$

where *C* (dimensionally an area) typifies the structural characteristics of the media independent of the fluid properties. The principal advantage of Equation (7) lies in its use when dealing with more than one fluid or with temperature variations. When employing a single relatively incompressible fluid sub-

TABLE 4.2 Some Typical Values of Coefficient of Permeability

Soil Type	Coefficient of Permeability k, cm/s
Clean gravel	1.0 and greater
Clean sand (coarse)	1.0–0.01
Sand (mixtures)	0.01–0.005
Fine sand	0.05–0.001
Silty sand	0.002–0.0001
Silt	0.0005–0.00001
Clay	0.000001 and smaller

jected to small changes in temperature, such as in groundwater and seepage-related problems, it is more convenient to use k as a single parameter. Some typical values for k are given in Table 4.2.

Although Darcy's law was obtained initially from considerations of one-dimensional macroscopic flow, its practical utility lies in its generalization into two or three spatial dimensions. Accounting for the directional dependence of the coefficient of permeability, Darcy's law can be generalized to

$$v_s = -k_s \frac{\partial h}{\partial s} \qquad (8)$$

where k_s is the coefficient of permeability in the "s" direction, and v_s and $\partial h/\partial s$ are the components of the velocity and the hydraulic gradient, respectively, in that direction.

4.2.3   Reynolds Number

There remains now the question of the determination of the extent to which Darcy's Law is valid in actual flow systems through soils. Such a criterion is furnished by the Reynolds number R (a pure number relating inertial to viscous force), defined as

$$R = \frac{vd\rho}{\mu} \qquad (9)$$

where

v = discharge velocity, cm/sec
d = average of diameter of particles, cm
ρ = density of fluid, g(mass)/cm^3
μ = coefficient of viscosity, g-sec/cm^2

The critical value of the Reynolds number at which the flow in aggregations of particles changes from laminar to turbulent flow has been found by various investigators (see Muskat, 1937) to range between 1 and 12. However, it will generally suffice to accept the validity of Darcy's law when the Reynolds number is taken as equal to or less than unity, or

$$\frac{vd\rho}{\mu} \leq 1 \qquad (10)$$

Substituting the known values of ρ and μ for water into Equation (10) and assuming a conservative velocity of 1/4 cm sec, we have d equal to 0.4 mm, which is representative of the average particle size of coarse sand.

4.2.4   Homogeneity and Isotropy

If the coefficient of permeability is independent of the direction of the velocity, the medium is said to be *isotropic*. Moreover, if the same value of the coefficient of permeability holds at all points within the region of flow, the medium is said to be *homogeneous* and *isotropic*. If the coefficient of permeability is dependent on the direction of the velocity and if this directional dependence is the same at all points of the flow region, the medium is said to be *homogeneous* and *anisotropic* (or *aleotropic*).

4.2.5   Streamlines and Equipotential Lines

Physically, all flow systems extend in three dimensions. However, in many problems the features of the motion are essentially planar, with the flow pattern being substantially the same in parallel planes. For these problems, for steady-state incompressible, isotropic flow in the xy plane, it can be shown (Harr, 1962) that the governing differential equation is

$$k_x \frac{\partial^2 h}{\partial x^2} + k_y \frac{\partial^2 h}{\partial y^2} = 0 \tag{11}$$

Here the function $h(x, y)$ is the distribution of the total head (of energy available to do work) within and on the boundaries of a flow region, and k_x and k_y are the coefficients of permeability in the x- and y-directions, respectively. If the flow system is isotropic, $k_x = k_y$, and Equation (11) reduces to

$$\frac{\partial^2 h}{\partial x^2} + \frac{\partial^2 h}{\partial y^2} = 0 \tag{12}$$

Equation (12), called *Laplace's equation*, is the governing relationship for steady-state, laminar-flow conditions (Darcy's law is valid). The general body of knowledge relating to Laplace's equation is called *potential theory*. Correspondingly, incompressible steady-state fluid flow is often called *potential flow*. The correspondence is more evident upon the introduction of the *velocity potential* ϕ, defined as

$$\phi(x, y) = -kh + C = -k\left(\frac{p}{\gamma_w} + z\right) + C \tag{13}$$

where h is the total head, p/γ_w is the pressure head, z is the elevation head, and C is an arbitrary constant. It should be apparent that, for isotropic conditions,

$$v_x = \frac{\partial \phi}{\partial x} \qquad v_y = \frac{\partial \phi}{\partial y} \tag{14}$$

and, Equation (12) will produce

$$\frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} = 0 \tag{15}$$

The particular solutions of Equations (12) or (15) which yield the locus of points within a porous medium of equal potential, curves along with $h(x, y)$ or $\phi(x, y)$ are equal to a series of constants, are called *equipotential lines*.

In analyses of groundwater flow, the family of flow paths is given by the function $\psi(x, y)$, called the *stream function*, defined in two-dimensions as (Harr, 1962)

$$v_x = \frac{\partial \psi}{\partial y} \qquad v_y = -\frac{\partial \psi}{\partial x} \qquad (16)$$

where v_x and v_y are the components of the velocity in the x- and y-directions, respectively.

Equating the respective potential and stream functions of v_x and v_y produces

$$\frac{\partial \phi}{\partial x} = \frac{\partial \psi}{\partial y} \qquad \frac{\partial \phi}{\partial y} = -\frac{\partial \psi}{\partial x} \qquad (17)$$

Differentiating the first of these equations with respect to y and the second with respect to x and subtracting, we obtain Laplace's equation

$$\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} = 0 \qquad (18)$$

We shall examine the significance of this relationship following a little more discussion of the physical meaning of the stream function.

Consider AB of Figure 4.3 as the path of a particle of water passing through point P with a tangential velocity v. We see from the figure that

$$\frac{v_y}{v_x} = \tan \theta = \frac{dy}{dx}$$

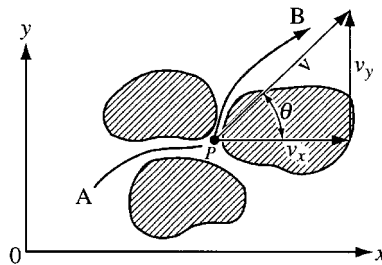and hence

$$v_y \, dx - v_x \, dy = 0 \qquad (19)$$

FIGURE 4.3 Path of flow.

Substituting Equation (16), it follows that

$$\frac{\partial \psi}{\partial x} \, dx + \frac{\partial \psi}{\partial y} \, dy = 0$$

which states that the total differential $d\psi = 0$ and

$$\psi(x, y) = \text{constant}$$

Thus we see that the family of curves generated by the function $\psi(x, y)$ equal to a series of constants are tangent to the resultant velocity at all points in the flow region and hence define the path of flow.

The potential, $\phi = -kh +$ constant, is a measure of the energy available at a point in the flow region to move the particle of water from that point to the tailwater surface. Recall the locus of points of equal energy, say $\phi(x, y) =$ constants, are called equipotential lines. The total differential along any curve, $\phi(x, y) =$ constant, produces

$$\mathrm{d}\phi = \frac{\partial \phi}{\partial x} dx + \frac{\partial \phi}{\partial y} dy = 0$$

Substituting for $\partial\phi/\partial x$ and $\partial\phi/\partial y$ from Equation (16), we have

$$v_x\, dx + v_y\, dy = 0$$

and

$$\frac{dy}{dx} = -\frac{v_x}{v_y} \tag{20}$$

Noting the negative reciprocal relationship between their slopes, Equations (19) and (20), we see that, within the flow domain, the families of streamlines $\psi(x, y) =$ constants and equipotential lines $\phi(x, y) =$ constants intersect each other at right angles. It is customary to signify the sequence of constants by employing a subscript notation, such as $\phi(x, y) = \phi_i$, $\psi(x, y) = \psi_j$ (Figure 4.4).
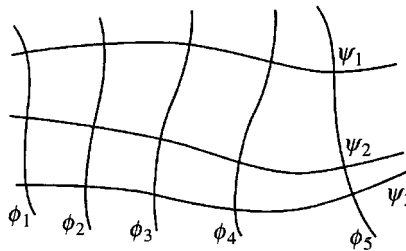
FIGURE 4.4 Streamlines and equipotential lines.

As only one streamline may exist at a given point within the flow medium, streamlines cannot intersect one another if flow is to occur between them. Consequently, if the medium is saturated, any pair of streamlines act to form a flow channel between them.

Consider the flow between the two streamlines ψ and $\psi + d\psi$ in Figure 4.5; v represents the resultant velocity of flow. The quantity of flow through the flow channel per unit length normal to the plane of flow (say, cubic feet per second per foot) is

$$dQ = v_x\, ds\, \cos\theta - v_y\, ds \sin\theta$$

$$= v_x\, dy - v_y\, dx$$

$$= \frac{\partial \psi}{\partial y} dy + \frac{\partial \psi}{\partial x} dx$$

and
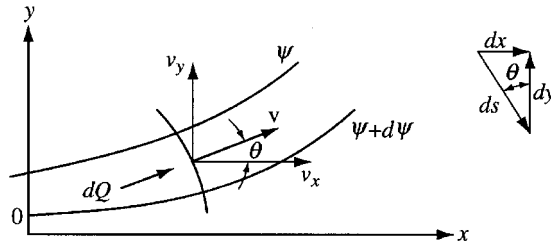
FIGURE 4.5 Flow between streamlines.

$$dQ = d\psi \tag{21}$$

Hence the quantity of flow (also-called the *quantity of discharge* and *discharge quantity*) between any pair of streamlines is a constant whose value is numerically equal to the difference in their respective ψ values. Thus, once a sequence of streamlines of flow has been obtained, with neighboring ψ values differing by a constant amount, their plot will not only show the expected direction of flow but the relative magnitudes of the velocity along the flow channels; that is, the velocity at any point in the flow channel varies inversely with the streamline spacing in the vicinity of that point.

An equipotential line was defined previously as the locus of points where there is an expected level of available energy sufficient to move a particle of water from a point on that line to the tailwater surface. Thus, it is convenient to reduce all energy levels relative to a tailwater datum. For example, a piezometer located anywhere along an equipotential line, say at 0.75h in Figure 4.6, would display a column of water extending to a height of $0.75h$ above the tailwater surface. Of course, the pressure in the water along the equipotential line would vary with its elevation, Equation (4). This will be illustrated in Section 4.4.
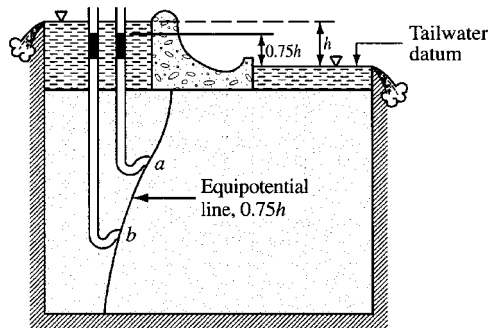
FIGURE 4.6 Pressure head along equipotential line.

4.3   Some Probabilistic Fundamentals

Within the context of engineering usage there are two primary definitions of the concept of probability: *relative frequency* and *subjective* interpretation. Historically, the measure first offered for the probability of an outcome was its *relative frequency*. If an outcome A can occur T times in N equally likely trials, the probability of the outcome A is

$$P[A] = \frac{T}{N} \tag{22a}$$

Implied in Equation (22a) is that the probability of an outcome A equals the number of outcomes favorable to A (within the meaning of the experiment) divided by the total number of possible outcomes, or

$$P[A] = \frac{\text{Favorable outcomes}}{\text{Total possible outcomes}} \qquad (22b)$$

Example 1: Find the probability of drawing a red card from an ordinary well-shuffled deck of 52 cards.

Solution: Of the 52 equally likely outcomes, there are 26 favorable (red card) outcomes. Hence,

$$P[\text{drawing red card}] = \frac{26}{52} = \frac{1}{2}$$

Understood in the example is that if one were to repeat the process a very large number of times, a red card would appear in one half of the trials. This is an example of the relative frequency interpretation. Now, what meaning could be associated with the statement, the probability of the failure of a proposed dam is 2% (P[failure] = 0.02)? The concept of repeated trials is meaningless: the structure will be built only *once* and it will either fail or be successful during its design lifetime. It can not do both. Here we have an example of the *subjective interpretation* of probability. It is a measure of information as to the likelihood of the occurrence of an outcome.

Subjective probability is generally more useful than the relative frequency concept in engineering applications. However, the basic rules governing both are identical. As an example, we note that both concepts specify the probability of an outcome to range between and to include numerical values from zero to one. The lower limit indicates there is *no* likelihood of occurrence: the upper limit corresponds to a *certain* outcome.

$$\textbf{<Axiom I>} \quad 0 \le P[A] \le 1 \qquad (23a)$$

The certainty of an outcome C is a probability of unity

$$\textbf{<Axiom II>} \quad P[C] = 1 \qquad (23b)$$

Equations (23) provide two of the three axioms of the theory of probability. The third axiom requires the concept of *mutually exclusive* outcomes. Outcomes are mutually exclusive if they cannot occur simultaneously. The third axiom states the probability of the occurrence of the sum of a number of mutually exclusive outcomes $A(1)$, $A(2)$, …, $A(N)$ is the sum of their individual probabilities (*addition rule*), or

$$\textbf{<Axiom III>} \quad P[A(1) + A(2) + \ldots + A(N)] = P[A(1)] + P[A(2)] + \ldots + P[A(N)] \qquad (23c)$$

As a very important application of these axioms, consider the proposed design of a spillway. After construction, only one of two outcomes can obtain in the absolute structural sense: either it is successful or it fails. These are mutually exclusive outcomes. They are also exhaustive in that, within the sense of the example, no other outcomes are possible. Hence, the second axiom, Equation (23b), requires

$$P[\text{success} + \text{failure}] = 1$$

Since they are mutually exclusive, the third axiom specifies that

$$P[\text{success}] + P[\text{failure}] = 1$$

The probability of the success of a structure is called its *reliability*, *R*. Symbolizing the probability of failure as *p*(*f*), we have the important expression

$$R + p(f) = 1 \tag{24}$$

4.3.1   Moments

Consider a system of *discrete* parallel (vertical) forces, $P(1)$, $P(2)$, …, $P(N)$, acting on a rigid beam at the respective distances $x(1)$, $x(2)$, …, $x(N)$ (Figure 4.7a). From statics we have that the magnitude of the *equilibrant* M, is

$$M = \sum_{i=1}^{N} P(i) \tag{25a}$$

and its point of application, $\bar{x}$, is

$$\bar{x} = \frac{\sum_{i=1}^{N} x(i)P(i)}{\sum_{i=1}^{N} P(i)} \tag{25b}$$

For a *continuously* distributed parallel force system (Figure 4.7b) over a finite distance, say from $x(a)$ to $x(b)$, the corresponding expressions are

$$M = \int_{x(a)}^{x(b)} p(x)\,dx \tag{26a}$$

and

$$\bar{x} = \frac{\int_{x(a)}^{x(b)} x p(x)\,dx}{M} \tag{26b}$$

Suppose that the discrete forces $P(i)$ in Figure 4.7a represent the frequencies of the occurrence of the N outcomes $x(1)$, $x(2)$, …, $x(N)$. As the distribution is exhaustive, from axiom II, Equation (23b), the magnitude of the equilibrant must be unity, $M = 1$. Hence, Equation (25b) becomes

$$<\text{Discrete}> \quad E[x] = \bar{x} = \sum_{i=1}^{N} x(i)P(i) \tag{27a}$$

Similarly, for the continuous distribution (Figure 4.7b), as all probabilities $p(x)\,dx$ must lie between $x(a)$ and $x(b)$, in Equation (26a) $M = 1$. Hence, Equation (26b) becomes

Discrete

(a)

Continuous

(b)
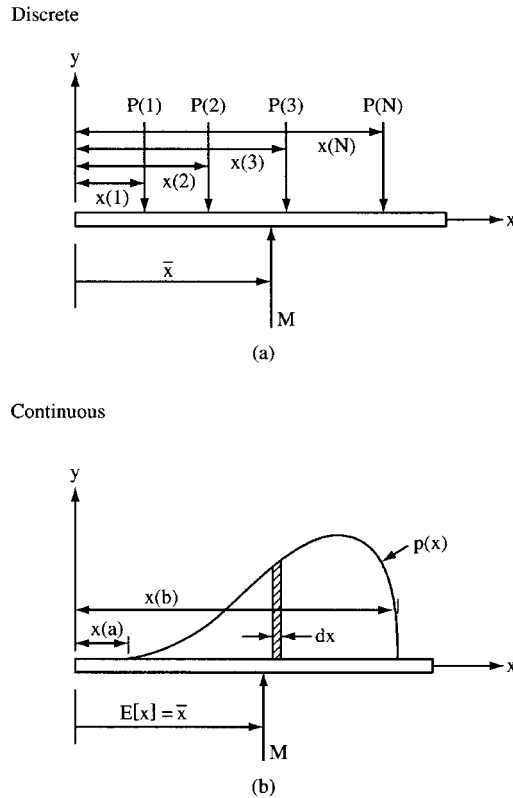
FIGURE 4.7 Equilibrant for discrete and continuous distributions.

$$<\text{Continuous}> \quad E[x] = \bar{x} = \int_{x(a)}^{x(b)} xp(x)\,dx \tag{27b}$$

The symbol $E[x]$ in Equations (27) is called the *expected value* or the *expectation* or simply the *mean* of the variable *x*. As is true of the equilibrant, it is a measure of the central tendency, the *center of gravity* in statics.

Example 2: What is the expected value of the number of dots that will appear if a fair die is tossed?

Solution: Here each of the possible outcomes 1, 2, 3, 4, 5, and 6 has the equal probability of $P(i) = 1/6$ of appearing. Hence, from Equation (27a)

$$E[\text{toss of a fair die}] = \frac{1}{6}\big[1 + 2 + 3 + 4 + 5 + 6\big] = 3.5$$

We note in the above example that the expected value of 3.5 is an impossible outcome. There is no face on the die that will show 3.5 dots; however, it is still the best measure of the central tendency.

Example 3: Find the expected value of a continuous probability distribution wherein all values are equally likely to occur (called a *uniform distribution* [Figure 4.8]) between $y(a) = 0$ and $y(b) = 1/2$.

Solution: From Equation (26a) as $M = 1$ and $p(y) = C$ is a constant, we have
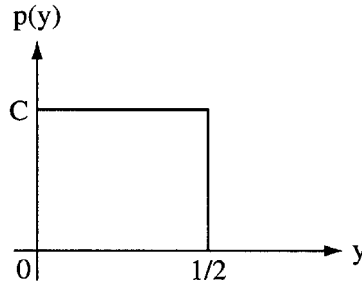
FIGURE 4.8 Uniform distribution.

$$1 = C \int_0^{1/2} dy, \quad C = 2$$

From Equation (27b), $E[y] = 2 \int_0^{1/2} y\,dy = 1/4$, as expected.

As the variables x and y in the above examples are determined by the outcomes of random experiments they are said to be *random variables*. In classical probability theory, random variables are generally represented by capital letters, such as *X* and *Y*. The individual values are customarily denoted by their corresponding lower-case letters *x* and *y*; however, no such distinction will be made here.

The expected value (mean) provides the locus of the central tendency of the distribution of a random variable. To characterize other attributes of the distribution, recourse is had to higher moments. Again, returning to statics, a measure of the dispersion of the distribution of the force system about the centroidal axis, at $x = E[x]$ in Figure 4.7b, is given by the *moment of inertia* (the *second central moment*),

$$I(y) = \int_{x(a)}^{x(b)} (x - \bar{x})^2 p(x)\,dx \tag{28}$$

The equivalent measure of the scatter (variability) of the distribution of a random variable is called its *variance*, denoted in this text as $v[x]$ (lower case "v") and defined as

$$\langle \text{Discrete} \rangle \quad v[x] = \sum_{\text{all } x(i)} [x(i) - \bar{x}]^2 P(i) \tag{29a}$$

$$\langle \text{Continuous} \rangle \quad v[x] = \int_{x(a)}^{x(b)} (x - \bar{x})^2 p(x)\,dx \tag{29b}$$

In terms of the expectation these can be written as

$$v[x] = E\left[(x - \bar{x})^2 \right] \tag{30}$$

which, after expansion, leads to a form more amenable to computations,

$$v[x] = E[x^2] - (E[x])^2 \tag{31}$$

This expression is the equivalent of the *parallel-axis theorem* for the moment of inertia.

Example 4: Find the expected value, and the variance of the *exponential distribution*, $p(x) = a \exp(-ax)$, $x > 0$, a is a constant.

Solution: We first show that $p(x)$ is a valid probability distribution

$$\int_0^\infty p(x)\,dx = a\int_0^\infty e^{-ax}\,dx = 1, \quad \text{Q.E.D.}$$

The expected value is

$$E[x] = a\int_0^\infty xe^{-ax}\,dx = \frac{1}{a}$$

Continuing,

$$E[x^2] = a\int_0^\infty x^2 e^{-ax}\,dx = \frac{1}{a^2}$$

whence, using Equation (31),

$$v[x] = \frac{2}{a^2} - \left(\frac{1}{a}\right)^2 = \frac{1}{a^2}$$

The variance has the units of the square of those of the random variable. A more meaningful measure of dispersion of a random variable (x) is the positive square root of its variance (compare with *radius of gyration* of mechanics) called the *standard deviation*, $\sigma[x]$,

$$\sigma[x] = \sqrt{v[x]} \tag{32}$$

From the results of the previous example, it is seen that the standard deviation of the exponential distribution is $\sigma[x] = 1/a$.

An extremely useful relative measure of the scatter of a random variable $\sigma[x]$ is its *coefficient of variation* $V(x)$, capital "V" with parenthesis, usually expressed as a percentage,

$$V(x) = \frac{\sigma[x]}{E[x]} \times 100(\%) \tag{33}$$

For the exponential distribution we found, $\sigma[x] = 1/a$ and $E[x] = 1/a$, hence V (exponential distribution) = 100%. In Table 4.3 are given representative values of the coefficients of variation. Original sources should be consulted for details.

TABLE 4.3 Representative Coefficient of Variation

Parameter	Coefficient of Variation, %	Source
Soil		
Porosity	10	Schultze (1972)
Specific gravity	2	Padilla and Vanmarcke (1974)
Water content		
Silty clay	20	Padilla and Vanmarcke (1974)
Clay	13	Fredlund and Dahlman (1972)
Degree of saturation	10	Fredlund and Dahlman (1972)
Unit weight	3	Hammitt (1966)
Coefficient of permeability	(240 at 80% saturation to 90 at 100% saturation)	Nielsen et al. (1973)
Compressibility factor	16	Padilla and Vanmarcke (1974)
Preconsolidation pressure	19	Padilla and Vanmarcke (1974)
Compression index		
Sandy clay	26	Lumb (1966)
Clay	30	Fredlund and Dahlman (1972)
Standard penetration test	26	Schultze (1975)
Standard cone test	37	Schultze (1975)
Friction angle ϕ		
Gravel	7	Schultze (1972)
Sand	12	Schultze (1972)
c, strength parameter (cohesion)	40	Fredlund and Dahlman (1972)
Structural Loads, 50-Year Maximum		
Dead load	10	Ellingwood et al. (1980)
Live load	25	Ellingwood et al. (1980)
Snow load	26	Ellingwood et al. (1980)
Wind load	37	Ellingwood et al. (1980)
Earthquake load	>100	Ellingwood et al. (1980)
Structural Resistance		
Structural steel		
Tension members, limit state, yielding	11	Ellingwood et al. (1980)
Tension members, limit state, tensile strength	11	Ellingwood et al. (1980)
Compact beam, uniform moment	13	Ellingwood et al. (1980)
Beam, column	15	Ellingwood et al. (1980)
Plate, girders, flexure	12	Ellingwood et al. (1980)
Concrete members		
Flexure, reinforced concrete, grade 60	11	Ellingwood et al. (1980)
Flexure, reinforced concrete, grade 40	14	Ellingwood et al. (1980)
Flexure, cast-in-place beams	8–9.5	Ellingwood et al. (1980)
Short columns	12–16	Ellingwood et al. (1980)
Ice		
Thickness	17	Bercha (1978)
Flexural strength	20	Bercha (1978)
Crushing strength	13	Bercha (1978)
Flow velocity	33	Bercha (1978)
Wood		
Moisture	3	Borri et al. (1983)
Density	4	Borri et al. (1983)
Compressive strength	19	Borri et al. (1983)
Flexural strength	19	
Glue-laminated beams		Borri et al. (1983)
Live load	18	Galambos et al. (1982)
Snow load	18	Galambos et al. (1982)

The coefficient of variation expresses a measure of the reliability of the central tendency. For example, given a mean value of a parameter of 10 with a coefficient of variation of 10% would indicate a standard deviation of 1, whereas a similar mean with a coefficient of variation of 20% would demonstrate a standard deviation of 2. The coefficient of variation has been found to be a fairly stable measure of variability for homogeneous conditions. Additional insight into the standard deviation and the coefficient of variation as measures of uncertainty is provided by *Chebyshev's inequality* [for the derivation see Lipschutz, 1965].

The spread of a random variable is often spoken of as its *range*, the difference between the largest and smallest outcomes of interest. Another useful measure is the range between the mean plus-and-minus-*h*-standard deviations, $\bar{x} \pm h\sigma [x]$, called the *h-sigma bounds* (Figure 4.9). If *x* is a random variable with mean value $\bar{x}$ and standard deviation σ, then Chebyshev's inequality states

$$P\left[\left(\bar{x} - h\sigma\right) \le x \le \left(\bar{x} + h\sigma\right)\right] \ge \frac{1}{h^2} \tag{34}$$
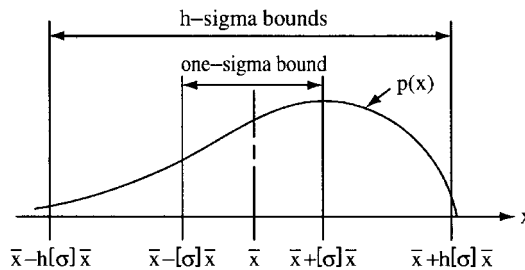
FIGURE 4.9 Range in *h*-sigma bounds.

In words, it asserts that for any probability distribution (with finite mean and standard deviation) the probability that random values of the variate will lie within *h*-sigma bounds is at least $(1 - 1/h^2)$. Some numerical values are given in Table 4.4. It is seen that quantitative probabilistic statements can be made without complete knowledge of the probability distribution function: only its expected value and coefficient of variation (or standard deviation) are required. In this regard, the values for the coefficients of variation given in the table may be used in the absence of more definitive information.

TABLE 4.4 Probabilities for Range of Expected Values ± *h*-Sigma Bounds

h	Chebyshev's Inequality	Gauss's Inequality	Exact Exponential Distribution	Exact Normal Distribution	Exact Uniform Distribution
1/2	0	0	0.78	0.38	0.29
1	0	0.56	0.86	0.68	0.58
2	0.75	0.89	0.95	0.96	1.00
3	0.89	0.95	0.982	0.9973	1.00
4	0.94	0.97	0.993	0.999934	1.00

Example 5: The expected value for the ϕ-strength parameter of a sand is 30°. What is the probability that a random sample of this sand will have a ϕ-value between 20° and 40°?

Solution: From Table 4.3, $V(\phi) = 12\%$ for sand; hence, $\sigma[\phi]$ is estimated to be $(0.12)(30) = 3.6°$ and $h = (\phi - \bar{\phi})/\sigma = 10°/3.6° = 2.8$. Hence, $P[20° \le \phi \le 40°] \ge 1 - (1/2.8)^2 \ge 0.87$. That is, the probability is at least 0.87 that the ϕ-strength parameter will be between 20° and 40°.

If the unknown probability distribution function is symmetrical with respect to its expected value and the expected value is also its maximum value (said to be *unimodal*), it can be shown that (Freeman, 1963)

$$P\left[\left(\bar{x} - h\sigma\right) \le x \le \left(\bar{x} + h\sigma\right)\right] \ge 1 - \frac{4}{9h^2} \tag{35}$$

This is sometimes called *Gauss' inequality.* Some numerical values are given in the third column of Table 4.4.

Example 6: Repeat the previous example if it is assumed that the distribution of the ϕ-value is symmetrical with its maximum at the mean value $(\bar{\phi} = 30°)$.

Solution: For this case, Gauss' inequality asserts

$$P\left[20° \le \phi \le 40°\right] \ge 1 - \frac{4}{9(2.8)^2} \ge 0.94$$

Recognizing symmetry we can also claim $P[\phi \le 20°] = P[\phi \ge 40°] = 0.03$.

Example 7: Find the general expression for the probabilities associated with h-sigma bounds for the exponential distribution, $h \ge 1$.

Solution: From Example 4, we have (with $E[x] = 1/a$, $\sigma[x] = 1/a$)

$$P\left[\left(\bar{x} - h\sigma\right) \le x \le \left(\bar{x} + h\sigma\right)\right] = \int_0^{(h+1)/a} ae^{-ax}\, dx = 1 - e^{-(h+1)}$$

Some numerical values are given in Table 4.4. The normal distribution noted in this table will be developed subsequently.

The results in Table 4.4 indicate that, lacking information concerning a probability distribution beyond its first two moments, from a practical engineering point of view, it may be taken to range within 3-sigma bounds. That is, in Figure 4.7b, $x(a) \approx \bar{x} - 3\sigma[x]$ and $x(b) \approx \bar{x} + 3\sigma[x]$.

For a symmetrical distribution all moments of odd order about the mean (central moments) must be zero. Consequently, any odd-ordered moment may be used as a measure of the degree of *skewness* or *asymmetry* of a probability distribution. The fourth central moment $E[(x - \bar{x})^4]$ provides a measure of the peakedness (called *kurtosis*) of a distribution.

As the units of the third central moment are the cube of the units of the variable, to provide an absolute measure of skewness, Pearson [1894, 1895] proposed that its value be divided by the standard deviation cubed to yield the dimensionless *coefficient of skewness,*

$$\beta(1) = \frac{E\left[\left(x - \bar{x}\right)^3\right]}{\left(\sigma[x]\right)^3} \tag{36}$$

If $\beta(1)$ is positive, the long tail of the distribution is on the right side of the mean: if it is negative the long tail is on the left side (Figure 4.10). Pearson also proposed the dimensionless *coefficient of kurtosis* $\beta(2)$ as a measure of peakedness

$$\beta(2) = \frac{E\left[\left(x - \bar{x}\right)^4\right]}{\left(\sigma[x]\right)^4} \tag{37}$$
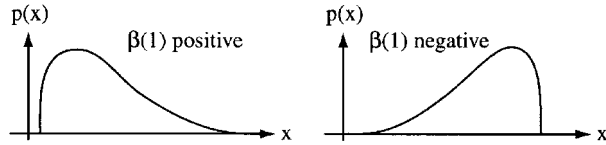
FIGURE 4.10 Coefficient of skewness.

4.3.2 Probability Distributions

In Figure 4.11 are shown the regions occupied by a number of probability distribution types identified by their coefficients of skewness and kurtosis. Examples of the various types are also shown schematically. We note in the figure that the type IV distribution and the symmetrical type VII are unbounded (infinite) below and above. From the point of view of practical groundwater engineering applications, this represents an extremely unlikely distribution. For example, *all* parameters or properties, see Table 4.3, are positive numbers (including zero).
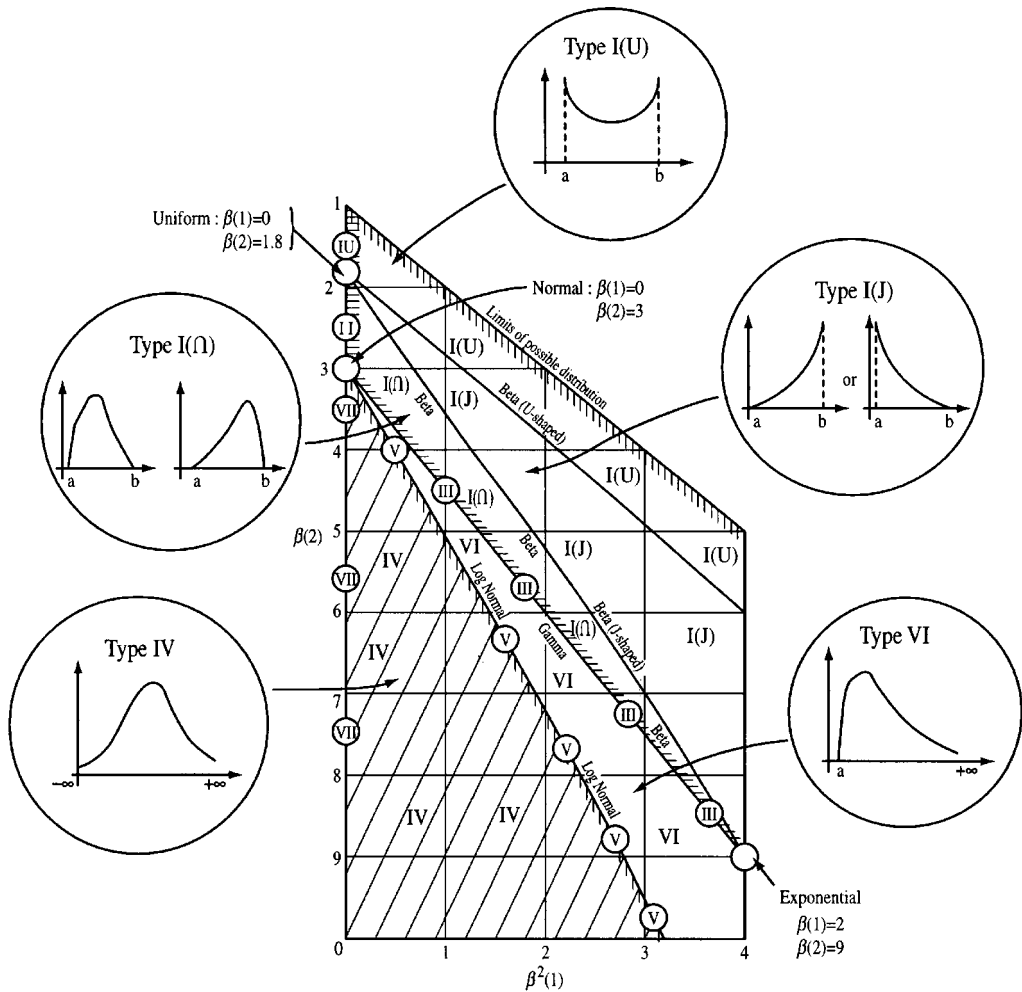
FIGURE 4.11 Space of probability distributions. (Adapted from Pearson, E.S. and Hartley, H.O. 1972. *Biometrika Tables for Statisticians,* vol. II. Cambridge University Press, London.)

The type V (the *lognormal distribution*), type III (the gamma), and type VI distributions are unbounded above. Hence, their use would be confined to those variables with an extremely large range of possible values. Some examples are the coefficient of permeability, the state of stress at various points in a body, distribution of annual rainfall, etc.

The *normal* (Gaussian) *distribution* [$\beta(1) = 0$, $\beta(2) = 3$], even though it occupies only a single point in the universe of possible distributions is the most frequently used probability model. Some associated properties were given in Table 4.4. The normal distribution is the well-known symmetrical bell-shaped curve (Figure 4.12). Some tabular values are given in Table 4.5. The table is entered by forming the standardized variable z, for the normal variate x as

$$z = \left| \frac{x - \bar{x}}{\sigma[x]} \right| \tag{38}$$

Tabular values yield the probabilities associated with the shaded areas shown in the figure: area = $\psi(z)$.
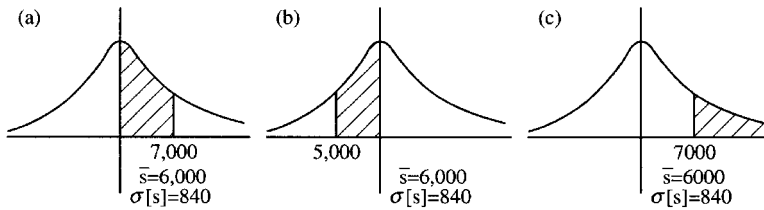
FIGURE 4.12 Example 8.

Example 8: Assuming the strength s of concrete to be used in a weir to be a normal variate with an expected value of $\bar{s} = 6000$ psi and a coefficient of variation of 14%, find (a) $P[6000 \le s \le 7000]$, (b) $P[5000 \le s \le 6000]$, and (c) $P[s \ge 7000]$.

Solution: The standard deviation is $\sigma[s] = (0.14)(6000) = 840$ psi. Hence, (see Figure 4.12),

   (a) $z = |(7000 - 6000)/840| = 1.19$; hence, from Table 4.5 $\psi(1.19) = 0.383$
   (b) By symmetry, $P[5000 \le s \le 6000] = 0.383$
   (c) $P[s \ge 7000] = 0.500 - 0.383 = 0.117$.

As might be expected from its name, the *lognormal distribution* (type V) is related to the normal distribution. If x is a normal variate and $x = \ln y$ or $y = \exp(x)$, then y is said to have a lognormal distribution. It is seen that the distribution has a minimum value of zero and is unbounded above. The probabilities associated with lognormal variates can be obtained very easily from those of mathematically corresponding normal variates (Table 4.5). If $E(y)$ and $V(y)$ are the expected value and coefficient of variation of a lognormal variate, the corresponding normal variate x will have the expected value and standard deviation (Benjamin and Cornell, 1970)

$$\left(\sigma[x] \right)^2 = \ln \left\{ 1 + \left[V(y) \right]^2 \right\} \tag{39a}$$

$$E[x] = \ln E(y) - \left(\sigma[x] \right)^2 \Big/ 2 \tag{39b}$$

Example 9: A live load of 20 tons acts on a footing. If the loading is assumed to be lognormally distributed, estimate the probability that a loading of 40 tons will be exceeded.

TABLE 4.5 Standardized Normal Variates N[0,1]

$$\text{for } z > 2.2,\ \psi(z) \approx \frac{1}{2} - \frac{1}{z}(2\pi)^{-1/2}\exp[-\frac{z^2}{2}] \qquad z = \frac{x-\bar{x}}{\sigma[x]}$$
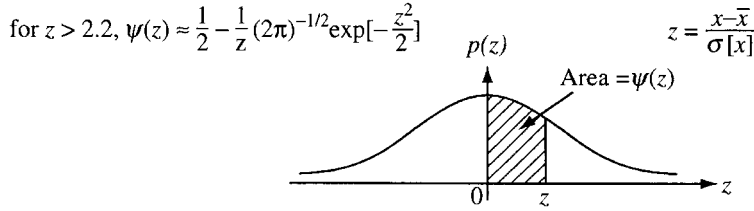
z	0	1	2	3	4	5	6	7	8	9
0	0	.003969	.007978	.011966	.015953	.019939	.023922	.027903	.031881	.035856
.1	.039828	.043795	.047758	.051717	.055670	.059618	.063559	.067495	.071424	.075345
.2	.079260	.083166	.087064	.090954	.094835	.098706	.102568	.106420	.110251	.114092
.3	.117911	.121720	.125516	.129300	.133072	.136831	.140576	.144309	.148027	.151732
.4	.155422	.159097	.162757	.166402	.170031	.174645	.177242	.180822	.184386	.187933
.5	.191462	.194974	.198466	.201944	.205401	.208840	.212260	.215661	.219043	.222405
.6	.225747	.229069	.232371	.235653	.238914	.242154	.245373	.248571	.251748	.254903
.7	.258036	.261148	.264238	.267305	.270350	.273373	.276373	.279350	.282305	.285236
.8	.288145	.291030	.293892	.296731	.299546	.302337	.305105	.307850	.310570	.313267
.9	.315940	.318589	.321214	.323814	.326391	.328944	.331472	.333977	.336457	.338913
1.0	.341345	.343752	.346136	.348495	.350830	.353141	.355428	.357690	.359929	.362143
1.1	.364334	.366500	.368643	.370762	.372857	.374928	.376976	.379000	.281000	.382977
1.2	.384930	.386861	.388768	.390651	.392512	.394350	.396165	.397958	.399727	.401475
1.3	.403200	.404902	.406582	.408241	.409877	.411492	.413085	.414657	.416207	.417736
1.4	.419243	.420730	.422196	.423641	.425066	.426471	.427855	.429219	.430563	.431888
1.5	.433193	.434476	.435745	.436992	.438220	.439429	.440620	.441792	.442947	.444083
1.6	.445201	.446301	.447384	.448449	.449497	.450529	.451543	.452540	.453521	.454486
1.7	.455435	.456367	.457284	.458185	.459070	.459941	.460796	.461636	.462462	.463273
1.8	.464070	.464852	.465620	.466375	.467116	.467843	.468557	.469258	.469946	.470621
1.9	.471283	.471933	.472571	.473197	.473610	.474412	.475002	.475581	.476148	.476705
2.0	.477250	.477784	.478308	.478822	.479325	.479818	.480301	.480774	.481237	.481691
2.1	.482136	.482571	.482997	.483414	.483823	.484222	.484614	.484997	.485371	.485738
2.2	.486097	.486447	.486791	.487126	.487455	.487776	.488089	.488396	.488696	.488989
2.3	.489297	.489556	.489830	.490097	.490358	.490613	.490863	.491106	.491344	.491576
2.4	.491802	.492024	.492240	.492451	.492656	.492857	.493053	.493244	.493431	.493613
2.5	.493790	.493963	.494132	.494297	.494457	.494614	.494766	.494915	.495060	.495201
2.6	.495339	.495473	.495604	.495731	.495855	.495975	.496093	.496207	.496319	.496427
2.7	.496533	.496636	.496736	.496833	.496928	.497020	.497110	.497197	.497282	.497365
2.8	.497445	.497523	.497599	.497673	.497744	.497814	.497882	.497948	.498012	.498074
2.9	.498134	.498293	.498250	.498305	.498359	.498411	.498462	.498511	.498559	.498605
3.0	.498650	.498694	.498736	.498777	.498817	.498856	.498893	.498930	.498965	.498999
3.1	.499032	.499065	.499096	.499126	.499155	.499184	.499211	.499238	.499264	.499289
3.2	.499313	.499336	.499359	.499381	.499402	.499423	.499443	.499462	.499481	.499499
3.3	.499517	.499534	.499550	.499566	.499581	.499596	.499610	.499624	.499638	.499651
3.4	.499663	.499675	.499687	.499698	.499709	.499720	.499730	.499740	.499749	.499758
3.5	.499767	.499776	.499784	.499792	.499800	.499807	.499815	.499822	.499828	.499835
3.6	.499841	.499847	.499853	.499858	.499864	.499869	.499874	.499879	.499883	.499888
3.7	.499892	.499896	.499900	.499904	.499908	.499912	.499915	.499918	.499922	.499925
3.8	.499928	.499931	.499933	.499936	.499938	.499941	.499943	.499946	.499948	.499950
3.9	.499952	.499954	.499956	.499958	.499959	.499961	.499963	.499964	.499966	.499967

Solution: From Table 4.3 we have that the coefficient of variation for a live load, say L, can be estimated as 25%; hence, from Equations (39) we have for the corresponding normal variate, say x,

$$\sigma[x] = \sqrt{\ln\left[1 + (0.25)^2\right]} = 0.25$$

and

$$E[x] = \ln 20 - (0.25)^2 \big/ 2 = 2.96$$

As $x = \ln L$, the value of the normal variate x equivalent to 40 tons is $\ln 40 = 3.69$. We seek the equivalent normal probability $P[3.69 \leq x]$. The standardized normal variate is $z = (3.69 - 2.96)/0.25 = 2.92$. Hence, using Table 4.5:

$$P[40 \leq L] = 0.50 - \psi(2.92) = 0.500 - 0.498 = 0.002$$

Particular notice should be given to the regions of Figure 4.11, labeled the beta distributions. It is seen that many distributions (such as the normal, lognormal, uniform, and exponential) can be obtained as special cases of the very versatile beta distribution. Inherent in selecting a probability distribution to characterize a random variable is the imposition of the limits or range of its applicability. For example, for the normal it is required that the variable range from $-\infty$ to $+\infty$; the range of the lognormal and the exponential is from 0 to $+\infty$. Such assignments may not be critical if knowledge of distributions are desired in the vicinity of their expected values and their coefficients of variation are not excessive (say, less than 25%). On the other hand, many probabilistic assessments (such as reliability and/or the probability of failure) are vested in the tails of distributions. It is in such characterizations that the beta distribution is of great value.

The beta distribution is treated in great detail by Harr (1977, 1987). The latter reference also contains FORTRAN programs for beta probability distributions. This writer modified Pearson's beta distribution by fitting it to the four parameters: the expected value, the standard deviation, the minimum, and the maximum values. Pearson's original work used the coefficients of skewness and kurtosis instead of the minimum and maximum values. Motivating the change was the sensitivity of the distribution to the coefficients. On the other hand, from an engineering point of view, one generally has a feel for the extremes of a design variable. Often zero is the optional minimum value. In the event that limits are not defined, the specification of a range as the mean plus or minus, say, three standard deviations generally places the generated beta distribution well within the accuracy required for most groundwater engineering applications (see Table 4.4).

We next return to groundwater and seepage.

4.4   The Flow Net

The graphical representation of special members of the families of streamlines and corresponding equipotential lines within a flow region form a *flow net*. The orthogonal network shown in Figure 4.13 represents such a system. Although the construction of a flow net often requires tedious trial-and-error adjustments, it is one of the more valuable methods employed to obtain the solutions for two-dimensional flow problems. Of additional importance, even a hastily drawn flow net will often provide a check on the reasonableness of solutions obtained by other means. Noting that, for steady-state conditions, Laplace's equation also models the action of (see Table 4.1) thermal, electrical, acoustical, odoriferous, torsional, and other systems, the flow net is seen to be a significant tool for analysis.

If, in Figure 4.13, Δw denotes the distance between a pair of adjacent streamlines and Δs is the distance between a pair of adjacent equipotential lines in the near vicinity of a point within the region of flow, the approximate velocity (in the mathematical sense) at the point, according to Darcy's law, will be

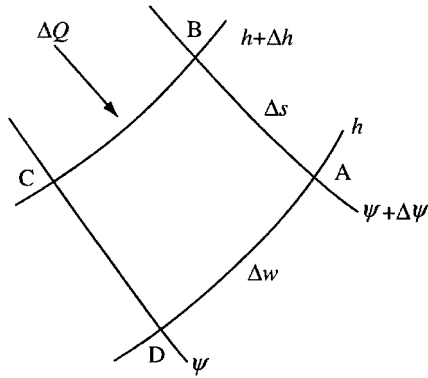$$v \approx \frac{k\,\Delta h}{\Delta s} \approx \frac{\Delta \psi}{\Delta w} \tag{40}$$

FIGURE 4.13 Flow at point.

As the quantity of flow between any two streamlines is a constant, ΔQ, and equal to $\Delta\psi$ (Equation 21) we have

$$\Delta Q \approx k \frac{\Delta w}{\Delta s} \Delta h \tag{41}$$

Equations (40) and (41) are approximate. However, as the distances Δw and Δs become very small, Equation (40) approaches the velocity at a point and Equation (41) yields the quantity of discharge through the flow channel.

In Figure 4.14 is shown the completed flow net for a common type of structure. We first note that there are four boundaries: the bottom impervious contour of the structure *BGHC*, the surface of the impervious layer *EF*, the headwater boundary *AB*, and the tailwater boundary *CD*. The latter two boundaries designate the equipotential lines $h = h$ and $h = 0$, respectively. For steady-state conditions the quantity of discharge through the section Q, and the head loss ($h = 16$ ft) must be constant. If the flow region is saturated, it follows that the two impervious boundaries are not only streamlines but their difference must be identically equal to the discharge quantity

$$Q = \psi_{BGHC} - \psi_{EF}$$

From among the infinite number of possible streamlines between the impervious boundaries, we sketch only a few specifying the same quantity of flow between neighboring streamlines. Designating N_f as the number of flow channels, we have, from above,*

$$Q = N_f \Delta Q = kN_f \frac{\Delta w}{\Delta s} \Delta h$$

Similarly, from among the infinite number of possible equipotential lines between headwater and tailwater boundaries, we sketch only a few and specify the same drop in head, say Δh, between adjacent equipotential lines. If there are N_e equipotential drops along each of the channels,

$$h = N_e \Delta h$$

and

---

* There is little to be gained by retaining the approximately equal sign $\approx$.
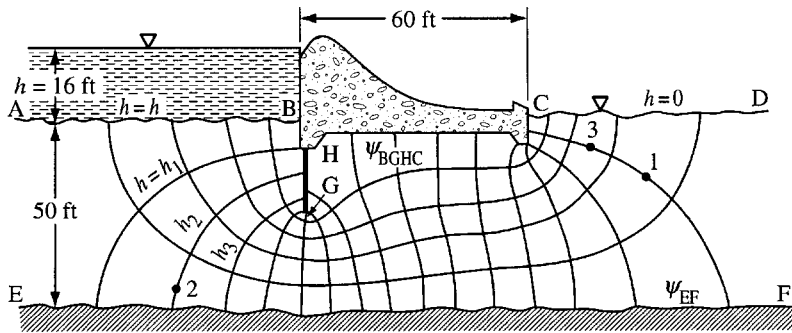
FIGURE 4.14 Example of flow net.

$$Q = k \frac{N_f}{N_e} \frac{\Delta w}{\Delta s} h \tag{42}$$

If, now, we also require that the ratio $\Delta w / \Delta s$ be the same at all points in the flow region, for convenience, and as a square is most sensitive to visual inspection, we take this ratio to be unity,

$$\frac{\Delta w}{\Delta s} = 1$$

and obtain

$$Q = k \frac{N_f}{N_e} h \tag{43}$$

Recalling that Q, k, and h are all constants, Equation (43) demonstrates that the resulting construction, with the obvious requirement that everywhere within the flow domain streamlines and equipotential lines meet at right angles, will yield a unique value for the ratio of the number of flow channels to the number of equipotential drops, N_f/N_e. In Figure 4.14, N_f equals about 5 and N_e equals 16; hence, N_f/N_e = 5/16.

The graphical technique of constructing flow nets by sketching was first suggested by Prasil (1913) although it was developed formally by Forchheimer (1930); however, the adoption of the method by engineers followed Casagrande's classical paper in 1940. In this paper and in the highly recommended flow nets of Cedergren (1967) are to be found some of the best examples of the art of drawing flow nets. Harr (1962) also warrants a peek!

Unfortunately, there is no "royal road" to drawing a good flow net. The speed with which a successful flow net can be drawn is highly contingent on the experience and judgment of the individual. In this regard, the beginner will do well to study the characteristics of well-drawn flow nets: *labor omnia vincit.*

In summary, a flow net is a sketch of distinct and special streamlines and equipotential lines that preserve right-angle intersections within the flow regime, satisfy the boundary conditions, and form curvilinear squares.* The following procedure is recommended:

---

* We except singular squares such as the five-sided square at point H in Figure 4.14 and the three-sided square at point G. (It can be shown [Harr, 1962, p. 84] that a five-sided square designates a point of turbulence.) With continued subdividing into smaller squares, the deviations, in the limit, act only at singular points.

1. Draw the boundaries of the flow region to a scale so that all sketched equipotential lines and streamlines terminate on the figure.
2. Sketch lightly three or four streamlines, keeping in mind that they are only a few of the infinite number of possible curves that must provide a smooth transition between the boundary streamlines.
3. Sketch the equipotential lines, bearing in mind that they must intersect all streamlines, including boundary streamlines, at right angles and that the enclosed figures must form curvilinear rectangles (except at singular points) with the same ratio of $\Delta s/\Delta s$ along a flow channel. Except for partial flow channels and partial head drops, these will form curvilinear squares with $\Delta w = \Delta s$.
4. Adjust the initial streamlines and equipotential lines to meet the required conditions. Remember that the drawing of a good flow net is a trial-and-error process, with the amount of correction being dependent upon the position selected for the initial streamlines.

Example 10: Obtain the quantity of discharge, Q/kh, for the section shown in Figure 4.15.

Solution: This represents a region of horizontal flow with parallel horizontal streamlines between the impervious boundaries and vertical equipotential lines between reservoir boundaries. Hence the flow net will consist of perfect squares, and the ratio of the number of flow channels to the number of drops will be $N_f/N_e = a/L$ and $Q/kh = a/L$.
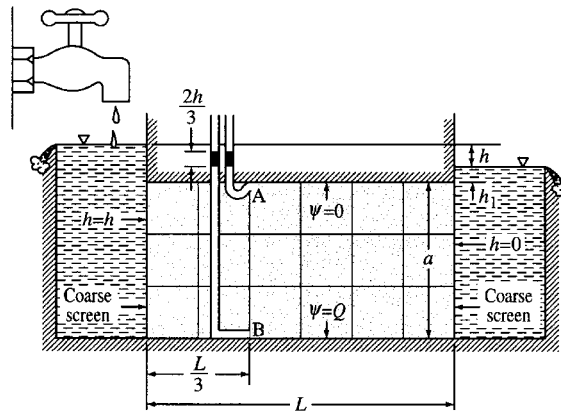

FIGURE 4.15 Example of flow regime.

Example 11: Find the pressure in the water at points A and B in Figure 4.15.

Solution: For the scheme shown in Figure 4.15, the total head loss is linear with distance in the direction of flow. Equipotential lines are seen to be vertical. The total heads at points A and B are both equal to $2h/3$ (datum at the tailwater surface). This means that a piezometer placed at these points would show a column of water rising to an elevation of $2h/3$ above the tailwater elevation. Hence, the pressure at each point is simply the weight of water in the columns above the points in question: $p_A = (2h/3 + h_1)\gamma_w$, $p_B = (2h/3 + h_1 + a)\gamma_w$.

Example 12: Using flow nets, obtain a plot of Q/kh as a function of the ratio s/T for the single impervious sheetpile shown in Figure 4.16a.

Solution: We first note that the section is symmetrical about the y axis, hence only one half of a flow net is required. Values of the ratio s/T range from 0 to 1; with 0 indicating no penetration and 1 complete cut-off. For $s/T = 1$, $Q/kh = 0$ (see point a in Figure 4.16b). As the ratio of s/T decreases, more flow channels must be added to maintain curvilinear squares and, in the limit as s/T approaches zero, Q/kh becomes unbounded (see arrow on Figure 4.16b). If $s/T = 1/2$ (Figure 4.16c), each streamline will evidence
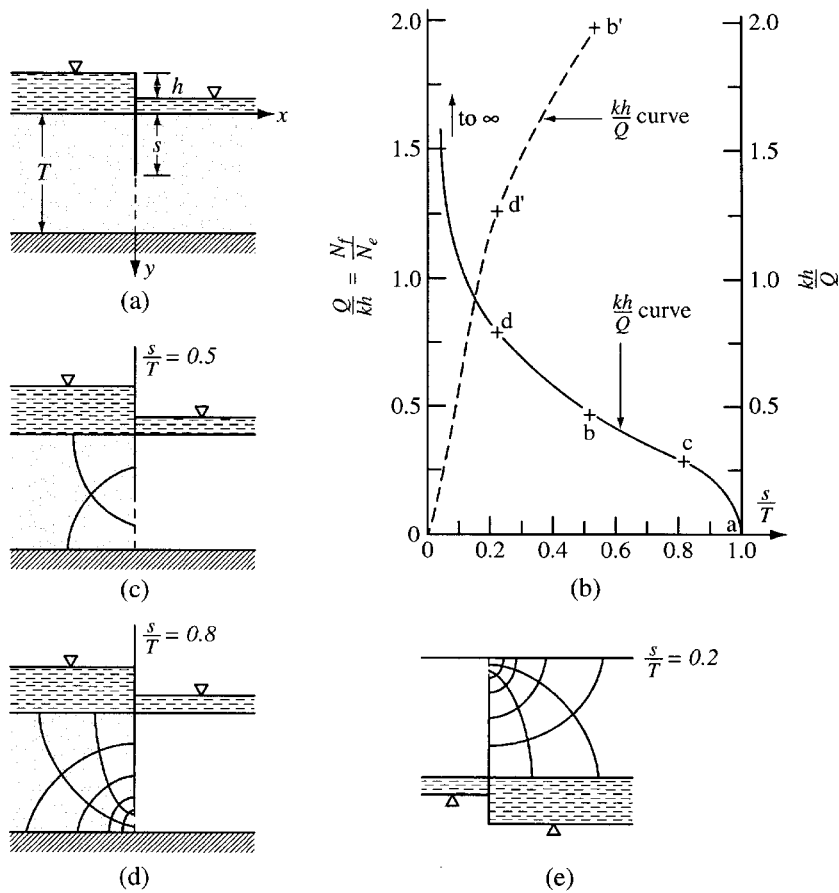
FIGURE 4.16 Example 12.

a corresponding equipotential line in the half-strip; consequently, for the whole flow region, $N_f/N_e = 1/2$ for $s/T = 1/2$ (point b in Figure 4.16b). Thus, without actually drawing a single flow net, we have learned quite a bit about the functional relationship between Q/kh and s/T. If Q/kh was known for $s/T = 0.8$ we would have another point and could sketch, with some reliability, the portion of the plot in Figure 4.16b for $s/T \geq 0.5$. In Figure 4.16d is shown one half of the flow net for $s/T = 0.8$, which yields the ratio of $N_f/N_e = 0.3$ (point c in Figure 4.16b). As shown in Figure 4.16e, the flow net for $s/T = 0.8$ can also serve, geometrically, for the case of $s/T = 0.2$, which yields approximately $N_f/N_e = 0.8$ (plotted as point d in Figure 4.16b). The portion of the curve for values of s/T close to 0 is still in doubt. Noting that for $s/T = 0$, $kh/Q = 0$, we introduce an ordinate scale of kh/Q to the right of Figure 4.16b and locate on this scale the corresponding values for $s/T = 0$, 0.2, and 0.5 (shown primed). Connecting these points (shown dotted) and obtaining the inverse, Q/kh, at desired points, the required curve can be obtained.

A plot giving the quantity of discharge (Q/kh) for symmetrically placed pilings as a function of depth of embedment (s/T), as well as for an impervious structure of width ($2b/T$), is shown in Figure 4.17. This plot was obtained by Polubarinova-Kochina (1952) using a mathematical approach. The curve labeled $b/T = 0$ applies for the conditions in Example 12. It is interesting to note that this whole family of curves can be obtained, with reasonable accuracy,* by sketching only two additional half flow nets (for special values of b/T).

---

* At least commensurate with the determination of the coefficient of permeability, k.

FIGURE 4.17 Quantity of flow for given geometry.
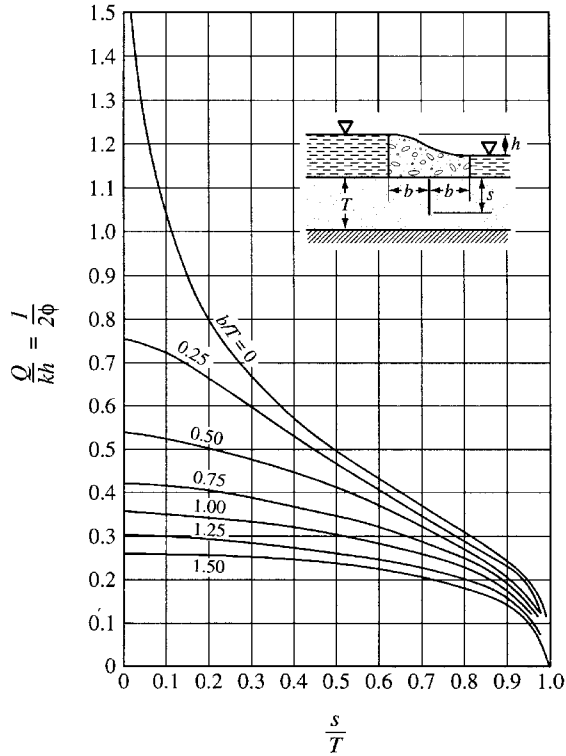
If $m = \cos\dfrac{\pi s}{2T} \sqrt{\tanh^2 \dfrac{\pi b}{2T} + \tan^2 \dfrac{\pi s}{2T}}$

for $m \leq 0.3$, $\dfrac{Q}{kh} = \dfrac{1}{\pi} \ln \dfrac{4}{m}$

for $m^2 \geq 0.9$, $\dfrac{Q}{kh} = \dfrac{-\pi}{2\ln(\frac{1-m^2}{16})}$

It was tacitly assumed in the foregoing that the medium was homogeneous and isotropic ($k_x = k_y$). Had isotropy not been realized, a transformation of scale* in the y direction of $Y = y(k_x/k_y)^{1/2}$ or in the x direction of $X = x(k_y/k_x)^{1/2}$ would render the system as an equivalent isotropic system (for details, see Harr, 1962, p. 29). After the flow net has been established, by applying the inverse of the scaling factor, the solution can be had for the anisotropic system. The quantity of discharge for a homogeneous and anisotropic section is

$$Q = \sqrt{k_x k_y}\, h \frac{N_f}{N_e} \tag{44}$$

We next introduce some powerful probabilistic tools.

---

* Here x and y are taken as the directions of maximum and minimum permeability.

4.5 The Point Estimate Method [PEM]

4.5.1 One Random Variable

Many methods have been offered in the past to accommodate uncertainty. The most common, by far, is to assign single-valued point estimates that reflect central tendencies or implied levels of conservatism. Analyses are then reduced to deterministic treatments. More direct probabilistic methods employ Monte Carlo simulations or truncated Taylor series (Harr, 1987). However, analyses rapidly become exceedingly difficult, if not impossible, for these methods for any but a very few uncorrelated random variables. A simple and very versatile procedure called the *point estimate method* is advocated by this writer and will be developed in some detail. This method, first presented by Rosenblueth (1975) and later extended by him in 1981, has since seen considerable use and expansion by this writer and his co-authors. The methodology is presented in considerable detail in Harr (1987).

Consider $y = p(x)$ to be the probability distribution of the random variable x, with expected value $\bar{x}$ and standard deviation σ. With analogy to Figure 4.7, we replace the load on the beam by *two* reactions, $p(-)$ and $p(+)$ acting at $x(-)$ and $x(+)$, as shown in Figure 4.18.
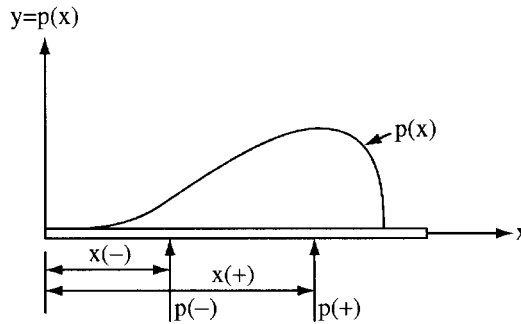
FIGURE 4.18 Point estimate approximations.

Pleading symmetry, probabilistic arguments produce for a random variable *x*:

$$p(+) = p(-) = \frac{1}{2} \tag{45a}$$

$$x(+) = \bar{x} + \sigma[x] \tag{45b}$$

$$x(-) = \bar{x} - \sigma[x] \tag{45c}$$

With the distribution $p(x)$ approximated by the point estimates $p(-)$ and $p(+)$, the moments of $y = p(x)$ are

$$E[y] = \bar{y} = p(-)y(-) + p(+)y(+) \tag{46a}$$

$$E[y^2] = p(-)y^2(-) + p(+)y^2(+) \tag{46b}$$

where $y(-)$ and $y(+)$ are the values of the function $p(x)$ at $x(-)$ and $x(+)$, respectively. These reduce to the simpler expressions

$$\bar{y} = \frac{y(+) + y(-)}{2} \tag{47a}$$

$$\sigma[y] = \left| \frac{y(+) - y(-)}{2} \right| \tag{47b}$$

Example 13: Estimate the expected value and the coefficient of variation for the well-known coefficient of active earth pressure $K_A = \tan^2(45 - \phi/2)$, if $E[\phi] = 30°$.

Solution: With the standard deviation of the ϕ-parameter not given, we again return to Table 4.3, $V(\phi)$ = 12%, and $\sigma[\phi] = 3.6°$. Hence, $\phi(+) = 33.6°$, $\phi(-) = 26.4°$. Therefore, $K_A(+) = 0.29$, $K_A(-) = 0.38$, and Equations (47) produce $\bar{K}_A = 0.34$, $\sigma[K_A] = 0.05$; hence, $V(K_A) = 13\%$.

Before proceding with functions of many random variables, we shall first consider regression and correlation.

4.5.2 Regression and Correlation

We next study the functional relationship between random variables called *regression analysis*. It is regression analysis that provides the grist of being able to predict the value of one variable from that of another or of others. The measure of the degree of correspondence within the developed relationship belongs to *correlation analysis*.

Let us suppose we have N pairs of data $[x(1), y(1)], \ldots, [x(N), y(N)]$ for which we postulate the *linear* relationship

$$y = Mx + B \tag{48}$$

where *M* and *B* are constants. Of the procedures available to estimate these constants (including *best fit by eye*), the most often used is the *method of least squares*. This method is predicated on minimizing the sum of the squares of the distances between the data points and the corresponding points on a straight line. That is, *M* and *B* are chosen so that

$$\Sigma (y - Mx - B)^2 = \text{minimum}$$

This requirement is met by the expressions

$$M = \frac{N \Sigma xy - \Sigma x \Sigma xy}{N \Sigma x^2 - (\Sigma x)^2} \tag{49a}$$

$$B = \frac{\Sigma x^2 \Sigma y - \Sigma x \Sigma xy}{N \Sigma x^2 - (\Sigma x)^2} \tag{49b}$$

It should be emphasized that a straight line fit was assumed. The reasonableness of this assumption is provided by the *correlation coefficient* ρ, defined as

$$\rho = \frac{\text{cov}[x, y]}{\sigma[x] \sigma[y]} \tag{50}$$

where $\sigma[x]$ and $\sigma[y]$ are the respective standard deviations and cov$[x,y]$ is their *covariance*. The covariance is defined as

$$\text{cov}[x, y] = \frac{1}{N} \sum_{i=1}^{N} [x(i) - \bar{x}][y(i) - \bar{y}] \tag{51}$$

With analogy to statics the covariance corresponds to the *product of inertia*.

In concept, the correlation coefficient is a measure of the tendency for two variables to vary together. This measure may be zero, negative, or positive; wherein the variables are said to be *uncorrelated, negatively correlated,* or *positively correlated.*

The variance is a special case of the covariance as

$$\text{cov}[x, x] = v[x] \tag{52}$$

Application of their definitions produces (Ditlevsen, 1981) the following identities (*a, b, c* are constants):

$$E[a + bx + cy] = a + bE[x] + cE[y] \tag{53a}$$

$$v[a + bx + cy] = b^2 v[x] + c^2 v[y] + 2bc\,\text{cov}[x, y] \tag{53b}$$

$$\text{cov}[x, y] \leq \sigma[x]\sigma[y] \tag{53c}$$

$$v[a + bx + cy] = b^2 v[x] + c^2 v[y] + 2bc\,\sigma[x]\sigma[y]\rho \tag{53d}$$

Equation (53c) demonstrates that the correlation coefficient, Equation (50), must satisfy the condition

$$-1 \leq \rho \leq +1 \tag{54}$$

If there is perfect correlation between variables and in the same sense of direction, $\rho = +1$. If there is perfect correlation in opposite directions (one variable increases as the other decreases), $\rho = -1$. If some scatter exists, $-1 < \rho < +1$, with $\rho = 0$ if there is no correlation. Some examples are shown in Figure 4.19

We shall next present an extension of the above correlation methodology that will accommodate very many random variables.

Suppose that 12 tests are performed and sample measurements of the three variables x_1, x_2, and x_3 yield the values in columns so labeled in Table 4.6. Also shown are the sample expected values $(\bar{x}_1)$ obtained as
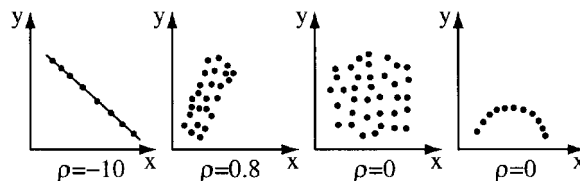

FIGURE 4.19 Example of scatter and correlation coefficients.

TABLE 4.6 Example of Extraction of Moments and Correlation Coefficients

Test	x_1	x_2	x_3	dx_1	dx_2	dx_3	ds_1	ds_2	ds_3
1	2	1	2	−5.25	−6.67	−4.92	−1.92	−1.85	−1.29
2	4	3	1	−3.25	−4.67	−5.92	−1.18	−1.30	−1.55
3	6	7	5	−1.25	−0.67	−1.92	−0.46	−0.19	−0.50
4	7	6	3	−0.25	−1.67	−3.92	−0.09	−0.46	−1.02
5	5	4	7	−2.25	−3.67	0.08	−0.82	−1.02	0.02
6	8	9	6	0.75	1.33	−0.92	0.27	0.37	−0.25
7	9	8	7	1.75	0.33	0.08	0.64	0.09	0.02
8	7	10	9	−0.25	2.33	2.08	−0.09	0.65	0.55
9	10	11	11	2.75	3.33	4.08	1.01	0.93	1.07
10	8	9	8	0.75	1.33	1.08	0.27	0.37	0.28
11	12	11	10	4.75	3.33	3.08	1.74	0.93	0.81
12	9	13	14	1.75	5.33	7.08	0.64	1.48	1.85
Expected value	7.25	7.67	6.92	0	0	0	0	0	0
Standard deviations	2.73	3.60	3.82	2.73	3.60	3.82	1	1	1

$$\bar{x}_i = E[x_i] = \frac{1}{12}\sum_{i=1}^{12} x_i, \qquad i = 1,2,3 \tag{55}$$

Columns labeled dx_1, dx_2, and dx_3 are *mean-corrected*; that is, the respective expected values have been subtracted from the raw data,

$$dx_i = x_i - \bar{x}_i, \qquad i = 1,2,3 \tag{56}$$

Designating the 12×3 matrix composed of the three columns (dx_1, dx_2, dx_3) as *D*, the *mean-corrected matrix*, and performing the matrix multiplication, *D′* is the transpose of the matrix *D*, produces the symmetrical *covariance matrix*, *C*:

$$C = \frac{1}{11} D'D = \begin{bmatrix} 7.48 & 8.73 & 7.84 \\ 8.73 & 12.97 & 12.24 \\ 7.84 & 12.24 & 14.63 \end{bmatrix} \tag{57a}$$

The elements on the principal diagonal, 7.48, 12.97, 14.63, are the respective variances of the variables x_1, x_2, and x_3. Hence, the respective standard deviations are 2.73, 3.60, and 3.82.

$$\sigma_1 = 2.73, \qquad \sigma_2 = 3.60, \qquad \sigma_3 = 3.82 \tag{57b}$$

The off-diagonal elements are the respective covariances; for example, $8.73 = \sigma_1\sigma_2\rho_{1,2}$, where $\rho_{1,2}$ is the correlation coefficient between the variables x_1 and x_2. Hence, $\rho_{1,2} = 8.73/(2.73 \times 3.60) = 0.89$. In general, dividing a row and column of the covariance matrix by the corresponding standard deviations produces the symmetrical *correlation matrix*, *K*, which for the example at hand becomes

$$K = \begin{bmatrix} 1.00 & 0.89 & 0.75 \\ 0.89 & 1.00 & 0.89 \\ 0.75 & 0.89 & 1.00 \end{bmatrix} \tag{58}$$

It should be noted that the correlation matrix could also have been obtained directly from the raw data, the columns of x_1, x_2, and x_3 in Table 4.6. The respective standard deviations (σ_i) are given by (Note: $12 - 1 = 11$):

$$\sigma_i^2 = \frac{1}{11} \sum_{i=1}^{12} \left(x_i - \bar{x} \right)^2, \quad i = 1, 2, 3 \tag{59}$$

The last three columns of Table 4.6, labeled ds_1, ds_2, and ds_3 are *standardized variates*, obtained by subtracting the expected values and dividing by the standard deviations, respectively; that is

$$ds_i = \frac{x_i - \bar{x}_i}{\sigma_i} \tag{60}$$

Then, forming the 12×3 matrix D_s, with the columns the same as the last three columns of Table 4.6, the correlation matrix is obtained as

$$K = \frac{1}{11} D_s' D_s \tag{61}$$

Thus, the data in Table 4.6 have been reduced to provide the expected values, standard deviations, and correlation coefficients of the three uncertain variables. It should be apparent that the foregoing concepts can be readily extended to m test results for n variables to produce the $n \times n$ correlation matrix and the respective expected values and standard deviations of the n-variates. The coefficient in Equations (59) and (61) would be $1/(m - 1)$. Computer software abounds that readily produces these results.

4.5.3   Point Estimate Method — Several Random Variables

Rosenblueth (1975) generalized his earlier methodology to any number of correlated variables. For example, for a function of three random variables; say, $y = y[x(1), x(2), x(3)]$; where $\rho(i,j)$ is the correlation coefficient between variables $x(i)$ and $x(j)$ he obtained:

$$E\left[y^N \right] = p(+++)y^N(+++) + p(++-)y^N(++-) + \ldots + p(---)y^N(---) \tag{62a}$$

where

$$y(\pm\pm\pm) = y\left[\bar{x}(1) \pm \sigma[x1], \bar{x}(2) \pm \sigma[x2], \bar{x}(3) \pm \sigma[x3] \right] \tag{62b}$$

$$p(+++) = p(---) = \frac{1}{2^3}\left[1 + \rho(1,2) + \rho(2,3) + \rho(3,1) \right]$$

$$p(++-) = p(--+) = \frac{1}{2^3}\left[1 + \rho(1,2) - \rho(2,3) - \rho(3,1) \right]$$

$$p(+-+) = p(-+-) = \frac{1}{2^3}\left[1 - \rho(1,2) - \rho(2,3) + \rho(3,1) \right]$$

$$p(+--) = p(-++) = \frac{1}{2^3}\left[1 - \rho(1,2) + \rho(2,3) - \rho(3,1)\right] \tag{62c}$$

where $\sigma[xi]$ is the standard deviation of $x(i)$. The sign of $\rho(i,j)$ is determined by the multiplication rule of i and j; that is, if the sign of $i = (-)$, and of $j = (+)$; $(i)(j) = (-)(+) = (-)$.

Equation (62a) has $2^3 = 8$ terms, all permutations of the three +'s and −'s. In general, for M variables there are 2^M terms and $M(M-1)/2$ correlation coefficients, the number of combinations of M objects taken two at a time. The coefficient on the right-hand side of Equation (62c), in general, is $(1/2)^M$.

Example 14: The recommendation of the American Concrete Institute (Galambos et al., 1982) for the design of reinforced concrete structures is (in simplified form)

$$R \geq 1.6D + 1.9L$$

where R is the strength of the element, D is the dead load, and L is the lifetime live load. If $E[D] = 10$, $E[L] = 8$, $V(D) = 10\%$, $V(L) = 25\%$, $\rho(D,L) = 0.75$: (a) Find the expected value and standard deviation of R for the case $R = 1.6I + 1.9L$. (b) If the results in part (a) generate a normal variate and the maximum strength of the element R is estimated to be 40, estimate the implied probability of failure.

Solution: The solution is developed in Figure 4.20.

Generalizations of the point estimate method to more than 3 random variables is given by Harr (1987). We next return to groundwater and seepage.

4.6 Method of Fragments

In spite of its many uses, a graphical flow net provides the solution for a singular problem configuration only. Should one wish to investigate the influence of a range of characteristic dimensions (such as is generally the case in a design problem) many flow nets would be required. Consider, for example, the section shown in Figure 4.21, and suppose we wish to investigate the influence of the dimensions A, B, and C on the characteristics of flow, all other dimensions being fixed. Taking only three values for each of these dimensions would require 27 individual flow nets. As noted previously, a rough flow net should always be drawn as a check. In this respect it may be thought of as being analogous to a free-body diagram in mechanics, wherein the physics of a solution can be examined with respect to satisfying conditions of necessity.

An approximate analytical method of solution, directly applicable to design, was developed by Pavlovsky in 1935 (Pavlovsky, 1956) and was expanded and advanced by Harr (1962, 1977). The underlying assumption of the method, called the *method of fragments*, is that equipotential lines at various critical parts of the flow region can be approximated by straight vertical lines (as, for example, the dotted lines in Figure 4.22) that divide the region into sections or fragments. The groundwater flow region in the figure is shown divided into four fragments.

Suppose, now, that one computes the discharge in the i*th* fragment of a structure with m such fragments as

$$Q = \frac{kh_i}{\Phi_i} \tag{63}$$

where h_i = head loss through the i*th* fragment, Φ_i = dimensional form factor in ith fragment, $\Phi_i = N_e/N_f$ in Equation (43). Then, as the discharge through all fragments must be the same,

a) $R = 1.6D + 1.9L$

Variable, x	$\bar{x}$	$\sigma[x]$	$x(+)$	$x(-)$
D	10	1	11	9
L	8	2	10	6

$$\rho(D,L) = +0.75$$

	$R(ij)$	$R(ij)^2$
$R(++)$:	36.6	1340
$R(+ -)$:	29.0	841
$R(- +)$:	33.4	1116
$R(- -)$:	25.8	666

$p(++) = 1/4\,(1+\rho) = 0.44$

$p(+-) = 1/4\,(1-\rho) = 0.06$

$p(-+) = 1/4\,(1-\rho) = 0.06$

$p(--) = 1/4\,(1+\rho) = 0.44$

$E[R]=\bar{R}=\Sigma R(ij)p(ij)$
     $=0.44(36.6+25.8)+0.06(29.0+33.4)$
     $=\underline{31.20}$

$E[R^2]=\Sigma R(ij)^2\,p(ij)$
     $=0.44(1340+666)+0.06(841+1116)$
     $=\underline{1000.06}$

$v[R]=E[R^2]-(E[R])^2 = 1000.06 - (31.20)^2 = \underline{26.62}$

$\sigma[R]=5.16; \quad V(R) = 16.5\%$

$$P_f=P[R\geq40] =\frac{1}{2} -\psi\left[\frac{40-31.20}{5.16}\right] = \frac{1}{2} -\psi[1.71] = \underline{0.044} \ (Table\,3)$$

The exact solution is $\underline{0.043}$
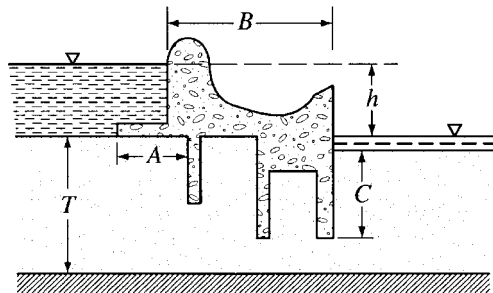
FIGURE 4.20 Solution to Example 14.

FIGURE 4.21 Example of complicated structure.

$$Q= \frac{kh_1}{\Phi_1} = \frac{kh_2}{\Phi_2} = \frac{kh_i}{\Phi_i} =\cdots = \frac{kh_m}{\Phi_m}$$
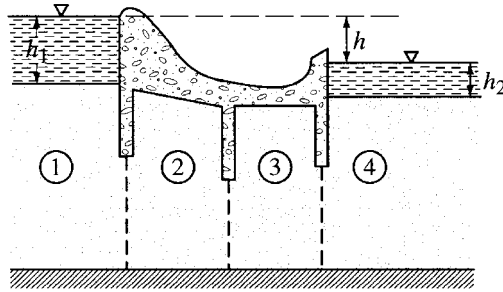
whence

$$\sum h_i = \frac{Q}{k}\sum \Phi_i$$

FIGURE 4.22 Four fragments.

and

$$Q = k \frac{h}{\sum_{i=1}^{m} \Phi_i} \tag{64}$$

where h (without subscript) is the total head loss through the section. By similar reasoning, the head loss in the i*th* fragment can be calculated from

$$h_i = \frac{h\Phi_i}{\sum \Phi} \tag{65}$$

Thus, the primary task is to implement this method by establishing a catalog of form factors. Following Pavlovsky, the various form factors will be divided into types. The results are summarized in tabular form, Figure 4.23, for easy reference. The derivation of the form factors is well documented in the literature (Harr 1962, 1977).

Various entrance and emergence conditions for Type VIII and IX fragments are shown in Figure 4.24. Briefly, for the entrance condition, when possible, the free surface will intersect the slope at right angles. However, as the elevation of the free surface represents the level of available energy along the uppermost streamline, at no point along the curve can it rise above the level of its source of energy, the headwater elevation. At the point of emergence the free surface will, if possible, exit tangent to the slope (Dachler, 1934). As the equipotential lines are assumed to be vertical, there can be only a single value of the total head along a vertical line, and, hence, the free surface cannot curve back on itself. Thus, where unable to exit tangent to a slope it will emerge vertical.

To determine the pressure distribution on the base of a structure (such as that along C′CC″ in Figure 4.25), Pavlovsky assumed that the head loss within the fragment is linearly distributed along the impervious boundary. Thus, if h_m is the head loss within the fragment, the rate of loss along E′C′CC″E″ will be

$$R = \frac{h_m}{L + s' + s''} \tag{66}$$

Once the total head is known at any point, the pressure can easily be determined by subtracting the elevation head, relative to the established (tailwater) datum.

Example 15: For the section shown in Figure 4.26a, estimate (a) the discharge and (b) the uplift pressure on the base of the structure.

Fragment type	Illustration	Form factor, Φ (h is head loss through fragment)	Fragment type	Illustration	Form factor, Φ (h is head loss through fragment)
I		$\Phi = \dfrac{L}{a}$	V		$L \leq 2s:$ $\Phi = 2 \ln\left(1+\dfrac{L}{2a}\right)$ $L \geq 2s:$ $\Phi = 2 \ln\left(1+\dfrac{s}{a}\right) + \dfrac{L-2s}{T}$
II		$\Phi = \dfrac{1}{2}\left(\dfrac{kh}{Q}\right)$, Fig. 4.17	VI		$L \geq 2s'+s'':$ $\Phi = \ln\left[\left(1+\dfrac{s'}{a'}\right)\left(1+\dfrac{s''}{a''}\right)\right]$ $+ \dfrac{L-(s'+s'')}{T}$ $L \leq 2s'+s'':$ $\Phi = \ln\left[\left(1+\dfrac{b'}{a'}\right)\left(1+\dfrac{b''}{a''}\right)\right]$ where $b' = \dfrac{L+(s'-s'')}{2}$ $b'' = \dfrac{L-(s'-s'')}{2}$
III		$\Phi = \dfrac{1}{2}\left(\dfrac{kh}{Q}\right)$, Fig. 4.17	VII		$\Phi = \dfrac{2L}{h_1+h_2}$ $Q = k\dfrac{h_1^2+h_2^2}{2L}$
IV		$b \leq s:$ $\Phi = \ln\left(1+\dfrac{b}{a}\right)$ $b \geq s:$ $\Phi = \ln\left(1+\dfrac{s}{a}\right)$ $+ \dfrac{b-s}{T}$	VIII		$Q = k\dfrac{h_1-h}{\cot \alpha}\ \ln\dfrac{h_d}{h_d-h}$
			IX		$Q = k\dfrac{a_2}{\cot \beta}\left(1+\ln\dfrac{a_2+h_2}{a_2}\right)$

FIGURE 4.23 Summary of fragment types and form factors.
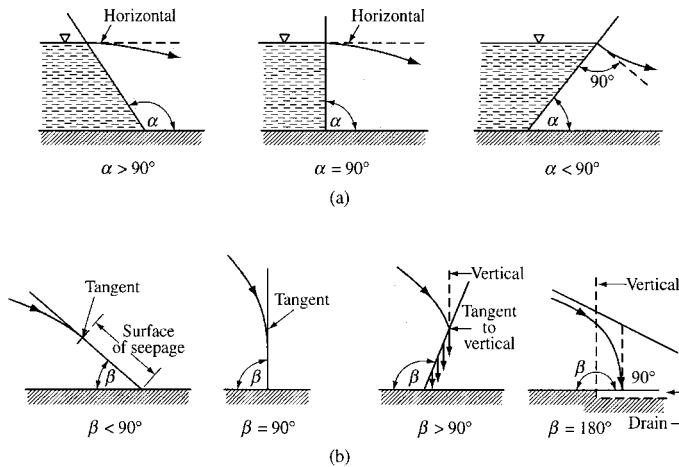
FIGURE 4.24 Entrance and emergence conditions.

Solution: The division of fragments is shown on the figure. Regions 1 and 3 are both type II fragments, and the middle section is type V with L = 2s. For regions 1 and 3, we have, from Figure 4.17, with b/T = 0, $\Phi_1 = \Phi_3 = 0.78$.

For region 2, as L = 2s, $\Phi_2 = 2 \ln(1 + 18/36) = 0.81$. Thus, the sum of the form factors is

$$\sum \Phi = 0.78 + 0.81 + 0.78 = 2.37$$

and the quantity of flow (Equation 64) is $Q/k = 18/2.37 = 7.6$ ft.

FIGURE 4.25 Illustration of Equation (66).
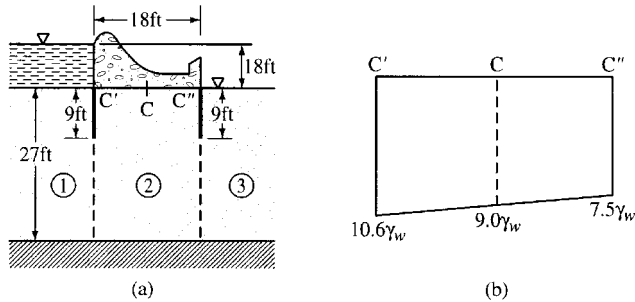
FIGURE 4.26 Example 15.

For the head loss in each of the sections, from Equation (64) we find

$$h_1 = h_3 = \frac{0.78}{2.37}(18) = 5.9 \, \text{ft}$$

$$h_2 = 6.1 \, \text{ft}$$

Hence the head loss rate in region 2 is (Equation 65)

$$R = \frac{6.1}{36} = 17 \text{ percent}$$

and the pressure distribution along C′CC″ is shown in Figure 4.26b.

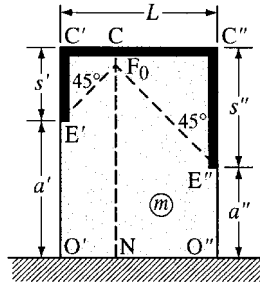**Example 16**[*]: Estimate the quantity of discharge per foot of structure and the point where the free surface begins under the structure (point A) for the section shown in Figure 4.27a.

Solution: The vertical line AC in Figure 4.27a is taken as the vertical equipotential line that separates the flow domain into two fragments. Region 1 is a fragment of type III, with the distance B as an unknown quantity. Region 2 is a fragment of type VII with $L = 25 - B$, and $h_1 = 10$ ft and $h_2 = 0$. Thus, we are led to a trial-and-error procedure to find B. In Figure 4.27b are shown plots of Q/k versus B/T for both regions. The common point is seen to be $B = 14$ ft, which yields a quantity of flow of approximately $Q = 100k/22 = 4.5k$.

---

[*] For comparisons between analytical and experimental results for mixed fragments (confined and unconfined flow) see Harr and Lewis (1965).
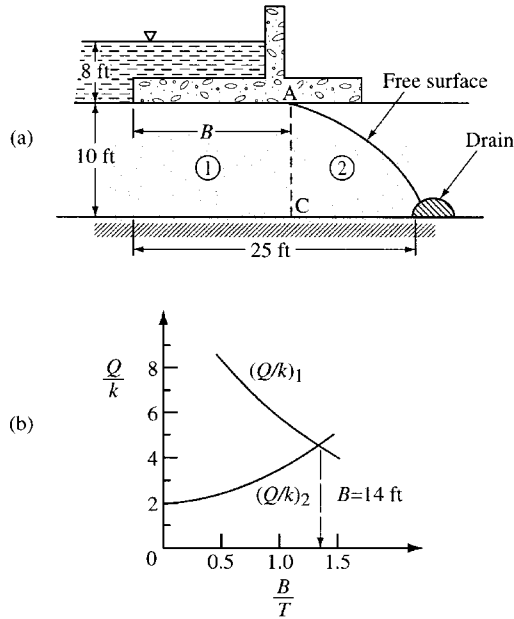
FIGURE 4.27 Example 16.
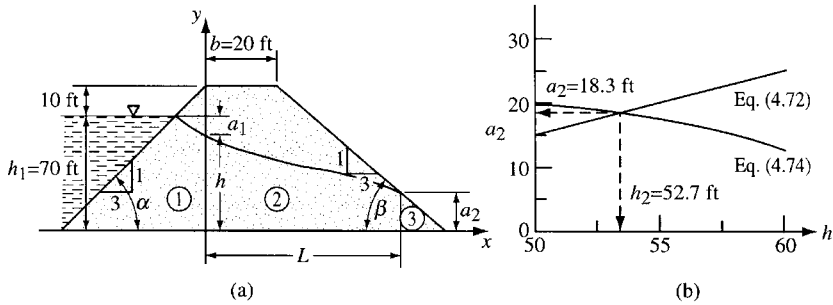
FIGURE 4.28 Example 17.

Example 17: Determine the quantity of flow for 100 ft of the earth dam section shown in Figure 4.28a, where k = 0.002 ft/min.

Solution: For this case, there are three regions. For region 1, a type VIII fragment $h_1 = 70$ ft, cot $\alpha = 3$, $h_d = 80$ ft, produces

$$\frac{Q}{k} = \frac{70 - h}{3} \ln \frac{80}{80 - h} \tag{67}$$

For region 2, a type VII fragment.

$$\frac{Q}{k} = \frac{h^2 - a_2^2}{2L} \tag{68}$$

With tailwater absent, $h_2 = 0$, the flow in region 3, a type IX fragment, with cot $\beta = 3$ produces

©1999 CRC Press LLC

$$\frac{Q}{k} = \frac{a_2}{3} \tag{69}$$

Finally, from the geometry of the section, we have

$$L = 20 + \cot \beta \left[h_d - a_2 \right] = 20 + 3 \left[80 - a_2 \right] \tag{70}$$

The four independent equations contain only the four unknowns, h, a_2, Q/k, and L, and hence provide a complete, if not explicit solution.

Combining Equations (68) and (69) and substituting for L in (70), we obtain, in general (b = crest width),

$$a_2 = \frac{b}{\cot \beta} + h_d - \sqrt{\left(\frac{b}{\cot \beta} + h_d \right)^2 - h^2} \tag{71}$$

and, in particular,

$$a_2 = \frac{20}{3} + 80 - \sqrt{\left(\frac{20}{3} + 80 \right)^2 - h^2} \tag{72}$$

Likewise, from Equations (66) and (68), in general,

$$\frac{a_2 \cot \alpha}{\cot \beta} = (h_1 - h) \ln \frac{h_d}{h_d - h} \tag{73}$$

and, in particular,

$$a_2 = (70 - h) \ln \frac{80}{80 - h} \tag{74}$$

Now, Equations (72) and (74), and (71) and (73) in general, contain only two unknowns (a_2 and h), and hence can be solved without difficulty. For selected values of h, resulting values of a_2 are plotted for Equations (72) and (74) in Figure 4.28b. Thus, $a_2 = 18.3$ ft, $h = 52.7$ ft, and $L = 205.1$ ft. From Equation (69), the quantity of flow per 100 ft is

$$Q = 100 \times 2 \times 10^{-3} \times \frac{18.3}{3} = 1.22 \, \text{ft}^3 / \text{min} = 9.1 \, \text{gal/min}$$

4.7   Flow in Layered Systems

Closed-form solutions for the flow characteristics of even simple structures founded in layered media offer considerable mathematical difficulty. Polubarinova-Kochina (1962) obtained closed-form solutions for the two layered sections shown in Figure 4.29 (with $d_1 = d_2$). In her solution she found a cluster of parameters that suggested to Harr (1962) an approximate procedure whereby the flow characteristics of structures founded in layered systems can be obtained simply and with a great degree of reliability.
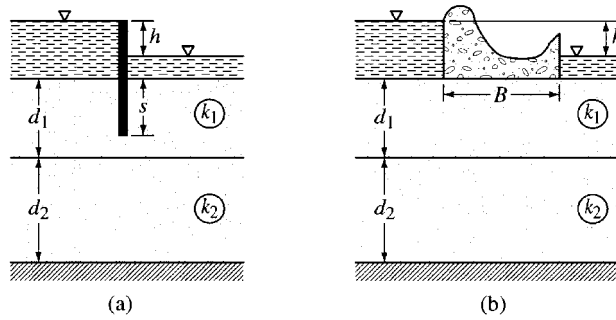
FIGURE 4.29 Example of two-layered systems.

The flow medium in Figure 4.29a consists of two horizontal layers of thickness d_1 and d_2, underlain by an impervious base. The coefficient of permeability of the upper layer is k_1 and of the lower layer, k_2. The coefficients of permeability are related to a dimensionless parameter ε by the expression

$$\tan \pi \varepsilon = \sqrt{\frac{k_2}{k_1}} \qquad (75)^*$$

Thus, as the ratio of the permeabilities varies from 0 to ∞, ε ranges between 0 and 1/2. We first investigate the structures shown in Figure 4.29a for some special values of ε.

1. $\varepsilon = 0$. If $k_2 = 0$, from Equation (75) we have $\varepsilon = 0$. This is equivalent to having the impervious base at depth d_1. Hence, for this case the flow region is reduced to that of a single homogeneous layer for which the discharge can be obtained directly from Figure 4.17.
2. $\varepsilon = 1/4$. If $k_1 = k_2$, the sections are reduced to a single homogeneous layer of thickness $d_1 + d_2$, for which Figure 4.17 is again applicable.
3. $\varepsilon = 1/2$. If $k_2 = \infty$, $\varepsilon = 1/2$. This represents a condition where there is no resistance to flow in the bottom layer. Hence, the discharge through the total section under steady-state conditions is infinite, or $Q/k_1 h = \infty$. However, of greater significance is the fact that the inverse of this ratio $k_1 h/Q = 0$. It can be shown (Polubarinova-Kochina, 1962), that for $k_2/k_1 \to \infty$,

$$\frac{Q}{k_1 h} \approx \sqrt{\frac{k_2}{k_1}} \qquad (76)$$

Thus, we see that for the special values of $\varepsilon = 0$, $\varepsilon = 1/4$, and $\varepsilon = 1/2$ measures of the flow quantities can be easily obtained. The essence of the method then is to plot these values, on a plot of $k_1 h/Q$ versus ε, and to connect the points with a smooth curve, from which intermediate values can be had.

Example 18: In Figure 4.29a, s = 10 ft, d_1 = 15 ft, d_2 = 20 ft, $k_1 = 4k_2 = 1 \times 10^{-3}$ ft/min, h = 6 ft. Estimate $Q/k_1 h$.

Solution: For $\varepsilon = 0$, from Figure 4.17 with $s/T = s/d_1 = 2/3$, $b/T = 0$, $Q/k_1 h = 0.39$, $k_1 h/Q = 2.56$.

For $\varepsilon = 1/4$, from Figure 4.17 with $s/T = s/(d_1 + d_2) = 2/7$, $Q/k_1 h = 0.67$, $k_1 h/Q = 1.49$.
For $\varepsilon = 1/2$, $k_1 h/Q = 0$.
The three points are plotted in Figure 4.30, and the required discharge, for $\varepsilon = 1/\Pi \tan^{-1}(1/4)^{1/2} = 0.15$, is $k_1 h/Q = 1.92$ and $Q/k_1 h = 0.52$; whence
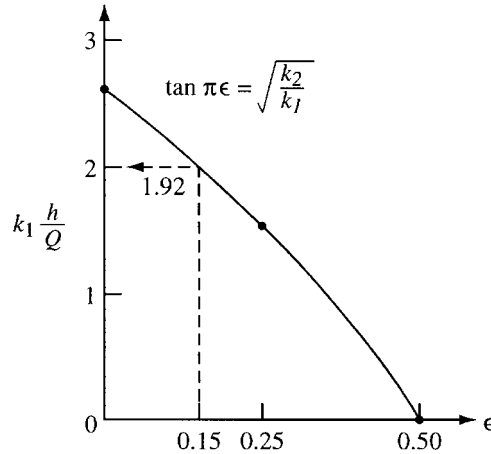
---

* π is in radian measure.

FIGURE 4.30 Example 18.

$$Q = 0.52 \times 1 \times 10^{-3} \times 6 = 3.1 \times 10^{-3} \text{ ft}^3/(\text{min})(\text{ft})$$

In combination with the method of fragments, approximate solutions can be had for very complicated structures.

Example 19: Estimate (a) the discharge through the section shown in Figure 4.31a, $k_2 = 4k_1 = 1 \times 10^{-3}$ ft/day and (b) the pressure in the water at point P.

Solution: The flow region is shown divided into four fragments. However, the form factors for regions 1 and 4 are the same. In Figure 4.31b are given the resulting form factors for the listed conditions. In Figure 4.31c is given the plot of $k_1 h/Q$ versus ε. For the required condition ($k_2 = 4k_1$), $\varepsilon = 0.35$, $k_1 h/Q = 1.6$ and hence $Q = (1/1.6) \times 0.25 \times 10^{-3} \times 10 = 1.6 \times 10^{-3}$ ft^3/(day)(ft).

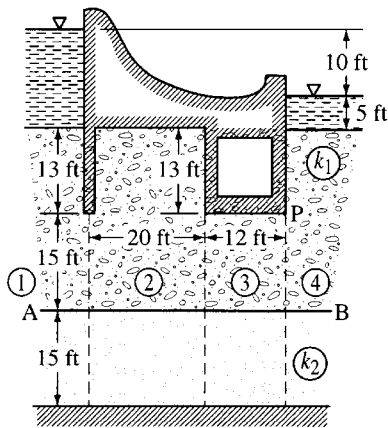The total head at point P is given in Figure 4.31b as Δh for region 4; for $\varepsilon = 0$, $h_p = 2.57$ ft and for $\varepsilon = 1/4$, $h^p = 3.00$. We require h_p for $k_2 = \infty$; theoretically, there is assumed to be no resistance to the flow in the bottom layer for this condition. Hence the boundary between the two layers (AB) is an equipotential line with an expected value of $h/2$. Thus $h_p = 10/2 = 5$ for $\varepsilon = 1/2$. In Figure 4.31d is given the plot of h_p versus ε. For $\varepsilon = 0.35$, $h_p = 2.75$ ft, and the pressure in the water at point P is $(3.75 + 5)\gamma_w = 8.75\gamma_w$.

The above procedure may be extended to systems with more than two layers.

Example 20: Given the weir section shown in Figure 4.32a. There are four random variables: the soil permeability, the thickness of the pervious layer, and the effective depth of both sheetpiles. The coefficient of variation for the soil permeability will be taken as 100%, with a mean value of 1×10^{-3} cm/sec. A coefficient of variation of 30% will be used for the sand layer thickness. A coefficient of variation of 15% will be assumed for the depth of the sheetpiles. Determine the expected value and standard deviation for the quantity of flow per unit of the section shown.

Solution: AB and CD in Figure 4.32a are taken as equipotential lines; hence, the flow regime is divided into three regions or fragments. In general, with four variables, 16 permutations would be required; however, as the form factors are not functions of soil permeability, only eight form factor terms are necessary for this example.

Region 1 is a type III fragment. This fragment contains both an impervious partial cutoff; say (s_1) and top blanket. The top blanket width of 10 feet is considered to be known in this example (no uncertainty). The form factor is a function of the ratios of b and sl to the pervious layer thickness (T). These ratios are entered in Figure 4.17 to determine the term $1/2\phi$ which is then used to calculate ϕ. In the PEM the

FIGURE 4.31 Example 19.

variability in sl and T are represented by two plus and minus terms. There are two b/T ratio terms (b/T_+, and b/T_-) and four s/T ratio terms (s_+/T_+, s_-/T_+, s_+/T_-, s_-/T_-) used to determine the four form factors. The corresponding PEM values for Region 1 are shown in Figure 4.32b.

Region 2 is a type VI fragment for which there are three variables. The corresponding PEM plus and minus terms are shown in Figure 4.32c. For this fragment type the equation for the form factor is dependent on the relationship between the length (L) and the sum of the partial cutoffs ($s1 + s2$). For the present example the length is greater than the sum of the cutoffs. Thus, the following equation is used (Figure 4.23):

A

$\triangledown$

20 ft

25 ft

5 ft

10 ft

10 ft

A

10 ft

A

S_1

13 ft   15 ft

45 ft

C S_2

Region 1   Region 2   Region 3

Impervious

B   B   D
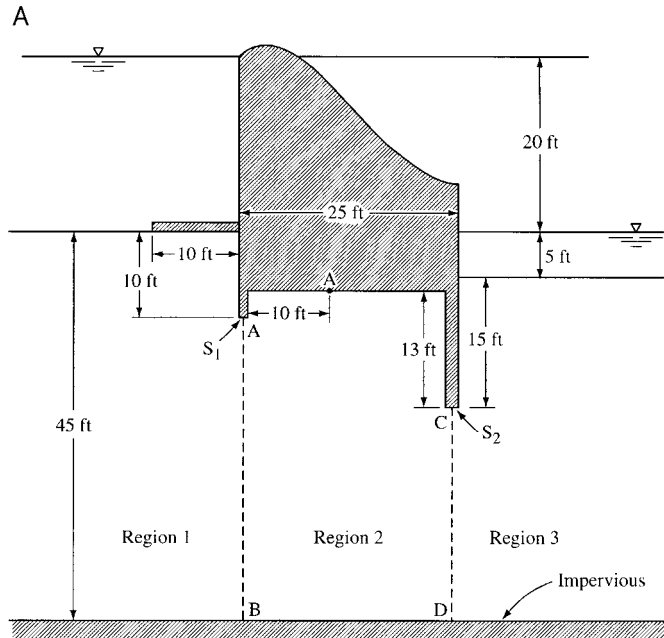
B   **Type III Fragment**

$b = 10$ ft

$\bar{S}_1 = 10$ ft, $V_{S_1} = 15\%$, $\sigma_{s_1} = 1.5$ ft

$s_{1+} = 11.5$ ft $s_{1-} = 8.5$ ft

$$\frac{Q}{kh} = \frac{1}{2\Phi}$$

$\bar{T} = 45$ ft, $V_T = 30\%$, $S_T = 13.5$ ft

$T_+ = 58.5$ ft $T_- = 31.5$ ft

		$\frac{1}{2\Phi}$	$\Phi_{\pm\pm}$	S_T
$\frac{b}{T_+}=0.171$	$\dfrac{S_+}{T_+} = \dfrac{11.5}{58.5} = 0.197$	.74	0.676	++
	$\dfrac{S_-}{T_+} = 0.145$	.78	0.641	-+
$\frac{b}{T_-}=0.317$	$\dfrac{S_+}{T_-} = 0.365$	.52	0.962	+-
	$\dfrac{S_-}{T_-} = 0.270$	.57	0.877	--

FIGURE 4.32 (A) Cross section for Example 20. (B) Region 1 variables and form factors.

C Type VI Fragment

$\bar{S}_1 = 3$ ft, $V_{S_1} = 15\%$, $\sigma_{S_1} = 0.45$ ft $\qquad S_{1+} = 3.45$ ft $\qquad S_{1-} = 2.65$ ft

$\bar{a}_1 = 35$ ft $\qquad a_{1+} = 34.55$ ft $\qquad a_{1-} = 35.33$ ft

$\bar{S}_2 = 13$ ft, $V_{S_2} = 15\%$, $\sigma_{S_2} = 1.95$ ft $\qquad S_{2+} = 14.95$ ft $\qquad S_{2-} = 11.05$ ft

$\bar{a}_2 = 25$ ft $\qquad a_{2+} = 23.05$ ft $\qquad a_{2-} = 26.95$ ft

$\bar{T} = 38$ ft, $V_T = 30\%$, $\sigma_T = 11.4$ ft $\qquad T_+ = 49.4$ ft $\qquad T_- = 26.6$ ft

$L = 25$ ft

L is always $> S_1 + S_2$

$$\Phi_2 = \ln\left[(1 + S_1/a_1)(1 + S_2/a_2)\right] + \frac{L - (S_1 + S_2)}{T}$$

$$\Phi_{S_1,S_2,T}^{+++} = \ln\left[\left(1 + \frac{3.45}{34.55}\right)\left(1 + \frac{14.95}{23.05}\right)\right] + \frac{25 - (3.45 + 14.95)}{49.4} = 0.729$$

$\Phi_{-++} = 0.722$

$\Phi_{+-+} = 0.651$

$\Phi_{--+} = 0.645$

$\Phi_{++-} = 0.843$

$\Phi_{-+-} = 0.851$

$\Phi_{+--} = 0.834$

$\Phi_{---} = 0.841$

D Type II Fragment

$\bar{S}_2 = 15$ ft, $V_{S_2} = 15\%$, $\sigma_{S_2} = 2.25$ ft $\qquad S_{2+} = 17.25$ ft $\quad S_{2-} = 12.75$ ft

$\bar{T} = 40$ ft, $V_T = 30\%$, $\sigma_T = 12$ ft $\qquad T_+ = 52$ ft $\quad T_- = 28$ ft

$b = 0$ thus $b/T = 0$

$\dfrac{S_+}{T_+} = 0.332$, $\dfrac{1}{2\Phi} = 0.64$ $\qquad\qquad \Phi_{++} = 0.781$

$\dfrac{S_-}{T_+} = 0.245$, $\dfrac{1}{2\Phi} = 0.73$ $\qquad\qquad \Phi_{-+} = 0.685$

$\dfrac{S_+}{T_-} = 0.616$, $\dfrac{1}{2\Phi} = 0.41$ $\qquad\qquad \Phi_{+-} = 1.220$

$\dfrac{S_-}{T_-} = 0.455$, $\dfrac{1}{2\Phi} = 0.53$ $\qquad\qquad \Phi_{--} = 0.943$

from Fig 4.17

FIGURE 4.32 (C) Region 2 variables and form factors. (D) Region 3 variables and form factors.

$$\phi = \ln\left[\left(1 + \frac{s1}{a1}\right)\left(1 + \frac{s2}{a2}\right)\right] + \frac{L - (s1 + s2)}{T} \tag{77}$$

The necessary eight form factors are shown in the f

Region 3 is a type II fragment which consists of only a partial cutoff (s2). The graph in Figure 4.17 is used again to determine the form factor based on the s_2/T ratio. The plus and minus terms of the ratio and the resulting form factors are shown on Figure 4.32d. As there are two random variables, only four form factors are required for this fragment.

Figure 4.33 shows the summation of the appropriate ϕ terms. Equation (64) is then used to calculate the quantity of flow for the sixteen permutations of plus and minus terms required for four random variables. Whereas the coefficient of variation of the soil permeability was taken to be 100%, the plus term is twice and mean value and the minus term is zero (Figure 4.34). Thus, half of the flow values will be zero. The mean value for the quantity of flow is seen to be 23.62 cubic feet per day with a standard deviation of 24.23 cubic feet per day. This results in a coefficient of variation of 102.6%.

$$\Sigma\Phi \text{ Terms}$$

S_1	S_2	T	Φ_1	Φ_2	Φ_3	$\Sigma\Phi_{\pm\pm\pm}$
+	+	+	.676	.729	.781	2.186
-	+	+	.641	.722	.781	2.144
+	-	+	.676	.651	.685	2.012
-	-	+	.641	.645	.685	1.971
+	+	-	.962	.843	1.220	3.025
-	+	-	.877	.851	1.220	2.948
+	-	-	.962	.834	.943	2.739
-	-	-	.877	.841	.943	2.661

FIGURE 4.33 Sum of form factors.

Example 21: Obtain the probability distribution of the quantity of flow for the weir in Example 20.

Solution: With a coefficient of variation of 102.6%, the normal distribution would lead to the absurdity of negative, $P[Q < 0] = 17\%$, flow quantities! Consequently, the beta distribution will be used with a lower boundary of zero. The upper boundary will be taken to be three standard deviations above the mean. Thus, the bounds for the distribution are 0 to 96.32 cubic feet per day. The resulting distribution is shown in Figure 4.35. This example illustrates the necessity of using the beta distribution rather than the normal distribution when estimating the quantity of flow.

4.8   Piping: Reliability

By virtue of the viscous friction exerted on a liquid flowing through a porous medium, an energy transfer is effected between the liquid and the solid particles. The measure of this energy is the head loss h between the points under consideration. The force corresponding to this energy transfer is called the *seepage force*. It is this seepage force that is responsible for the phenomenon known as *quicksand*, and its assessment is of vital importance in the stability of structures founded in soils and subject to the action of flowing water (seepage).

The first rational approach to the determination of the effects of seepage forces was presented by Terzaghi (1922) and forms the basis for all subsequent studies. Consider all the forces acting on a volume of particulate matter through which a liquid flows.

$Q = \dfrac{kh}{\Sigma\Phi}$ Random Variables k, $\Phi(S_1, S_2, T)$

(thus 2^4 or 16 terms)

$h = 20$ ft

$\bar{k} = 1 \times 10^{-3}$ cm/sec $= 2.83$ ft/day, $V_k = 100\%$, $S_k = 2.83$ ft/day

$k_+ = 5.66$ ft/day $k_- = 0.0$

$E[Q] = \displaystyle\sum_{i=1}^{16} P_{\pm i} \, Q_{\pm i}^2$ $P_{\pm i} = \dfrac{1}{16}$ since all terms are taken to be uncorrelated

$\sigma_Q^2 = \displaystyle\sum_{i=1}^{16} P_{\pm i} \, Q_{\pm i}^2 - \left(E[Q]^2\right)$

S_1	S_2	T	k	Q_i	Q_i^2
+	+	+	+	51.950	2698.85
-	+	+	+	52.627	2769.57
+	~	+	+	56.431	3184.43
-	-	+	+	57.258	3278.53
+	+	-	+	37.421	1400.37
-	+	-	+	38.399	1474.48
+	-	-	+	41.329	1708.08
-	-	-	+	42.540	1809.69
+	+	+	-	0.0	0.0
-	+	+	-	0.0	0.0
+	-	+	-	0.0	0.0
-	-	+	-	0.0	0.0
+	+	-	-	0.0	0.0
-	+	-	-	0.0	0.0
+	-	-	-	0.0	0.0
-	-	-	-	0.0	0.0

$\Sigma = 377.955$ 18324.0 $\Sigma/16 = 1145.25$

$E[Q] = 23.62$ ft^3/day

$\sigma_Q^2 = 1145.25 - (23.62)^2 = 587.24$

$\sigma_Q = 24.23$ ft^3/day

$V_Q = 1.026$ or 102.6%

FIGURE 4.34 Flow quantities.
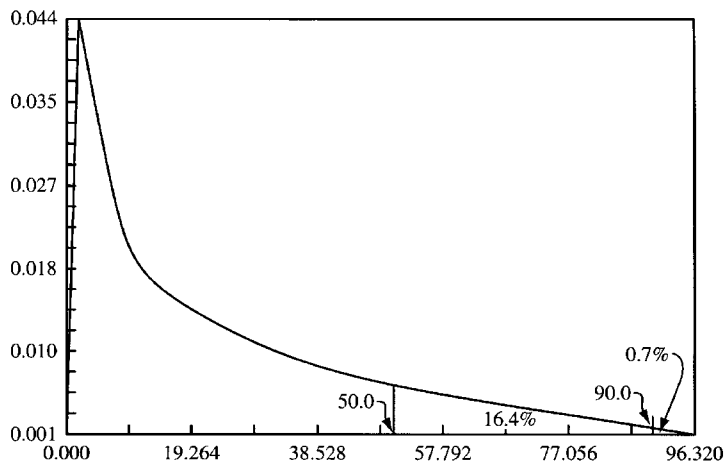
FIGURE 4.35 Beta distribution of flow quantity for Example 20.

1. The total weight per unit volume, the mass unit weight is

$$\gamma_m = \frac{\gamma_1(G+e)}{1+e}$$

where e is the void ratio, G is the specific gravity of solids, and γ_1 is the unit weight of the liquid.
2. Invoking Archimedes' principle of buoyancy that *a body submerged in a liquid is buoyed up by force equal to the weight of the liquid displaced*, the effective unit weight of a volume of soil, called the *submerged unit weight*, is

$$\gamma'_m = \gamma_m - \gamma_1 = \frac{\gamma_1(G-1)}{1+e} \tag{78}$$

To gain a better understanding of the meaning of the submerged unit weight, consider the flow condition shown in Figure 4.36a. If the water column (AB) is held at the same elevation as the discharge face CD ($h = 0$), the soil will be in a submerged state and the downward force acting on the screen will be

$$F\downarrow = \gamma'_m LA \tag{79}$$

where $\gamma'_m = \gamma_m - \gamma_w$. Now, if the water column is slowly raised (shown dotted to A'B'), water will flow up through the soil. By virtue of this upward flow, work will be done to the soil and the force acting on the screen will be reduced.
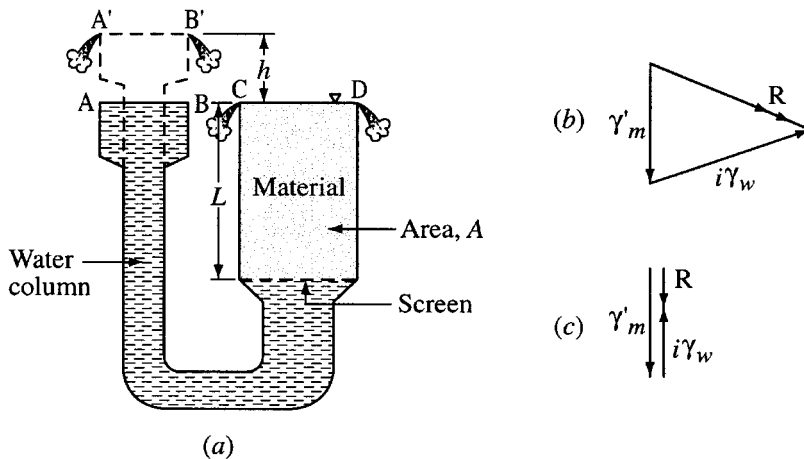
FIGURE 4.36 Development of seepage force concept.

3. The change in force through the soil is due to the increased pressure acting over the area or

$$F\uparrow = h\gamma_w A$$

Hence, the change in force, granted steady-state conditions, is

$$\Delta F = \gamma'_m LA - h\gamma_w A \tag{80}$$

Dividing by the volume AL, the resultant force per unit volume acting at a point within the flow region is

$$R = \gamma'_m - i\gamma_w \qquad (81)$$

where i is the hydraulic gradient. The quantity $i\gamma_w$ is called the seepage force (force per unit volume). In general, Equation (81) is a vector equation, with the seepage force acting in the direction of flow (Figure 4.36b). Of course, for the flow condition shown in Figure 4.36a, the seepage force will be directed vertically upward (Figure 4.36c). If the head h is increased, the resultant force R in Equation (81) is seen to decrease. Evidently, should h be increased to the point at which $R = 0$, the material would be at the point of being washed upward. Such a state is said to produce a *quick* (meaning alive) condition. From Equation (81) it is evident that a quick condition is incipient if

$$i_{cr} = \frac{\gamma'_m}{\gamma_w} = \frac{G-1}{1+e} \qquad (82)$$

Substituting typical values of $G = 2.65$ (quartz sand) and $e = 0.65$ (for sand $0.57 \le e \le 0.95$), we see that as an average value the critical gradient can be taken as

$$i_{cr} \approx 1 \qquad (83)$$

When information is lacking as to the specific gravity and void ratio of the material, the critical gradient is generally taken as unity.

At the critical gradient, there is no interparticle contact ($R = 0$); the medium possesses no intrinsic strength and will exhibit the properties of liquid of unit weight

$$\gamma_q = \left(\frac{G+e}{1+e}\right)\gamma_1 \qquad (84)$$

Substituting the above values for G, e, and $\gamma_1 = \gamma_w$, $\gamma_q = 124.8$ lb/ft^3. Hence, contrary to popular belief, a person caught in quicksand would not be sucked down, but would find it almost impossible to avoid floating.

Many hydraulic structures, founded on soils, have failed as a result of the initiation of a local quick condition which, in a chain-like manner, led to severe internal erosion called *piping*. This condition occurs when erosion starts at the exit point of a flow line and progresses backward into the flow region, forming pipe-shaped watercourses which may extend to appreciable depths under a structure. It is characteristic of piping that it needs to occur only locally and that once begun it proceeds rapidly, and is often not apparent until structural failure is imminent.

Equations (82) and (83) provide the basis for assessing the safety of hydraulic structures with respect to failure by piping. In essence, the procedure requires the determination of the maximum hydraulic gradient along a discharge boundary, called the *exit gradient*, which will yield the minimum resultant force (R_{min}) at this boundary. This can be done analytically, as will be demonstrated below, or from flow nets after a method proposed by Harza (1935).

In the graphical method, the gradients along the discharge boundary are taken as the macrogradient across the contiguous squares of the flow net. As the gradients vary inversely with the distance between adjacent equipotential lines, it is evident that the maximum (exit) gradient is located where the vertical projection of this distance is a minimum, such as at the toe of the structure (point C) in Figure 4.14. For example, the head lost in the final square of Figure 4.14 is one-sixteenth of the total head loss of 16 ft, or 1 ft, and as this loss occurs in a vertical distance of approximately 4 ft, the exit gradient at point C is approximately 0.25. Once the magnitude of the exit gradient has been found, Harza recommended

that the factor of safety be ascertained by comparing this gradient with the critical gradient of Equations (82) and (83). For example, the factor of safety with respect to piping for the flow conditions of Figure 4.14 is 1.0/0.25 or 4.0. Factors of safety of 4 to 5 are generally considered reasonable for the graphical method of analysis.

The analytical method for determining the exit gradient is based on determining the exit gradient for the type II fragment, at point E in Figure 4.23. The required value can be obtained directly from Figure 4.37 with h_m being the head loss in the fragment.
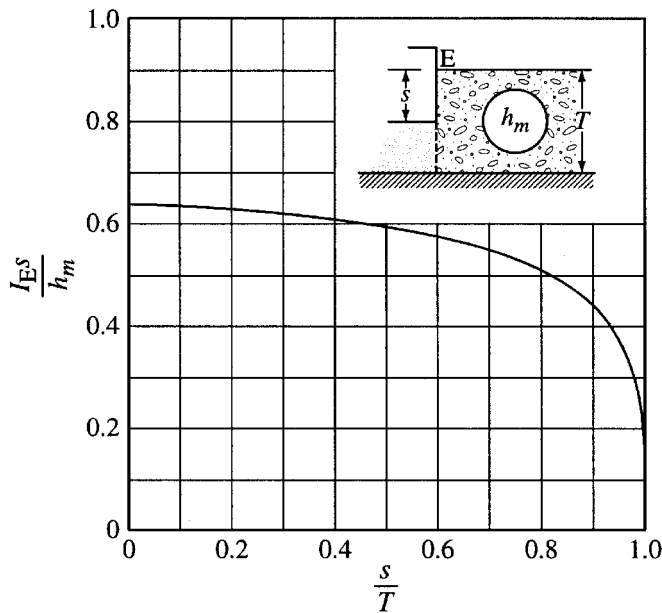
FIGURE 4.37 Exit gradient.

Example 22: Find the exit gradient for the section shown in Figure 4.26a.

Solution: From the results of Example 15, the head loss in fragment 3 is $h_m = 5.9$ ft. With $s/T = 1/3$, from Figure 4.37 we have $I_E s/h_m = 0.62$; hence,

$$I_E = \frac{0.62 \times 5.9}{9} = 0.41 \quad \text{or} \quad FS = 1/0.41 = 2.44$$

To account for the deviations and uncertainties in nature, Khosla, Bose, and Taylor (1954) recommend that the following factors of safety be applied as critical values of exit gradients; gravel, 4 to 5; coarse sand, 5 to 6; and fine sand, 6 to 7.

The use of reverse filters on the downstream surface, or where required, serves to prevent erosion and to decrease the probability of piping failures. In this regard, the Earth Manual of the U.S. Bureau of Reclamation, Washington, 1974, is particularly recommended.

For Further Information

Ang, A. H.-S. and Tang, W. H. 1975. *Probability Concepts in Engineering Planning and Design, Vol. I – Basic Principles*, John Wiley & Sons, New York.

Guymon, G. L., Harr, M. E., Berg, R. L., and Hromadka, T. V. 1981. A probabilistic-deterministic analysis of one-dimensional ice segregation in a freezing soil column, *Cold Reg. Sci. Tech.* 5, 127-140.

Hahn, G. J. and Shapiro, S. S. 1967. *Statistical Models in Engineering*, John Wiley & Sons, New York.

Jayne, E. T. 1978. Where do we stand on maximum entropy?, in *The Maximum Entropy Formalism*, Eds. R. D. Levine and M. Tribus, MIT Press, Cambridge, MA.

Tribus, M. 1969. *Rational Descriptions, Decisions and Designs*, Pergamon Press, New York.

Whitman, R. Y. 1984. Evaluating calculated risk in geotechnical engineering, *J. Geotech. Eng.*, *ASCE*, 110(2).

References

Bear, J. 1972. *Dynamics of Fluids in Porous Media*. American Elsevier Publication Company, New York.

Bear, J., Zaslavsky, D., and Irmay, S. 1966. *Physical Principles of Water Percolation and Seepage*. Technion, Israel.

Benjamin, J. R. and Cornell, C. A. 1970. *Probability, Statistics, and Decision for Civil Engineers*. McGraw-Hill, Inc. New York.

Bercha, F. G. 1978. Application of probabilistic methods in ice mechanics. Preprint 3418, ASCE.

Borri, A., Ceccotti, A., and Spinelli, P. 1983. Statistical analysis of the influence of timber defects on the static behavior of glue laminated beams, vol. 1, in 4*th Int. Conf. Appl. Prob. Soil Struct. Eng.*, Florence, Italy.

Brahma, S. P. and Harr, M. E. 1962. Transient development of the free surface in a homogeneous earth dam, *Géoetechnique*, 12, 4.

Casagrande, A. 1940. Seepage through dams, "Contributions to Soil Mechanics 1925-1940," Boston Society of Civil Engineering, Boston.

Cedergren, H. R. 1967. *Seepage, Drainage and Flow Nets*. John Wiley & Sons, inc., New York.

Dachler, R. 1934. Ueber den Strömungsvorgang bei Hanquellen, *Die Wasserwirtschaft*, 5.

Darcy, H. 1856. *Les Fontaines Publique de la Ville de Dijon*. Paris.

Domenico, P. A. and Schwartz, F. W. 1990. *Physical and Chemical Hydrogeology*. John Wiley & Sons, New York.

Ditlevsen. 1981. *Uncertainty Modeling*. McGraw-Hill, New York.

Dvinoff, A. H. and Harr, M. E. 1971. Phreatic surface location after drawdown, *J. Soil Mech. Found. Div.*, *ASCE*, January.

Earth Manual. 1974. Water Resources Technology Publication, 2nd ed., U.S. Department of Interior, Bureau of Reclamation, Washington.

Edris, E. V. 1990. Probabilistic applications, U.S. Army Corps of Engineers, Waterways Experiment Station, Vicksburg, Mississippi.

Ellingwood, B., Galambos, T. V., MacGregor, J. G., and Cornell, C. A. 1980. Development of a probability based load criterion for American National Standard A58. *Nat. Bur. Stand. Spec. Publ.* 577, Washington, D.C.

Forchheimer, P. 1930. *Hydraulik*. Teubner Verlagsgesellschaft, mbh, Stuttgart.

Fredlund, D. G. and Dahlman, A. E. 1972. Statistical geotechnical properties of glacial Lake Edmonton sediments, in *Statistics and Probability in Civil Engineering*. Hong Kong University Press (Hong Kong International Conference), Ed. P. Lumb, distributed by Oxford University Press, London.

Freeman, H. 1963. *Introduction to Statistical Inference*. John Wiley & Sons, New York.

Freeze, R. A. and Cherry, J. A. 1979. *Groundwater*. Prentice-Hall, Englewood Cliffs, NJ.

Galambos, T. V., Ellingwood, B., MacGregor, J. G., and Cornell, C. A. 1982. Probability based load criteria: Assessment of current design practice. *J. Struct. Div.*, *ASCE*, 108(ST5).

Grivas, D. A. and Harr, M. E. 1977. Reliability with respect to bearing capacity failures of structures on ground, 9*th Int. Conf. Soil Mech. Found.* Tokyo, Japan.

Grivas, D. A. and Harr, M. E. 1979. A reliability approach to the design of soil slopes, 7*th Eur. Conf. on S.M.A.F.E.* Brighton, England.

Hammitt, G. M. 1966. Statistical analysis of data from a comparative laboratory test program sponsored by ACIL, United States Army Engineering Waterways Experiment Station, Corps of Engineers, Miscellaneous Paper No. 4-785.

Harr, M. E. 1962. *Groundwater and Seepage*. McGraw-Hill Inc., New York.

Harr, M. E. and Lewis, K. H. 1965. Seepage around cutoff walls. *RILEM*, Bulletin 29, December.

Harr, M. E. 1976. Fundamentals of probability theory. *Transp. Res. Rec.* 575.

Harr, M. E. 1977. *Mechanics of Particulate Media: A Probabilistic Approach*, McGraw-Hill Inc., New York.

Harr, M. E. 1987. *Reliability-Based Design in Civil Engineering*, McGraw-Hill Inc., New York.

Harza, L. F. 1935. Uplift and seepage under dams on sand, *Transactions ASCE*, vol. 100.

Haugen, E. B. 1968. *Probabilistic Approaches in Design*, John Wiley & Sons, New York.

Khosla, R. B. A. N., Bose, N. K., and Taylor, E. McK. 1954. *Design of Weirs on Permeable Foundations*. Central Board of Irrigation, New Delhi, India.

Lamb, H. 1945. *Hydrodynamics*. Dover Publications, New York.

Lipshutz, S. 1965. *Schaum's Outline of Theory and Problems of Probability*, McGraw-Hill Inc., New York.

Lumb, P. 1972. Precision and accuracy of soils tests, in *Statistics and Probability in Civil Engineering*, Hong Kong University Press (Hong Kong International Conference), ed. P. Lumb, distributed by Oxford University Press, London.

Lumb, P. 1974. Application of statistics in soil mechanics, in *Soil Mechanics – New Horizons*, ed. I. K. Lee, American Elsevier, New York.

Muskat, M. 1937. *The Flow of Homogeneous Fluids through Porous Media*, McGraw-Hill Inc., New York. Reprinted by J. W. Edwards, Ann Arbor, 1946.

Padilla, J. D. and Vanmarcke, E. H. 1974. *Settlement of structures on shallow foundations: A Probabilistic Analysis*, Research Report R74-9, M.I.T.

Pavlovsky, N. N. 1956. *Collected Works*, Akad. Nauk USSR, Leningrad.

Pearson, E. S. and Hartley, H. O. 1972. *Biometrika Tables for Statisticians*, Vol. II, Cambridge University Press, London.

Pearson, K. 1894,1895. Skew variations in homogeneous material, contributions to the mathematical theory of evolution, *Philos. Trans. R. Soc.*, Vol. 185 and Vol. 186.

Polubarinova-Kochina, P. Ya. 1941. Concerning seepage in heterogeneous two-layered media, *Inzhenernii Sbornik*, 1, 2.

Polubarinova-Kochina, P. Ya. 1952. *Theory of the Motion of Ground Water*, Gostekhizdat, Moscow.

Prasil, F. 1913. *Technische Hydrodynamik*, Springer-Verlag, Berlin.

Rosenblueth, E. 1975. Point estimates for probability moments, *Proc. Natl. Acad. Sci*, USA, 72(10).

Rosenblueth, E. 1981. Two-point estimates in probabilities, *Appl. Math. Modeling*, Vol. 5.

Scheidegger, A. E. 1957. *The Physics of Flow Through Porous Media*, The MacMillan Company, New York.

Schultze, E. 1972. Frequency distributions and correlations of soil properties, in *Statistics and Probability in Civil Engineering*, Hong Kong University Press (Hong Kong International Conference), ed. P. Lumb, distributed by Oxford University Press, London.

Terzaghi, K. 1922. Der Grundbruch and Stauwerken und seine Verhütung, *Die Wasserkraft*, 445.

Glossary

Capacity   The ability to resist an induced demand; resistance or strength of entity.

Coefficient of Permeability   Coefficient of proportionality between Darcy's velocity and the hydraulic gradient.

Correlation Coefficient   Measure of the compliance between two variables.

Demand   Applied loading or energy.

Expected Value, Expectation   Weighted measure of central tendency of a distribution.

Flow Net   Trial-and-error graphical procedure for solving seepage problems.

Hydraulic Gradient   Space rate of energy dissipation.

Method of Fragments   Approximate analytical method for solving seepage problems.

Piping   Development of a "pipe" within soil by virtue of internal erosion.

Probability   Quantitative measure of a state of knowledge.

Quick Condition   Condition when soil "liquifies."

Random Variable   An entity whose measure cannot be predicted with certainty.

Regression   Means of obtaining a functional relationship among variables.

Reliability   Probability of an entity (or system) performing its required function adequately for a specified period of time under stated conditions.

Standard deviation   Square root of variance.

Variance   Measure of scatter of variable.

5

Soil Properties and Moisture Movement in the Unsaturated Zone

Randell Haverkamp,
Faycal Bouraoui,
Christian Zammit, and
Rafael Angulo-Jaramillo

Laboratoire d'Etude des Transferts en Hydrologie et Environnement Grenoble, France

5.1 Introduction

The *unsaturated zone*, sometimes called the *vadose zone*, is the zone between the ground surface and the water table. The term unsaturated zone is somewhat of a misnomer because the capillary fringe above the groundwater or the rain-saturated top soil are portions that are saturated. For this reason some authors (e.g., Bouwer, 1978) prefer the term vadose zone. The unsaturated zone is the hydrological connection between the surface water component of the hydrologic cycle and the groundwater component (Figure 5.1). The surface water component includes the precipitation as it reaches the land surface and artificial water application such as irrigation, surface runoff, stream flow, lakes, and artificial impoundments. The rainfall may infiltrate and some of the surface water may percolate through the unsaturated zone. Recharge of the groundwater usually occurs through the unsaturated zone. The unsaturated zone may lose water through evaporation, evapotranspiration, and drainage.

The unsaturated zone plays a crucial role in the transfer of pollutants. Many constituents present in the surface waters eventually find their way into the groundwater through the unsaturated zone.
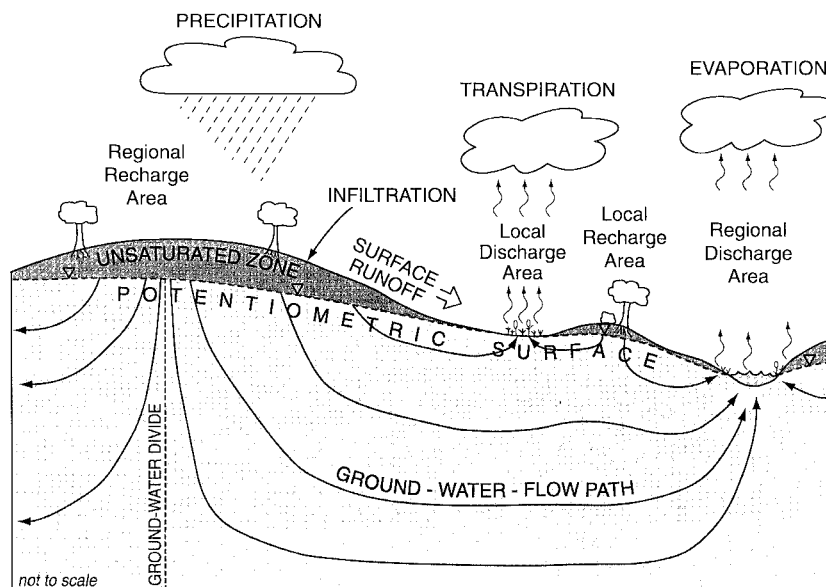
FIGURE 5.1 The unsaturated and saturated zones in the hydrologic cycle.

Accidental spills of chemicals, application of fertilizers and pesticides on the land surface, leaks from gasoline storage tanks, septic tank drainage, and leaching from landfills are examples of anthropogenic activities that contribute to the leaching of contaminants through the unsaturated zone into the ground-water (see Chapter 2, Figure 2.22). Chapter 16 succinctly develops the multiphase, multispecies equations used to represent the soil, water, and air interactions in the unsaturated zone. Figure 16.1 shows the principal contaminant sources and pathways in the soil-chemical system.

   An introduction to water in the unsaturated zone is given in Chapter 1, Section 1.5. This chapter is concerned with the movement of moisture in the unsaturated zone. It covers the conceptual aspects of unsaturated soil water flow related to soil physical properties and soil water characteristics. In Section 5.2 we will deal with the soil physical properties, the soil hydraulic properties, and the choice of the functional relationships used to describe the soil hydraulic characteristics. Additionally, we will present some of the measurement and prediction techniques used for the estimation of the soil hydraulic characteristics parameters. We will then address the conceptual aspects of the different unsaturated soil water transfer equations used in subsurface hydrology (Section 5.3). An overview of the most significant analytical solutions is presented in Section 5.4. The different empirical infiltration laws are discussed in more detail in Chapter 6. Section 5.5 sets out the details of a new scaling approach of the water transfer equation (Richards, 1931) and its implications for the soil characteristic scale parameters. Concluding remarks will be given in Section 5.6.

5.2   Physical and Hydraulic Soil Properties

5.2.1   Soil Phases

Soils comprises three phases: the solid, the liquid, and the gaseous phases. The *solid phase* includes the mineral particles of gravel, sands, silts, and clays. Gravel, sand and silt have a dense spherical or angular shape. Clays have a plate-like structure (with dimensions between 1 and 2 μm) and their capability to adsorb sodium or calcium controls whether they are dispersed or flocculent. Particle-size properties are determined from the size distribution of individual particles in a soil sample. Soil particles smaller than 2 mm are divided into the three soil texture groups: sand, silt, and clay. Particles larger than 2 mm are
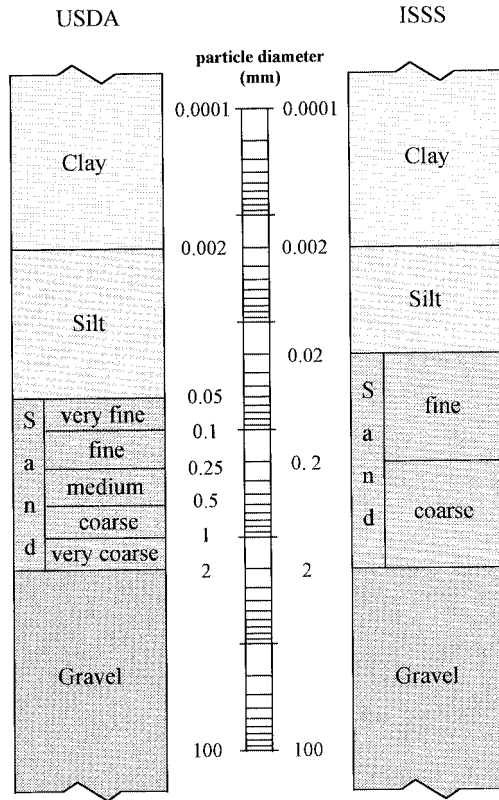
FIGURE 5.2 Particle-size limits according to the USDA and ISSS soil classification schemes.

grouped into the gravel class. Figure 5.2 shows the particle-size limits according to the U.S. Department of Agriculture (Soil Survey Laboratory Staff, 1992) and the International Soil Science Society (Yong and Warkentin, 1966). Figure 5.3 shows the USDA soil textural triangle which allows the classification of soils as a function of their particle-size distribution. For example, a soil consisting of 30% sand, 40% silt, and
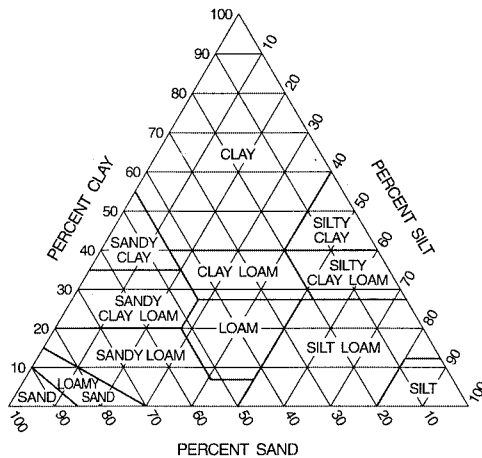
FIGURE 5.3 USDA soil textural classification chart showing the percentages of clay (below 2 μm), silt (2 to 50 μm), and sand (50 to 2000 μm).

30% clay by weight is a clay loam. Another grouping is the Unified Soil Classification System (USCS). It is based on the particle size, the plasticity index, and the liquid limit and is described in geotechnical engineering texts such as Lambe and Whitman (1979). The mineral particles, along with the interconnecting pores, form the *soil matrix*. The solid phase may also include organic material such as peat.

The fluids in the soil comprise liquids and gases. The *liquid phase* usually consists of water that can move through the pores of the soil. The water can include numerous dissolved minerals and organic compounds. Other liquids may be present. They may be miscible or immiscible in water and generally result from agricultural and industrial activities or accidental spills.

The principal component of the *gaseous phase* is air. Like water, air can also move through the soil and contains other dissolved gases like water vapor and volatile components, usually of pollutants. The principal characteristics of the unsaturated zone depend on the proportion of liquid and gas in the soil matrix. In this three-phase system, water and air can flow simultaneously.

5.2.2   Physical Properties of Soils

Let V_S, V_l, V_g, V_t, M_S, M_l, M_g, M_t and W_S, W_l, W_g, W_t be the volumes, masses, and weights of solid, liquid, gas, and total, respectively. Then the density of the solid particles (ρ_S), the dry bulk density (ρ_d), and the porosity (ε) are given by:

$$\rho_s = \frac{M_s}{V_s} = \frac{W_s}{gV_s} \tag{1}$$

$$\rho_d = \frac{M_s}{V_t} = \frac{W_s}{gV_t} \tag{2}$$

$$\varepsilon = \frac{V_l + V_g}{V_t} = \frac{V_t - V_s}{V_t} = 1 - \frac{\rho_d}{\rho_s} \tag{3}$$

where g is gravitational acceleration (≈ 9.82 cm/s^2). The particle density (ρ_S) is normally assumed to be equal to 2.65 g/cm^3. The void ratio (e) is useful when V_t is not constant due to swelling or shrinking and is defined as:

$$e = \frac{V_l - V_g}{V_s} = \frac{\varepsilon}{1 - \varepsilon} \tag{4}$$

Typical values of porosity and density of soils are given in Chapter 16, Table 16.2. A classification of pores and typical pore volumes are listed in Tables 16.3 and 16.4. If the definition of Equation (3) is applied to a very small elementary volume, the porosity would be equal to one if the element is centered in a pore or would be equal to zero if the element is centered in the solid. The element under consideration has to be equal to or larger than a minimum volume called *representative elementary volume* (REV) to reach a stable value. The continuum approach is used in this chapter. That is to say, a representative elementary volume is assumed such that the fluid flow can be represented as the product of a transfer coefficient and an energy gradient. For a detailed discussion of the concept of the REV, see Bear and Bachmat (1984).

When small stones and porous coarse fragments (diameter > 2 mm) occur in the elementary volume, the whole soil dry bulk density (ρ_{ap}) measured over the whole soil sample is different from the fine soil dry bulk density. The correction formula is:

$$\rho_d = \frac{1 - \alpha - \beta}{\dfrac{1}{\rho_{ap}} - \dfrac{\alpha}{\rho_s} - \dfrac{\beta}{\rho_c}} \tag{5}$$

where ρ_d is the dry bulk density of the fine soil material; ρ_c is the specific weight of the porous coarse fragments; and α and β are the stone and porous rock weight fractions, respectively, of the whole soil sample. The whole soil porosity (ε_{ap}) and the porosity of the porous coarse fragments (ε_c) are then defined as:

$$\varepsilon_{ap} = 1 - \frac{\rho_{ap}}{\rho_s} \tag{6}$$

$$\varepsilon_c = 1 - \frac{\rho_c}{\rho_s} \tag{7}$$

Dry bulk density and soil porosity are important soil morphological parameters because they are closely related to soil structure. Hence, they influence the soil water transfer properties to a large extent.

5.2.3 Hydraulic Properties of Soils

Regardless of the scales involved, the soil hydraulic properties which affect the flow behavior are incorporated into two fundamental characteristics: (1) the soil water retention curve describing the relation between volumetric soil water content and soil water pressure; and (2) the relation between volumetric water content and hydraulic conductivity.

Soil water content. The soil water content can be expressed by mass (w) or by volume (θ):

$$w = \frac{M_l}{M_s} = \frac{W_l}{W_s} \tag{8}$$

$$\theta = \frac{V_l}{V_t} = \frac{w\rho_d}{\rho_w} \tag{9}$$

where ρ_w is the specific density of water ($\rho_w \approx 1$ g/cm³). In most hydrologic applications, volumetric soil water content is used in non-dimensional form : $\theta^* = \theta/\theta_S$, where θ_S is the *volumetric soil water content at natural saturation.* Due to air entrapment, water content at natural saturation (θ_S) seldom reaches saturation of the total pore space (ε), e.g., Rogowski (1971) gives $\theta_S = 0.9\ \varepsilon$. Another commonly used water content term is *residual volumetric soil water content* (θ_r). Conceptually, residual water content can be associated with the immobile water present within a dry soil profile in films on particle surfaces, in interstices between particles, and within soil pores. When volumetric water content is expressed as the ratio to total pore space (ε) it is referred to as the degree of saturation (S):

$$S = \frac{V_l}{V_l + V_g} = \frac{\theta}{\varepsilon} \tag{10}$$

Soil water pressure. Matric potential or *capillary potential* (ϕ_c) is the measure of the energy status of water retained in a soil by capillarity and surface adsorption (see also Chapter 1, Section 1.4). It is one of the three prevalent components of *total soil water potential* (Φ) given by:

$$\Phi = \phi_g + \phi_c + \phi_o \qquad (11)$$

where ϕ_g is the *gravitational potential* and ϕ_o is the *osmotic potential* (e.g., van Bavel, 1969). The total soil water potential (Φ) measures the energy per unit quantity (volume, mass, or weight) that is required to move a small volume of water from a pool of water at a datum elevation and at standard atmospheric pressure, to the soil at the point under consideration. The sum of the matric potential and the gravity potential is often referred to as the *hydraulic potential*. When there is a density gradient of dissolved substances it is necessary to add an osmotic potential. Following the International Society of Soil Science (Bolt, 1976) two other components can be included: the *air potential* (ϕ_a), and the *envelope potential* (ϕ_e). The air potential accounts for the difference in pressure that can exist between the air inside the pores and the outside atmospheric pressure over the reference water. When the external pressure due to the overburden is transmitted by the water envelope around the particles, the envelope potential can be significant.

The potential can be expressed per unit volume, per unit mass, or per unit weight. If the potential is expressed in terms of a volume, the units are joules per cubic meter, or equivalently in units of pressure: newtons per square meter or pascals. If the potential is expressed per unit mass or per unit weight, the units are joules per kilogram and meters, respectively. The latter option is conventionally used for most hydrologic studies. When taken per unit weight, the *gravity potential* corresponds to the elevation head and the *capillary or matric potential* corresponds to the *soil water pressure head* (h):

$$\varphi_c = \rho_w g h \qquad (12)$$

The pressure head h is positive below the water table and negative above it (see also Chapter 2, Figure 2.1). In the unsaturated zone the negative of the pressure is also called the *matric suction* or *tension* which, hence, is expressed as a positive number. A fully detailed description of the different notations is given by Kutilek and Nielsen (1994). In dry soils the pressure head can reach extremely high values; for computational convenience its value is often expressed in the logarithmic mode:

$$pF = \log|h| \qquad (13)$$

The value of $pF = 4.2$ (i.e., a pressure head of 15,000 cm of water) is the threshold pressure value at which plant transpiration can take place. This point, which is referred to as the *wilting point*, is usually associated with the volumetric soil water content value θ_r (van Genuchten, 1980).

Water retention characteristic. The variation of volumetric soil water content (θ) with soil water suction (h) is referred to as the water retention characteristic ($h(\theta)$). It describes the soil's ability to store or release water. The water retention characteristic is a highly nonlinear S-shaped curve. Figure 5.4 gives an illustration of the typical behavior of the water retention characteristic. It clearly shows that the water retention curve varies with soil structure. For sandy soils the shape of $h(\theta)$ is characterized by a typical step form; for clay soils, on the other hand, the water retention curve is rather steep. Hence, for identical pressure head values clay soils measure generally higher water content values than sandy soils.

The characteristic curves may be different for drying (desorbing) and wetting (absorbing) soils. This phenomenon, referred to as *hysteresis*, can roughly be explained by the ink bottle effect, namely, the fact that the opening radius r at the top of any pore is generally smaller than the radius R of the main pore. Following Laplace's equation, the capillary rise in a tube of radius r is known to be inversely proportional to the radius:

$$h = \frac{2\sigma_{LV}\cos(\omega)}{rg\rho_w} \qquad (14)$$
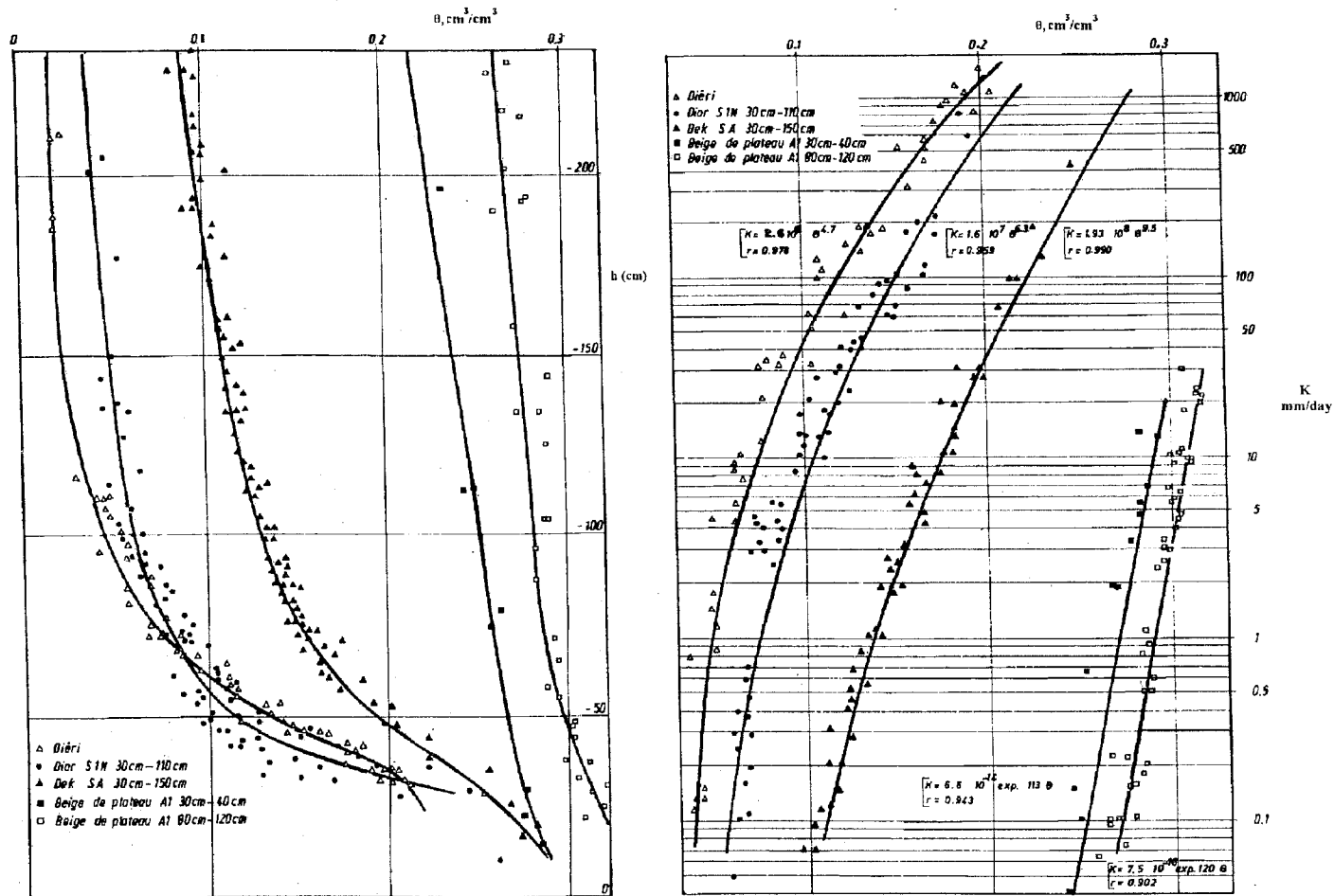
FIGURE 5.4 Water retention and hydraulic conductivity curves of five different soil types measured in Senegal (Adapted from Hamon, G. 1980. Mise en oeuvre et critique de méthodes de caractérisation hydrodynamique de la zone non-saturée du sol. Applications aux sols de culture du Sénégal. *Thèse Docteur-Ingénieur*, Université Scientifique et Médicale de Grenoble, France.).

where σ_{LV} is the liquid–vapor surface tension of water; and ω is the contact angle between liquid and solid (see also Equation (7), Chapter 1). At 20°C, with ρ_w equal to 1 g/cm^3 and ω equal to zero, Equation (14) is reduced to:

$$h \approx \frac{0.149}{r} \tag{15}$$

where soil water suction (h) and pore radius (r) are expressed in cm (e.g., Rose, 1966). Upon drying, the pore remains full until the capillary rise surpasses $0.149/r$. For wetting, the water starts filling the pore when h drops to the lower value $0.149/R$. The water content of the drying pore thus surpasses that of the wetting pore causing the hysteresis. Theoretically, for a given soil geometry, it should be possible to predict hysteretic effects from first principles. However, this problem remains largely unsolved, and our rather sketchy understanding of soil structure suggests that only a few soil models will yield to this approach. Instead, hysteresis in soils remains largely based on Poulovassilis' (1962) analysis obtained by applying the independent domain theory to soils (see also Topp, 1971; and Mualem, 1974). Parlange (1976 and 1980) proposed an alternative theory which has proven to be extremely precise and robust in spite of its operational simplicity. It requires only one boundary curve of the hysteresis envelope to predict the other boundary and all scanning curves in between (see also Hogarth et al., 1988; Liu et al., 1995; and Parlange et al., 1996). An example of a hysteresis-affected water retention curve measured on a vertical soil column, is shown in Figure 5.5.

The phenomenon of hysteresis is more important for sands than for clay soils. Laboratory experiments have clearly shown that hysteresis effects can be important for infiltration into soils with dry, nonuniform initial water content profiles (Vachaud and Thony, 1971). However, its influence under field conditions is often masked by heterogeneities and spatial variability.

Hydraulic conductivity. The isothermal hydraulic conductivity was originally introduced by Darcy (1856) for saturated soils and extended to unsaturated soils by Buckingham (1907) and Richards (1931). The hydraulic conductivity is a measure of the ability of the soil to route water. From a theoretical view point $K(\theta)$ can be expressed as:

$$K(\theta) = k \frac{\rho_w g}{\mu_w} k_{rw}(\theta) \tag{16}$$

where k is intrinsic permeability; $k_{rw}(\theta)$ is relative water permeability (namely the ratio of the unsaturated to the saturated water permeabilities), which varies from 0 for completely dry soils to 1 for fully saturated soils; and μ_w is the water viscosity. Equation (16) shows that the soil conductivity depends on the soil matrix (k), the moving fluid (ρ_w and μ_w), and the fluid content in the soil ($k_{rw}(\theta)$). The dependence of K on θ is unaffected by hysteresis. The hydraulic conductivity at or above saturation ($h \geq 0$) is referred to as the *hydraulic conductivity at natural saturation* (K_S).

When expressed as a function of the volumetric soil water content, the hydraulic conductivity function $K(\theta)$ is strongly nonlinear. Generally speaking, it behaves like a power function. For decreasing soil water content, the hydraulic conductivity decreases rapidly. Figure 5.4 illustrates the typical behavior of hydraulic conductivity. The hydraulic conductivity function depends highly on soil structure. For sandy soils, the hydraulic conductivity at natural saturation (K_S) is usually higher than that for clay soils, even though the porosity is higher in clay soils.

Besides the water retention and hydraulic conductivity functions, alternative combinations of soil characteristics can be chosen to characterize the soil moisture behavior of soils, such as $K(h)$ (e.g., Rijtema, 1965; Gardner, 1958), $D(\theta)$ (e.g., Childs and Collis-George, 1950), and/or $C(\theta)$ (e.g., Richards, 1931), where D refers to *diffusivity* defined by $D = K(\theta) \, dh/d\theta$ and C is *specific capacity* defined by $C(\theta) = d\theta/dh$. Both functions, $C(\theta)$ and $D(\theta)$, will be addressed in more detail in Section 5.3.1. The quintessence is that two soil characteristics are always required to model soil water movement in the vadose zone.
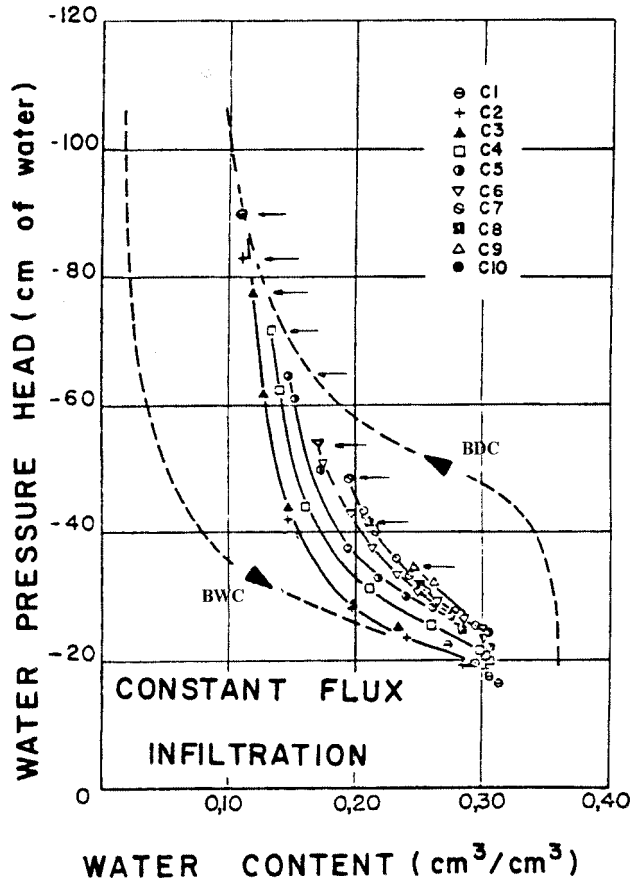
FIGURE 5.5 Hysteretic water retention curves observed during constant flux infiltration. The boundary wetting and drying curves are represented by the dotted lines. The measurements show the primary and secondary wetting curves. (Adapted from Vachaud, G. and Thony, J. L. 1971. Hysteresis during infiltration and redistribution in a soil column at different initial water contents. *Water Resour. Res.* 7: 111-127.)

5.2.4 Functional Relationships

In the past, many different functional relations have been proposed in the literature based on various combinations of the dependent variables θ, h and K, and a certain number of fitting parameters (e.g., Gardner, 1958; Brooks and Corey, 1964; Brutsaert, 1966; van Genuchten, 1980; Haverkamp and Vauclin, 1981). Independent of the correct physical meaning of the fitting parameters, their values are submitted to constraints imposed by the use of the transfer equations such as the Fokker Planck and/or Richards equation (see Section 5.4 of this chapter). The most frequently used in the literature are the following water retention and hydraulic conductivity expressions:

Water retention curves: The Brooks and Corey (1964) equation:

$$\begin{cases} \dfrac{\theta-\theta_r}{\theta_s-\theta_r} = \left[\dfrac{h_{bc}}{h}\right]^{\lambda} & \text{for} \quad h \le h_{bc} \\ \theta = \theta_s & \text{for} \quad h_{bc} \le h \le 0 \end{cases} \qquad (17)$$

and the van Genuchten (1980) water retention equation:

$$\frac{\theta - \theta_r}{\theta_s - \theta_r} = \left[1 + \left(\frac{h}{h_g}\right)^n\right]^{-m} \tag{18}$$

where the water pressure head (h) is usually taken as negative and expressed in cm of water; h_{bc} is the Brooks and Corey *pressure scale parameter*; h_g is the van Genuchten *pressure scale parameter*; and λ, m, and n are *water retention shape parameters*. The water retention shape parameters m and n are assumed to be linked by

$$m = 1 - \frac{k_m}{n} \quad \text{with} \quad n > k_m \tag{19}$$

where k_m is an integer value initially introduced by van Genuchten (1980) to calculate closed-form analytical expressions for the hydraulic conductivity function when substituted in the predictive conductivity models of Burdine (1953) or Mualem (1976) (see Equation [32]). For the Mualem theory, parameter k_m takes the value $k_m = 1$, and for the Burdine theory $k_m = 2$. For high pressure head values, the van Genuchten water retention equation (Equation [18]) behaves like the Brooks and Corey equation (Equation [17]) with $\lambda = m\, n$. However, it should be noted that this identity is only confirmed for soils with shape parameter values $m \leq 0.1$ (for the case where the Burdine condition is used: $k_m = 2$).

Hydraulic conductivity functions: The Brooks and Corey (1964) equation:

$$\frac{K}{K_s} = \left[\frac{\theta - \theta_r}{\theta_s - \theta_r}\right]^{\eta} \tag{20}$$

and the van Genuchten (1980) hydraulic conductivity equation:

$$\frac{K}{K_s} = \left[\frac{\theta - \theta_r}{\theta_s - \theta_r}\right]^{\frac{1}{2}} \left[1 - \left\{1 - \left(\frac{\theta - \theta_r}{\theta_s - \theta_r}\right)^{\frac{1}{m}}\right\}^m\right]^2 \tag{21}$$

where η is a *conductivity shape parameter*.

Through an extensive study, Fuentes et al. (1992) concluded that only the combination of the van Genuchten water retention equation (Equation [18]), $h(\theta)$, based on the Burdine theory ($m = 1 - 2/n$) together with the Brooks and Corey conductivity equation (Equation [20]) stays valid for all different types of soil encountered in practice without becoming inconsistent with the general water transfer theory. This is due to the rather limiting constraint which exists for shape parameter m when using the Mualem theory : $0.15 \leq m \leq 1$. Even though the residual water content (θ_r) has a well-defined physical meaning (see above in this section), the parameter θ_r which enters in Equations (17) to (21) is somewhat of a misnomer because it usually behaves as a pure fitting parameter without any physical meaning. For practical purposes, it can easily be set equal to zero (Kool et al., 1987).

For modeling purposes the soil characteristic equations are often expressed in dimensionless form (e.g., Haverkamp and Vauclin, 1979; Warrick et al., 1985). The following dimensionless soil variables are used:

$$\theta^* = \frac{\theta}{\theta_s}; \qquad h^* = \frac{h}{h_s}; \qquad \text{and} \qquad K^* = \frac{K}{K_s} \tag{22}$$

where the superscript * refers to the nondimensional form of the different variables.

The use of Equations (17) through (21) combined with appropriate initial and boundary conditions allows the description of water transfer in the vadose zone of soil in a fully deterministic way (e.g., in the form of Richards [1931] equation, see Section 5.4). However, such a comprehensive approach has its price in complexity, since it requires the determination of five unknown soil parameters:

- the two dimensionless shape parameters (m and η); and
- the three scale parameters θ_S, h_g, and K_S.

Both shape parameters are strongly linked to the textural soil properties, whereas the scale parameters are related to soil structure.

Before going to the following Section on the measurements of soil characteristics, it is important to underline two major difficulties in applying the foregoing formulism to field studies:

- The Equations (17) through (21) only refer to the capillary water retention and hydraulic conductivity soil properties. Side effects due to *macroporosity*, such as soil cracks, root holes, worm channels, and large pores formed for various biological or mechanical reasons, are not incorporated.
- The soil characteristic equations only give "point" scale information (i.e., *local scale*) with a critical scale area of 1 m². Consequently, when dealing with the modeling of soil moisture dynamics at field scales, there exists a clear mismatch between the scale at which the functional soil characteristics are defined (and measured!) and the characteristic spatial scale for which models are trying to make predictions.

5.2.5 Measurement of Soil Characteristics

Many experimental works have been devoted over the last decades to the development of measurement techniques for estimating soil hydraulic characteristics. Roughly speaking, two categories of methods can be distinguished for the determination of the unknown soil hydraulic parameters: (1) the measurement techniques (direct or indirect) and (2) the predictive methods. The first category of techniques has been developed mainly for the measurement of the soil hydraulic parameters at a local scale and is difficult to apply over large areas. The second category is more flexible and could possibly be used at larger scales. For that reason we have preferred to address the predictive methods separately in Section 5.2.6.

In general, the measurement techniques rely on precise and time-consuming experimental procedures that can be categorized as being either laboratory- or field-based. While laboratory methods allow accurate measurement of flow processes, they are performed on samples taken from the field, and as a result, their representativity of field conditions can be questioned. The presence of aggregates, stones, fissures, fractures, tension cracks, and root holes, commonly encountered in unsaturated soil profiles, is difficult to represent in small-scale laboratory samples. Field techniques can be more difficult to control, but they have the advantage of estimating *in situ* soil hydraulic properties that are more representative, which is of considerable value in the subsequent use of the hydraulic information. Therefore it is desirable to aim at field methods that can alleviate, to some extent, the time-consuming constraints.

Soil water content. Soil water content is determined using direct or indirect methods. Direct techniques, such as the gravimetric method, involve the measurement of water losses through the process of evaporation or drainage. Chemical reactions can also be used to displace water. Indirect methods, such as electrical or radiological techniques, estimate the physical properties of some factor known to have an influence on soil water content. Then through a calibration procedure, these properties are related to soil water content. A complete and detailed description of the different methods is given by Gardner (1986).

1. *Gravimetric method.* This method usually involves the weighing of a wet soil sample, the removal of water by evaporation through oven drying or microwave heating, and then the reweighing of the dry sample. The soil water content is determined as the ratio of the net weights of the water removed to the dry sample. A high accuracy in soil water content can be achieved by drying the

soil sample until a constant weight is reached. Even though it is a destructive and time-consuming approach, it is often used routinely because it does not require any calibration and it is easy to perform.

2. *Electrical resistance methods.* Electrical resistance methods are based on the strong correlation between the electrical resistance of a porous media and its water content. Two electrodes inserted in a porous block made of a variety of material resistant to degradation once buried (gypsum, nylon, fiber glass) are placed into the soil. Once equilibrium is reached (identical matric potential for the porous block and the soil), the water content of the porous block is an indicator of the soil water content. This nondestructive technique is affected by hysteresis; due to technical limitations, only a drying calibration curve can be determined. The use of such techniques is highly limited because of a reduced accuracy, especially for high water contents.

3. *Neutron Thermalization.* The basis of this method is the property of hydrogen atoms to slow down and scatter fast, high-energy colliding neutrons. The successive collisions result in a loss of energy until thermal equilibrium with the surrounding atoms is reached (*thermalization*). A source of fast neutrons is lowered in an access tube. Through collisions with the hydrogen atoms, these neutrons become thermalized. A counting system determines the concentration of thermal neutrons by calculating the number of particles emitted by the reaction of the low-energy neutrons and boron trifluoride present in a detector. Through calibration, the counts of the thermal neutrons are transformed to a water content. This method is nondestructive, but it requires calibration and its use is to be limited in the near future because of stricter environmental regulations. Another limiting factor is the volume of measurement and its dependency on water content. Detailed information on the error analysis in estimating soil water content from neutron probe measurements is given by Haverkamp et al. (1984) and Vauclin et al. (1984).

4. *Gamma rays attenuation method.* This nondestructive method consists of measuring the attenuation of intensity of a beam of gamma rays when crossing a soil column. The attenuation depends on the soil constituents and bulk density. Measuring the attenuation at two different gamma ray energies yields both the water content (θ) and dry bulk density (ρ_d). The apparatus includes a source of gamma rays and a scintillator detector. An accurate calibration is required.

5. *Capacitance method: time domain reflectometry (TDR) methods.* These methods are based on the measurement of the dielectric constant of the soil. Since the water has a very large dielectric constant compared to that of the solid phase and that of air, it has a large influence on the dielectric constant of the soil. TDR consists of measuring the transfer time of electromagnetic waves along parallel metallic rods (two or more) of known length inserted in the soil:

$$C_e = \left[\frac{c t_{tr}}{2l} \right]^2 \qquad (23)$$

where C_e is the relative dielectric constant of the soil; t_{tr} is the wave travel time from the entrance into the soil to the end of the rods; l is the length of the rods; and c is the light velocity. The next step is to calibrate the C_e versus a known value of volumetric water content. The technique is described in a comprehensive paper by Topp et al. (1980). Although the authors suggested that the procedure is insensitive to variations in dry bulk density, temperature, mineral composition, and salinity (i.e., one single calibration curve could be applied to nearly all soils), it is evident that the sensitivity of the measurements depends on the dielectric constant of the material between and/or around the wire probes (e.g., Knight, 1992). Thus, local nonuniformities due to small air gaps or material of different density can cause significant errors in the measured water content values. It complicates significantly the interpretation of the field water content measurements and often requires site-specific calibration curves. Haverkamp et al. (1997a) proposed a procedure to correct the apparent (i.e., TDR) soil water content values for the effects of air gaps, small stones, and water-holding coarse fragments. The advantage of the TDR method is that it is nondestructive.

Since the use of the neutron probe will be limited for environmental reasons, the TDR method is becoming more and more a standard.

6. *Remote sensing*. Remote sensing involves measurements, usually from satellites or airplanes, of the electromagnetic signature of a body or surface. For inferring soil moisture, measurements are usually done in the visible, infrared, and microwave spectra. Microwave sensors are of increasing interest since they measure the dielectric properties of a surface body without being affected by cloud cover. Two microwave sensors are used: passive sensors, also known as *radiometers*, which measure microwave emission, and active sensors or *radars*, which send a microwave signal and measure its reflection. One major advantage of remote sensing is that it integrates soil moisture over a certain area. However the measurements are limited to the very first centimeters of the soil and are affected by surface roughness and vegetation density. The calibration procedure is extremely difficult because of the temporal and spatial variability of ground measurement of soil moisture.

Soil water pressure head. Depending on the pressure range measured, two different devices are most widely used: tensiometers and psychrometers.

1. *Tensiometers*. A tensiometer is composed of a porous cup, a connecting tube or barrel, and a pressure measuring device that can be a water manometer, mercury manometer, pressure transducer, or a vacuum gauge. The porous cup is filled with degased water and has to remain saturated for the whole soil water pressure range. The porous cup and the surrounding soil are in hydrostatic equilibrium, yielding then a measure of the soil water pressure. The use of tensiometers is limited to pressure heads ranging from 0 to 800 cm. For higher pressures, the use of psychrometers is recommended. A detailed description of the different types of tensiometers is given by Cassel and Klute (1986).

2. *Psychrometer*. In this method, the soil water potential is related to the relative humidity defined by the ratio of the water vapor pressure in equilibrium with the liquid phase (p) to the saturated vapor pressure (p_0) as:

$$h = \frac{R_g T_l}{M_{mol}} \ln\left(\frac{p}{p_0}\right) \tag{24}$$

where M_{mol} is the molecular weight of water, R_g is the universal gas constant, and T_l is the temperature (Kelvin) of the liquid phase. Details about the thermocouple phsychrometer design and functioning are given by Rawlins and Campbell (1986).

3. *Other methods*. Additional methods for measuring soil water potential include *electric resistance sensors*, which can be gypsum, nylon or fiberglass. Once potential equilibrium is reached between the resistance sensor and the soil matrix, the electric conductivity is measured and then transformed to water potential through pre-established calibration curves. More methods, such as *heat dissipation*, are described by Campbell and Gee (1986).

Hydraulic conductivity. Hydraulic conductivity is a soil characteristic that cannot be measured directly. Unlike the water retention characteristic, which can be considered a quasistatic soil property, the hydraulic conductivity function is always related to water movement into or through the soil. The transient and steady movement of water in the unsaturated zone of soils depends to a large extent upon the pore network resulting from the assemblage of soil particles and aggregates. This dependence of $K(\theta)$ on the pore-size distribution underlines the complicity that exists between the water retention and hydraulic conductivity functions. It implies at the same time that the optimization of soil water characteristic parameters over measured field data should be carried out simultaneously with a combined objective function. As only very little information is available in the literature on this problem (e.g., Yates et al., 1992), the authors believe that combined optimization of water retention and hydraulic conductivity parameters is unfortunately seldom used in practice.

In the following, three methods developed for *in situ* determination of hydraulic soil properties are presented. The *instantaneous profile method* must be considered as a fixed ground based experiment, while the two other methods, *tension disk infiltrometer* and *pressure ring infiltrometer*, can be looked upon as mobile experiments. The two latter methods are based on inverse techniques. *Inverse methods* consist of procedures that use solutions for a flow process inversely against observations of that process. The hydraulic parameters are then estimated either via reorganization of the solution into an explicit expression, or numerical optimization. Numerical solutions usually have high computation overheads. A more attractive approach is to use analytical solutions which can yield better appreciation of the role of the different soil parameters interfering in the flow processes, and also require less computation. For a possible successful application of inverse methods, it is crucial to describe the flow process as precisely as possible and to control perfectly well the imposed initial and boundary conditions.

1. *Instantaneous profile method.* In field studies, the instantaneous profile method is a transient method used to observe the natural water movement into the soil. During the time interval $\Delta t = t_2 - t_1$, the change per unit surface in the total water storage of the vadose zone (ΔS), from the soil surface to the reference depth z_m (Figure 5.6), is calculated as:

$$\Delta S(z) = \int_0^{z_m} \left[\theta(z, t_2) - \theta(z, t_1) \right] dz \tag{25}$$
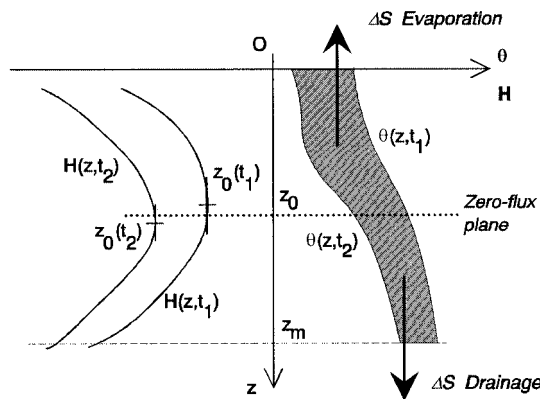
FIGURE 5.6 Instantaneous profile method to estimate unsaturated hydraulic conductivity.

Without any water supply to the soil profile during the period $[t_1, t_2]$, the storage variation corresponds to the water volume drained across the z_m plane or evaporated (i.e., evapotranspiration) through the plane $z = 0$, or even to a combination of both processes simultaneously. The flow direction can be determined from the gradient of the measured profiles of hydraulic head $H(z,t)$. A negative gradient, $dH/dz < 0$, indicates a positive, or downward-oriented, water flux in the Darcy's law (see also Section 5.3, Equation [38]). This flux is associated with percolation or drainage. A positive gradient, $dH/dz > 0$, corresponds to a negative or upward-oriented water flux (Figure 5.6), associated with evaporation. A *zero flux plane z_0* is defined when the hydraulic gradient $dH/dz = 0$. During a redistribution, the zero flux plane moves downward.

To determine the mean water flux per surface unit, at any soil depth, $q(z)$, it is possible to integrate the mass conservation equation from the zero flux plane z_0 to the chosen depth z as:

$$q(z) = \frac{\Delta S(z_0, z)}{t_2 - t_1} \tag{26}$$

where z_0 is the position of the zero flux plane averaged over the period $[t_1, t_2]$. The application of Darcy's law (Equation [38]) gives the hydraulic conductivity directly:

$$K(\theta) = -\frac{q(z)}{dH/dz} \tag{27}$$

The hydraulic gradient dH/dz in Equation (27) is the averaged value for the period Δt calculated at the depth z.

2. *Tension disk infiltrometer.* A representation of three tension disk infiltrometers with 250, 80, and 48 mm diameter bases (after the design of Perroux and White, 1988) is shown in Figure 5.7. A graduated reservoir tower provides the water supply; the bubble tower with a moveable air-entry tube (C_1) permits us to impose different boundary condition pressure head values at the cloth base (with a mesh of 20 µm). A thin layer of fine sand is placed over the soil surface to ensure a good hydraulic contact between the disk and the soil. *Cumulative infiltration*, $I(t)$, is recorded by measuring the water level drop in the reservoir tower (for the proper definition of cumulative infiltration, see Section 5.3). The transient infiltration flux is then given by the derivative: $q(t) = dI/dt$. By the use of an inverse procedure, these flux data allow the calculation of the spot values of hydraulic conductivity and sorptivity valid for the initial and boundary conditions chosen (for the proper definition of sorptivity see Section 5.4). Under field conditions, the contact layer of sand may be important. When this is the case, the water volume stored during the early stages of infiltration can be significant. Noting I_{sa} as the water volume necessary to fill the sand layer over the time period t_{sa}, the infiltration measurements have to be corrected by the coordinate transformation: $(I - I_{sa})$ and $(t - t_{sa})$.
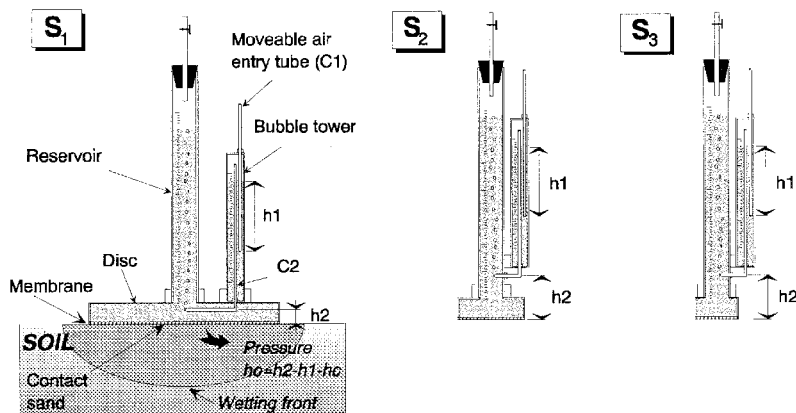
FIGURE 5.7 Tension disk infiltrometer. (Adapted from the design of Perroux, K. M., and White, I. 1988. Design of disc permeameters. *Soil Sci. Soc. Am. J.* 52: 1205-1215.)

In most cases slightly negative supply pressures are applied (Perroux and White, 1988), which allow the determination of hydraulic conductivity values close to saturation representative for the fine soil matrix without being biased by the possible influence of macroporosity. The fact that the initial and boundary conditions are well controlled makes the disk infiltrometer experiments particularly appropriate for data analysis through inverse procedures. The standard analysis uses Wooding's (1968) solution for three-dimensional steady-state infiltration valid for infinite time and uniform initial conditions. Unfortunately, these conditions are seldom met in the field. In order to overcome these limitations, a three-dimensional analytical solution of infiltration has been derived (Smettem et al., 1994; Haverkamp et al., 1994), which allows the description of transient three-dimensional infiltration behavior.

3. *Pressure ring infiltrometer.* The pressure ring infiltrometer is formed by a metallic ring that is driven into the soil to a given depth and connected to a reservoir system where the cumulative infiltrated volume could be measured for constant or falling hydraulic head. The reservoir system has to be adapted to the permeability characteristics of each soil. For soils with high permeability, water is supplied to the soil surface at a constant head (H_1) through the sealed top lid of a Mariotte bottle, with a moveable air tube allowing a wide range of different head values H_1 to be applied. For soils with low permeability, water must be supplied from a small capillary tube also acting as a measuring burette. This tube can be positioned either horizontally for water to infiltrate at a constant head, or vertically for water to infiltrate at a continually falling head. The cumulative infiltration, $I(t)$, is obtained from readings of the water supply tube. The saturated hydraulic conductivity is then deduced from the cumulative infiltration measurements by the use of analytical infiltration solutions developed for one-dimensional positive head infiltration subject to uniform initial conditions (Elrick et al., 1995). As these conditions are rarely met under field conditions, the results should be interpreted with caution. The initial condition is determined from undisturbed soil water content samples taken near the single ring.

4. *Other methods.* Over the last decade the *multistep outflow* method has gained some interest in the literature (e.g., Kool et al., 1987; Kool and Parker, 1988; Echings and Hopmans, 1993; van Dam et al., 1994). It concerns a laboratory-based method which allows the estimation of the soil characteristics by inverse modeling of a series of outflow experiments. For a successful application of the inverse optimization procedure it is crucial to ensure the uniqueness of the solution; consequently, various laboratory outflow experiments subject to different initial and/or boundary conditions are generally required. The multistep outflow method becomes more and more time-consuming as the number of unknown parameters to be determined increases. As for most laboratory methods, the representativity of the results for field conditions is questionable.

5.2.6 Estimation Techniques of Soil Characteristics

As shown in the previous section, direct measurements, either *in situ* or in laboratory, may be extremely time-consuming and expensive, in particular for large hydrological studies. Furthermore, it is extremely useful to have efficient methods for estimating soil transfer characteristics in areas where the amount of available information is limited. So many attempts were made at estimating soil characteristics from readily available data, such as *textural soil properties* (namely, particle-size distribution, and porosity), which are the most common measured soil data across the world. These relationships will be referred to hereafter as *pedotransfer functions*. This approach can be extremely powerful since it can either be used at the local scale using point textural properties or at the watershed scale, where textural information has been aggregated (soil maps). However, these pedotransfer functions should be used with caution, since they often rely on statistical regression equations which make them site (data) specific.

Water retention relation. In general three different approaches have been developed to predict soil water characteristics from the particle size distribution: (1) discrete matric potential regression methods, (2) functional regression methods, and (3) semiphysical approaches.

1. *Discrete matric potential regression methods.* In this approach, a multiple linear regression analysis is conducted to relate specific potential to particle-size distribution, porosity, organic matter content, and bulk density. These methods make no assumptions concerning the form of the soil water retention curve. Gupta and Larson (1979) developed regression equations between specific matric potential, particle-size distribution, and organic matter content. Rawls and Brakensiek (1982) developed three regression models. The first uses particle-size distribution, organic matter content, and bulk density as fitting variables. To improve these estimates, Rawls and Brakensiek (1982) introduced soil water content values at –1500 kPa and –33 kPa, respectively, in the second and third model. Tietje and Tapkenhinrichs (1993) as well as Vereecken et al. (1992), who tested the point matric potential approach extensively, concluded that the discrete methods often give poor results because the regression equations are usually based on measurements conducted on disturbed soil samples, limiting their representativeness and applicability.

2. *Functional regression methods.* In this approach, a preliminary shape of the soil water retention curve is assumed (e.g., the Brooks and Corey function), and its parameters are derived through fitting (e.g., Clapp and Hornberger, 1978; Bloemen, 1980; Vereecken et al., 1989; Wösten and van Genuchten, 1988). The method is more adapted for unsaturated flow modeling since it gives a continuous functional description of the water retention curve. McCuen et al. (1981) found that the mean and standard deviation of the Brooks and Corey parameters (i.e., λ and h_{bc} of Equation [17]) vary across soil textural classes. Cosby et al. (1984) extended this work and found that the textural soil properties can explain most of the variations of these parameters. This led to the development of regression equations between the mean and standard deviation of the soil hydraulic parameters and soil textural classes (Table 5.1). Rawls and Brakensiek (1985) developed a set of regression equations to estimate the water retention curve from particle-size distribution and porosity (Table 5.2). Notwithstanding that most of these empirical models clearly show correlation trends between water retention and textural/structural soil data, the application validity is in general restricted to the soils tested for each study, hence, limiting the transportability of the methods. The models present the advantage of easiness of use; however, when applied for predictions, they give hazardous results with important errors on water content, especially in the wet range of $h(\theta)$, which is mostly due to the fact that the models fail to predict correctly the soil structure related water content and pressure scale parameters (θ_S and h_{bc} or h_g).

TABLE 5.1 Regression Coefficients for the Brooks and Corey Parameters

| Coefficients | Mean (μ) | | | | Standard deviation (σ) | | | |
| | $\mu = c_1 + c_2 Cl + c_3 Sa + c_4 Si$ | | | | $\sigma = c_1 + c_2 Cl + c_3 Sa + c_4 Si$ | | | |
	log (h_{bc})	$1/\lambda$	θ_S	log (K_S)	log (h_{bc})	$1/\lambda$	θ_S	log (K_S)
c_1	1.54	3.10	50.5	−0.6000	0.72	0.92	8.23	0.43
c_2	0	0.157	−0.037	−0.0064	0.0012	0.0492	−0.0805	0.0011
c_3	−0.0095	−0.003	−0.142	0.0126	0	0	0	0
c_4	0.0063	0	0	0	−0.0026	0.0144	−0.0070	0.0032

From Cosby, B. J., Hornberger, G. M., Clapp, R. B., and Ginn, T. R. 1984. A statistical exploration of the relationship of soil moisture characteristics to the physical properties of soils. *Water Resour. Res.* 20: 682-690.
Cl, Si, and Sa represent the clay, silt, sand fractions (%), respectively.

3. *Semiphysical methods.* In spite of the quite unsatisfying results obtained by the functional regression methods (2) for the prediction of the water content (θ_S) and pressure (h_{bc} or h_g) scale parameters, these methods put in evidence the correlation trend between the shape parameters of the water retention and cumulative particle-size distribution functions. This *shape similarity* formed the main hypothesis for the few semiphysical models proposed in the literature (e.g., Arya and Paris, 1981; Haverkamp and Parlange, 1986). The model of Arya and Paris (1981), which is the most commonly used, involves dividing the cumulative particle-size distribution function into a number of fractions, assigning a pore volume and a volumetric water content to each fraction, and then computing a representative mean pore radius (R_i) and a corresponding water pressure head (h_i) value. They proposed the following nonlinear relationship to relate the pore radius (R_i) to the mean particle radius (D_{p_i}):

$$R_i = D_{p_i} \left[\frac{2enb_i^{1-\alpha}}{3} \right]^{0.5} \tag{28}$$

where e is the void ratio; nb_i is the number of particles in each size class; and α is an empirical factor referred to as the *tortuosity factor.* The value of α was fixed at $\alpha = 1.38$. Even though the parameter α seems to vary close to an average value for a large range of particle radii ($D_{p_i} < 100$ (μm), many soils,

TABLE 5.2 Regression Coefficients for the Brooks and Corey Parameters

	Function Type[1]			
	exp(f)	exp(f)	(f)	exp(f)
Regression Coefficients	h_{bc}	λ	θ_r	K_S
c_1	5.3396738	−0.7842831	−0.01824820	−8.968470
c_2	0.1845038	0	0.00513488	−0.028212
c_3	0	0.0177544	0.00087269	0
c_4	−2.48394546	−1.06249800	0.02939286	19.523480
c_5	0.00213853	−0.00273493	−0.00015395	−0.0094125
c_6	−0.61745089	0	0	0
c_7	0	−0.00005304	0	0.00018107
c_8	−0.04356349	−0.03088295	−0.00108270	0.077718
c_9	0	1.11134946	0	−8.395215
c_{10}	0.50028060	−0.00674491	−0.00235940	0
c_{11}	0.00000540	0	0	−0.0000035
c_{12}	0.00895359	0.00798746	0.00030703	0.0273300
c_{13}	−0.00001282	−0.00000235	0	0.0000173
c_{14}	−0.00855375	−0.00610522	−0.00018233	−0.0194920
c_{15}	−0.00072472	0	0	0.0014340
c_{16}	0	0	0	0
c_{17}	0.00143598	0.00026587	0	−0.00298

[1] Function $f = c_1 + c_2\,Cl + c_3\,Sa + c_4\,\varepsilon + c_5\,Cl^2 + c_6\,Cl\,\varepsilon + c_7\,Sa^2 + c_8\,Sa\,\varepsilon + c_9\,\varepsilon^2 + c_{10}\,Cl\,\varepsilon^2 + c_{11}\,Cl^2\,Sa + c_{12}\,Cl^2\,\varepsilon + c_{13}\,Cl\,Sa^2 + c_{14}\,Cl^2\,\varepsilon^2 + c_{15}\,Sa^2\,\varepsilon + c_{16}\,Cl^2\,Sa^2 + c_{17}\,Sa^2\,\varepsilon^2$. Cl represents the clay fraction (5% < Cl < 60%), Sa represents the sand fraction (5% < Sa < 70%), and ε is porosity.

From Rawls, W. J. and Brakensiek, D. L. 1985. Prediction of soil water properties for hydrologic modeling. Watershed Management in the Eighties. *Proc. Irrig. and Drain. Div.*, ASCE, Denver, Colorado

especially sandy soils, exhibit a large decline in the α value for the very fine particle class, and an α value larger than 2 for coarse sand particle radius ranges. The fluctuation of α is mainly due to the fact that the hypothesis of constant porosity (or void ratio) over the whole range of particle size classes is mostly not verified. Figure 5.8 shows the evolution of α as a function of θ_r/θ_S calculated for more than 1000 soils taken from the two soil databases UNSODA (Leij et al., 1996) and GRIZZLY (Haverkamp et al., 1997b). The second difficulty related to the use of the model of Arya and Paris (1981) is that hysteresis effects are not taken into account. The cumulative particle-size distribution is unique for a given soil and, hence, can only produce one associated water retention curve. The model has the advantage of easiness of use.

The second semiphysical model was proposed by Haverkamp and Parlange (1986). Their method allows the direct estimation of the parameters of Equation (17) for sandy soils without organic matter. The predicted $h(\theta)$ curve was then associated to the boundary wetting curve (BWC). Coupled with the hysteresis model proposed by Parlange (1976), the full family of wetting curves can be predicted. For the sake of simplicity, the authors assumed a linear relationship between the mean pore radius (R) and the corresponding particle radius (D_p):

$$R = \gamma D_p \tag{29}$$

where γ is an empirical packing factor to be determined through fitting from textural data. This method presents the advantage of interpreting the cumulative particle-size distribution function in its continuous form, however, it uses the extremely simple relationship between R and D_p (Equation [29]) which is perhaps valid for structureless soils such as pure sand soils, but which is certainly too crude for most field soils. Moreover, and similar to all the other methods presented before, it requires some statistical calibration for the estimation of γ.
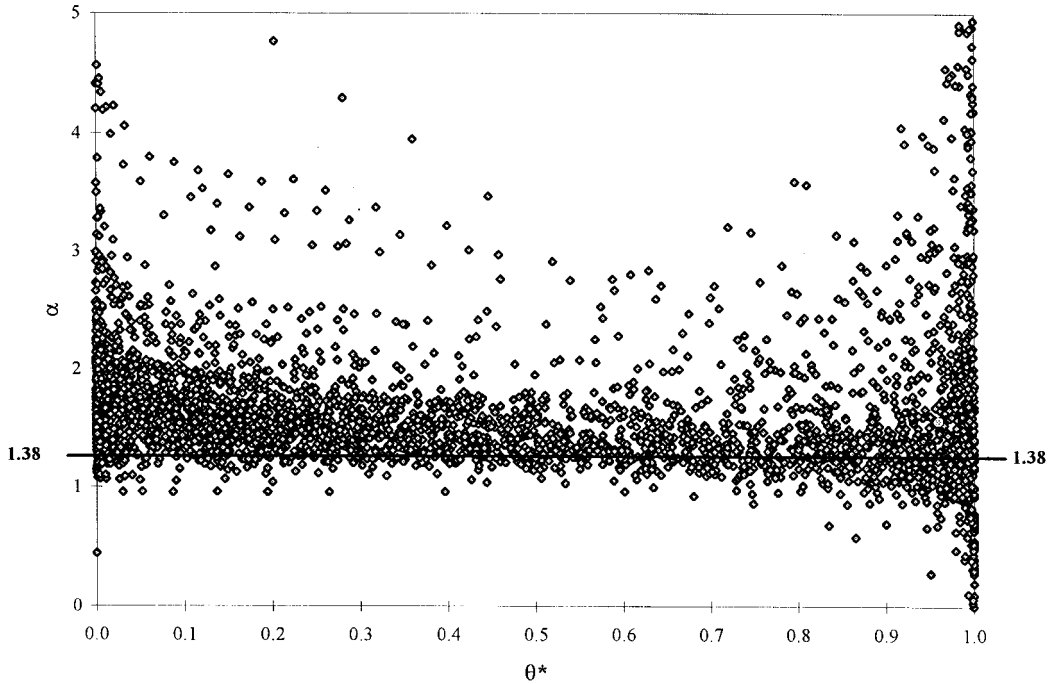
FIGURE 5.8 Evolution of the tortuosity factor α of the Arya and Paris (1981) model as a function of θ_i/θ_S calculated for more than 1000 different soils.

In a very recent study, Haverkamp et al. (1997c) proposed an improved physically based approach for the estimation of water retention curve parameters from textural soil properties. The method relies upon the concept of shape similarity and uses the method of geometrical scaling. The approach involves three complementary steps: the first concerns the link between the main wetting branch of the water retention curve and the cumulative pore-size distribution; the second step defines the relation between cumulative pore-size and particle-size distribution functions; and the third entails the problem imposed by hysteresis. The authors distinguish the difference between hydraulic pore radius and matric pore radius and take into account the effect of tortuosity. Choosing the van Genuchten (1980) type function (Equation [18]) to describe the soil water retention curve, the cumulative particle-size distribution function, $F(D_p)$, is written in the form:

$$F(D_p) = \left[1 + \left(\frac{D_g}{D_p} \right)^N \right]^{-M} \qquad \text{with} \qquad M = 1 - \frac{k_M}{N} \tag{30}$$

where D_g is the particle-size scale parameter; and M and N are the particle-size shape parameters linked to each other in a similar way as the shape parameters used for the water retention function (Equation [19]). However, the value of k_M is not obligatory equal to k_m. The ratio k_m/k_M is a function of tortuosity and porosity. As cumulative particle size data are easily accessible under field conditions, the values of D_g, M, and N are considered to be accessible from field measurements. The water retention shape parameters m and n are then calculated from M and N through:

$$M = m(1 + p\beta) \qquad \text{and} \qquad n = N \tag{31}$$

where $p\beta$ is a tortuosity factor defined as function of k_M. The water retention scale parameter h_g is calculated from D_g. A major advantage of the method is that it does not require calibration. The authors tested the model on two extensive independent soil databases (UNSODA (Leij et al., 1996) and GRIZZLY (Haverkamp et al., 1997b)) including more than 1000 soils. When compared with the other predictive models described in this section, the results obtained by this physically based model show significant improvement.

Table (5.3) is a summary of the mean soil characteristics found on the GRIZZLY database (Haverkamp et al., 1997b) when fitting the measured $h(\theta)$ points with the van Genuchten (1980) water retention equation (Equation [18]) subject to Burdine's condition ($k_m = 2$, Equation [19]). The values listed are quite data specific and vary when using other databases such as that of Rawls and Brakensiek (1985). However, an interesting point can be inferred from Table 5.3. It is obvious by looking at the clay fraction and the van Genuchten shape parameter (n) that the textural soil properties affect mainly the shape parameters, which agrees with the approach developed by Cosby et al. (1984), while the impact on the scale parameters (θ_S and h_g) is not as clear.

TABLE 5.3 Mean Soil Characteristic Parameters Found on the GRIZZLY Database Using van Genuchten Water Retention Equation (18) with $\theta_r = 0$ and Subject to Burdine's Condition (Equation [19] with $k_m = 2$)

Soil classes (USDA)	Clay fraction (%)	Sand fraction (%)	Porosity cm³/cm³	θ_s cm³/cm³	$|h_g|$ cm	n
Sand	0	93	0.39	0.38	26.95	2.94
Loamy sand	2	80	0.39	0.37	35.85	2.32
Sandy loam	8	65	0.40	0.37	77.87	2.35
Silt loam	17	25	0.47	0.44	162.40	2.23
Loam	18	39	0.44	0.42	141.41	2.21
Sandy clay loam	19	53	0.41	0.36	163.62	2.20
Clay loam	32	28	0.44	0.43	75.87	2.12
Silty clay loam	34	13	0.48	0.46	83.91	2.11
Silty clay	48	4	0.52	0.49	122.09	2.11
Clay	58	10	0.50	0.48	139.57	2.09

Hydraulic conductivity relation. Methods used to estimate the hydraulic conductivity curve rely on capillary model hypotheses. Most models proposed in the literature (e.g., Millington and Quirk, 1961; Mualem, 1976) are based on the Burdine's equation (1953) and make use of the water retention equation $h(\theta)$:

$$K^* = \left[\theta^*\right]^{p\beta} \left[\frac{\displaystyle\int_0^{\theta^*} \frac{1}{h\left(\bar{\theta}^*\right)^2} \, d\bar{\theta}^*}{\displaystyle\int_0^1 \frac{1}{h\left(\theta^*\right)^2} \, d\theta^*} \right] \tag{32}$$

where θ^* and K^* are the dimensionless water content and hydraulic conductivity, respectively, given by Equation (22); and $p\beta$ is a parameter generally associated with the effect of tortuosity. The differences between the various models arise from the hypotheses introduced to describe the pore structure (i.e., tortuosity) and its interaction with the relative permeability Equation [16]. By the use of Equation (32) the hydraulic conductivity curve shape parameter (i.e., η of Equation [20]) can be expressed as a function of the water retention shape parameters (i.e., λ or m of Equations [17] or [18], respectively) and the tortuosity factor:

$$\eta = \frac{2}{\lambda} + 2 + p\beta \tag{33}$$

Different values of $p\beta$ can be derived dependent on the capillary model chosen, i.e., Childs and Collis-George (1950) used $p\beta = 0$; Mualem (1976) $p\beta = 1/2$; Burdine (1953) $p\beta = 1$; and Millington and Quirk (1961) $p\beta = 4/3$. Although these $p\beta$-values have to be interpreted with considerable caution (as they were mostly based on pure intuition of the different authors combined with some subjective statistical analyses), they definitely have the benefit of indicating the existence of some complicity between the shape parameters of the water retention and hydraulic conductivity functions. The ideal case would be if the $p\beta$-value could be individualized for each soil.

The scale parameter K_S is strongly related to soil structure. Among the different soil hydraulic characteristic parameters the saturated hydraulic conductivity is the parameter which is the most influenced by effects such as macropores, stones, fissures, cracks, and other irregularities formed for various biological and mechanical reasons. Hence, it is the parameter which is the most difficult to predict. The models proposed in the literature either give estimations of the capillary conductivity value or are based on site- and soil-specific databases. The results should therefore be considered with caution when applied to field studies.

Mishra and Parker (1990) used the Mualem model (1976) with the van Genuchten water retention function (Equation [18]) to obtain a closed-form expression of the saturated hydraulic conductivity:

$$K_s = c_1 \frac{\left[\theta_s - \theta_r\right]^{2.5}}{h_g^2} \tag{34}$$

where c_1 is a constant including the effects of fluid characteristics and the porous media geometric factor; it has a value of 108 cm³/s when K_S is expressed in cm/s; h_g is the van Genuchten (1980) pressure scale parameter. The authors derived a similar predictive equation by the use of the Brooks and Corey (1964) water retention equation (Equation [17]):

$$K_s = c_1 \frac{\left[\theta_s - \theta_r\right]^{2.5}}{h_{bc}^2} \left[\frac{\lambda}{1+\lambda}\right]^2 \tag{35}$$

where λ and h_{bc} are the Brooks and Corey (1964) shape and scale parameters, respectively. Ahuja et al. (1985) used the general Kozeny–Carman approach to determine the saturated hydraulic conductivity from the effective porosity $(\varepsilon - \theta_r)$:

$$K_s = c_2 \left[\varepsilon - \theta_r\right]^{c_3} \tag{36}$$

where c_2 is equal to 1058 cm/h when K_S is expressed in cm/h; and c_3 takes a value of 4 or 5.

Cosby et al. (1984) derived a statistical regression between the saturated hydraulic conductivity and the clay, silt, and sand fractions of soil (Table 5.1). In a similar way, Rawls and Brakensiek (1989) presented a statistical model using particle-size distribution and porosity (Table 5.2).

5.2.7   Spatial Variability of Soil Water Properties

The difficulty of parametrization of soil moisture dynamics lies in the characterization and understanding of the different processes involved and the scale at which they take place: microscale, macroscale, and megascale. As underlined by Kabat et al. (1997), water transfer in the vadose zone is characterized by laminar flow with small mixing ratios, whereas atmospheric fluxes occur in the turbulent domain causing

large mixing ratios, resulting in different characteristic spatial scales. There is an urgent need to characterize and incorporate the spatial variability into current modeling approaches. The use of distributed models can only represent explicitly heterogeneities of processes with a characteristic length scale larger than the model grid size. This type of approach still uses some lumping at the grid level, and there remains the problem of characterizing the variability at the subgrid level.

Water movement in the unsaturated zone is affected by intrinsic parameters such as soil characteristics and by external factors such as rainfall and management practices. The movement is thus characterized by a high spatial variability. It is important, concerning the concept of spatial variability, to separate the concept of modeling of the unsaturated flow process and the parametrization of the soil hydraulic characteristics. In addition to the obvious lack of theoretical knowledge of the behavior of the flow equation in the unsaturated zone across temporal and spatial scales, the nonlinearity of the flow equations adds also to the complexity of representing spatial variability. The validity of extrapolating concepts developed at the local scale to larger scale may be questioned. A second major problem arises for the model parametrization. Is there a physical meaning for a characteristic curve representing 1 km² grid cell?

As mentioned earlier (Section 5.2.4, Equations [17] through [20]), characterizing soil properties requires the determination of five unknown parameters: two shape parameters (i.e., m and η, or λ and η) and three scale parameters (θ_S, K_S and h_g or h_{bc}). The shape parameters as shown by Haverkamp et al. (1997c) are strongly linked to soil texture, while the scale parameters depend mainly on the soil structural properties. Textural properties are characterized by a smaller underlying variability than structural properties. This is mainly due to the fact that anthropogenic factors, such as management practices, have a much larger impact on structural properties than on textural properties. Thus, it would be expected to be able to capture and determine the shape parameters λ, m, and η (*static characteristics*), across different scales by the use of pedotransfer functions. However, the scale parameters θ_S and K_S (*dynamic characteristics*) will unlikely be determined from textural information, especially for K_S, which is a driving parameter for modeling surface and subsurface flow. It is furthermore indispensable to note that soil characteristics have a meaning for capillary flow and that this concept will not hold in field conditions where macropore and preferential flow may be present and prevail over capillary flow. Thus the use of pedotransfer functions for estimating the dynamic parameters are doomed to fail. Haverkamp et al. (1996) proposed an alternate method to determine the scale parameters by coupling the soil characteristics with the unsaturated flow equation. This methodology, which will be presented in more detail in a later section, consists in the upscaling of the transfer equation and the derivation of scale-invariant soil parameters.

Another approach for incorporating spatial variability is the determination of effective parameters which are supposed to be valid over a representative elementary volume (e.g., Woods et al., 1995). However, because of the nonlinearity between soil water fluxes and the soil properties, and the extrapolation of theories developed at the local scale to large scale, the validity of the concept of effective parameters can be questioned.

5.3   Conceptual Aspects of Unsaturated Soil Water Flow

For the following analysis soil water movement is supposed to be isothermal and one-dimensional, whereas the influence of swelling and shrinking of the soil porous material is not taken into consideration. For a more detailed description of water movement in swelling soils we refer to work published by Smiles and Rosenthal (1968), Philip (1969a), Smiles (1974), Groenevelt and Parlange (1974), Sposito (1973, 1975a, b, c), Smiles (1995), and Gérard-Marchant et al. (1997). Moreover, the case of two-phase flow will not be addressed in this study. Under general field conditions, the effects due to air displacement can be considered negligible because of the large differences in dynamic viscosity. Only in particular situations (such as flooding, storm rains, or stratified soil profiles) may air be not free to escape, resulting in an increased air pressure ahead of the wetting front and, hence, in a reduced infiltration rate (see also Peck, [1965a, b]; Brustkern and Morel-Seytoux, [1970]; Phuc and Morel-Seytoux, [1972]; Vachaud et al. [1973]; Vachaud et al. [1974]; Starr et al. [1978]; and Touma et al. [1984]).

5.3.1 General Flow Equations

For one-dimensional vertical flow the continuity equation takes the form:

$$\frac{\partial \theta}{\partial t} = -\frac{\partial q}{\partial z} \tag{37}$$

where θ is volumetric soil water content; q is soil water flux; z is depth taken positive downward; and t is time. Combined with the generalized Darcy's law (1856):

$$q = -K(\theta)\frac{\partial H}{\partial z} \tag{38}$$

the soil water transfer equation can be written as:

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial z}\left[K(\theta)\left(\frac{\partial h}{\partial z} - 1\right)\right] \tag{39}$$

where h is soil water pressure head relative to atmospheric pressure ($h \leq 0$); K is hydraulic conductivity as a function of θ; and H is hydraulic head defined as:

$$H = h(\theta) - z \tag{40}$$

For nonswelling soils the hydraulic head (H) represents the energy of soil water per unit of weight at a given depth (z); the possible osmotic component which arises from the presence of solutes in the soil water is neglected (see also Section 5.2.3). Equation (39) can be expressed as a θ-dependent equation by introducing the concept of soil water diffusivity (Childs and Collis-George, 1950):

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial z}\left[D(\theta)\frac{\partial \theta}{\partial z} - K(\theta)\right] \tag{41}$$

where $D(\theta)$ is diffusivity defined by

$$D(\theta) = K(\theta)\frac{\mathrm{d}h}{\mathrm{d}\theta} \tag{42}$$

Equation (41) is generally known as the Fokker-Planck equation.

In a similar way Equation (39) can be expressed as a function of only soil water pressure head (h) by introducing the concept of specific capacity:

$$C(\theta)\frac{\partial h}{\partial t} = \frac{\partial}{\partial z}\left[K(\theta)\left(\frac{\partial h}{\partial z} - 1\right)\right] \tag{43}$$

where C is specific capacity given by $C(\theta) = d\theta/dh$. Equation (43) referred to as the Richards (1931) equation, is the one-dimensional, isothermal, unsaturated soil water transfer equation generally used in vadose zone hydrology, especially when coupled saturated and unsaturated flow problems are considered. The use of Equation (41) with volumetric water content (θ) as independent variable, without doubt

causes computational difficulties when applied to numerical simulation of water movement in regions close to saturation.

The solution of Equations (41) and/or (43) subject to given initial and boundary conditions describes the evolution of the water content profiles $\theta(z,t)$ as functions of depth and time. The initial condition imposed on $\theta(z,t)$ is given by:

$$\theta(z,0) = \theta_0(z) \tag{44}$$

where θ_0 is the initial water content value; and the upper boundary condition on $\theta(z,t)$ is either the *Dirichlet* concentration condition:

$$h(0,t) = h_1(t) \qquad \text{and/or} \qquad \theta(0,t) = \theta_1(t) \tag{45}$$

or the *Neumann* flux condition:

$$\left[K - K \frac{\partial h}{\partial z} \right]_{z=0} = q_1(t) \tag{46}$$

where h_1 is the surface soil water pressure, θ_1 is the corresponding volumetric soil water content, and q_1 is the flux at the soil surface. Both values of h_1 and q_1 may be positive, zero, or negative.

5.3.2 Infiltration

Infiltration concerns the physical process of water entry into the soil through the surface. It is governed by either the concentration type boundary condition (e.g., water ponding at the soil surface) or the flux concentration type surface boundary condition (e.g., rainfall). For a nonuniform initial water content profile, $\theta_0(z)$, the one-dimensional cumulative infiltration (I) expressed in volume per unit surface is defined by the integral of the flux at the soil surface:

$$I(t) = \int_0^t q_1(\bar{t}) \, d\bar{t} \tag{47}$$

For the particular case of infiltration under uniform initial conditions:

$$\theta(z,0) = \theta_0, \qquad t = 0 \qquad for \qquad z \geq 0 \tag{48}$$

and constant surface boundary conditions:

$$
\begin{aligned}
\theta(0,t) &= \theta_1, & z = 0 & \qquad for \qquad & t \geq 0 \\
h(0,t) &= h_1, & z = 0 & \qquad for \qquad & t \geq 0
\end{aligned}
\tag{49}
$$

the cumulative infiltration equation (Equation [47]) takes the form:

$$I(t) = K_0 t + \int_0^{z_f} \left[\theta(z,t) - \theta_0 \right] dz \tag{50}$$

where $z_f(t)$ is depth of the infiltration wetting front; and K_0 is the hydraulic conductivity at initial water content (θ_0). The constant water pressure value h_1 can be negative or positive equal to the ponded water depth (h_{surf}) imposed at the soil surface ($h_1 = h_{surf} \geq 0$).

Generally speaking, the time limits of cumulative infiltration are well defined:

$$\lim_{t \to 0} I(t) = 0 \qquad \text{and} \qquad \lim_{t \to \infty} I(t) = \infty \tag{51}$$

The infiltration rate, q_1 at the soil surface is defined as $q_1 = dI/dt$ with the time limits:

$$\lim_{t \to 0} q_1(t) = \infty \qquad \text{and} \qquad \lim_{t \to \infty} q_1(t) = K_1 \tag{52}$$

where K_1 is the hydraulic conductivity corresponding to the surface boundary condition $\theta = \theta_1$ and/or $h = h_1$; obviously, K_1 becomes equal to $K_1 = K_S$ for $\theta_1 = \theta_S$ and/or $h_1 \geq 0$.

5.4 Analytical Solutions of the Unsaturated Flow Equation

Although the solution of the Richards (and/or Fokker Planck) equation under general initial and boundary conditions can only be obtained through numerical simulation, exact analytical solutions under idealized conditions are now available. As most of the numerical codes are now available from reliable commercialized software packages (i.e., SWMS-2D, code for simulating water flow and solute transport in variably saturated media*), we rather prefer to explore in this Section some aspects of the analytical solutions. Good summaries of most of the numerical problems involved in solving Equation (43) can be found in Vauclin et al. (1979) and Haverkamp and Vauclin (1979 and 1981). They carried out a systematic comparative analysis of more than 40 numerical schemes based on the finite difference technique. Most of the classical solutions used for the problem of *linearization* (time) and *weighting* (space) were studied in terms of stability, convergence, and computational time. The results were compared to quasianalytical and experimental solutions. As an example, Figure 5.9 gives the comparison between cumulative infiltration obtained experimentally and calculated by the various methods of estimating finite difference interblock hydraulic conductivity values.

The mathematical difficulties associated with the analytical solutions for $\theta(z,t)$ and/or infiltration, can be best discussed through the initial and boundary conditions given by Equations (48) and (49). The constant pressure head (h_1) can be chosen negative, zero, or positive. In this Section we only consider the problem of constant pressure head infiltration. The flux condition is fully discussed in Chapter 6. The concept of sorptivity is addressed in some detail because it represents an important integral variable which links water retention and hydraulic conductivity characteristics. We then give a short discussion of two positive constant head infiltration equations required for the last section dealing with the spatial variability of water flow.

5.4.1 Constant Negative (or Zero) Pressure Head Condition at the Soil Surface

Different analytical solutions have been proposed in the literature (e.g., Philip, 1955 and 1957a; Parlange, 1971b and 1975; Philip and Knight, 1974; Babu, 1976; Parlange et al., 1982; and Broadbridge, 1990). A good summary can be found in Parlange et al. (1996).

Without gravity effects. In our discussion we first consider the case where gravity effects can be ignored (such as would be the case for horizontal infiltration). The water movement equation (Equation [41]) is then reduced to:

---

* US Salinity Laboratory, Agricultural Research Service, US Department of Agriculture, Riverside, Cal.
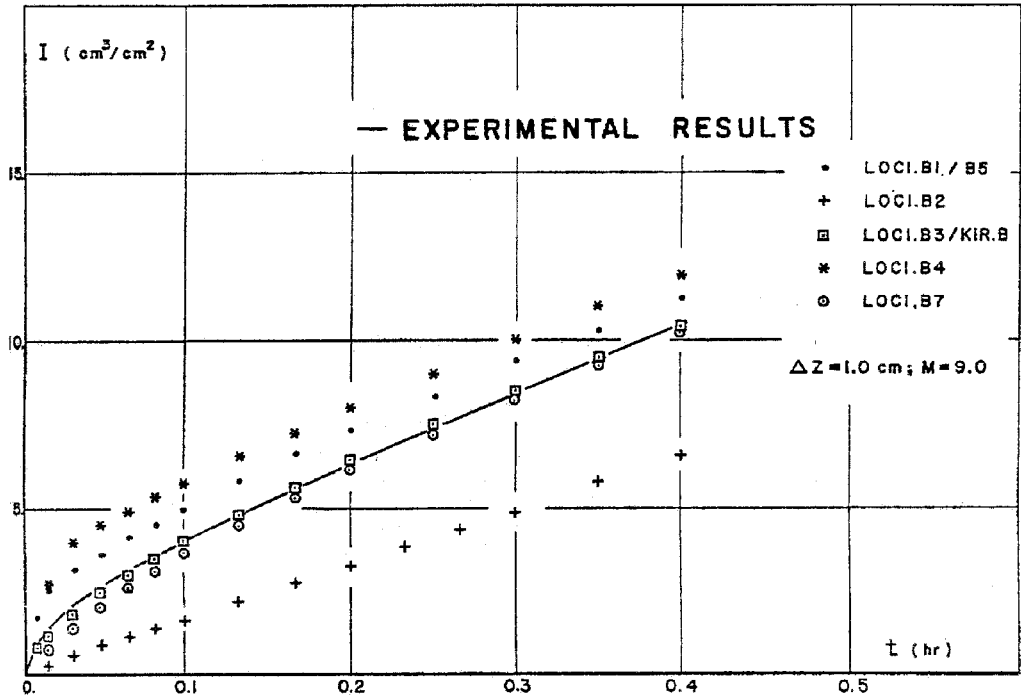
FIGURE 5.9 Comparison between cumulative infiltration data obtained experimentally and calculated by different numerical weighting modes of the interblock hydraulic conductivity values. (More details are given by Haverkamp and Vauclin, 1979).

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial z}\left[D(\theta)\frac{\partial \theta}{\partial z}\right] \qquad (53)$$

subject to the conditions:

$$\theta(z,0) = \theta_0, \qquad t = 0 \qquad for \qquad z \geq 0$$

$$\theta(0,t) = \theta_1, \qquad z = 0 \qquad for \qquad t \geq 0 \qquad (54)$$

The partial differential equation (Equation [53]) along with corresponding initial and boundary conditions can be reduced to an ordinary integral-differential equation by introducing the Boltzmann (1894) similarity variable χ_1 defined by $\chi_1(\theta) = z/t^{1/2}$:

$$\int_{\theta_0}^{\theta} \chi_1(\bar{\theta})\, d\bar{\theta} = -2D(\theta)\frac{d\theta}{d\chi_1(\theta)} \qquad (55)$$

subject to the condition:

$$\theta = \theta_1 \qquad for \qquad \chi_1(\theta_1) = 0 \qquad (56)$$

Equations (55) through (56) are generally known as the similarity equation of Bruce and Klute (1956). It shows that there exists a fundamental dynamic similarity between z and t associated with the solution of the diffusivity equation (Equation [55]); it reveals that the advance of the normalized wetting front is proportional to $t^{1/2}$.

Integrating the Boltzmann transform from θ_1 to θ_0, yields the cumulative infiltration equation:

$$I(t) = S_1(\theta_1, \theta_0)\sqrt{t} \tag{57}$$

where $S_1(\theta_1, \theta_0)$ is the *sorptivity* defined by:

$$S_1(\theta_1, \theta_0) = \int_{\theta_0}^{\theta_1} \chi_1(\theta)\, d\theta \tag{58}$$

The sorptivity S_1, which is clearly specific for the initial (θ_0) and boundary (θ_1) conditions encountered during each infiltration event, characterizes the ability of the soil to absorb water in the absence of gravity. It was initially presented by Green and Ampt (1911) and later coined by Philip (1957b).

Many solutions of Equation (55) have been proposed over time. One of the first reported in the literature is that of Crank and Henry (1949). However, their iterative method suffered from slow convergence and was often found slightly erroneous in the region close to initial water content θ_0. To overcome this problem Philip (1955) proposed a modified iterative procedure with the value of χ_1 close to θ_0 truncated at a fixed value calculated from an analytical inverse "erfc"-function. Even though the procedure of Philip improved the solution of Crank and Henry (1949) with respect to the convergence criterion, it still remained rather imprecise close to θ_0, and today this method is considered somewhat tedious and rather imprecise for the calculation of sorptivity.

A different solution of Equation (55) was proposed by Parlange (1971a) who started with a slightly modified version of Equation (53):

$$\frac{\partial z}{\partial t} + \frac{\partial}{\partial \theta}\left[D(\theta)\frac{\partial \theta}{\partial z}\right] = 0 \tag{59}$$

Integration of Equation (59) is carried out iteratively in two steps yielding the relation for $\chi_1(\theta)$ and hence $z(\theta, t)$ of the form:

$$\chi_1(\theta) = \frac{z(\theta, t)}{\sqrt{t}} = \int_{\theta_0}^{\theta_1} \frac{D(\bar{\theta})}{F_{PA}(\bar{\theta})}\, d\bar{\theta} \tag{60}$$

where $F_{PA}(\theta)$ is given by:

$$F_{PA}(\bar{\theta}) = \frac{\displaystyle\int_{\theta_0}^{\bar{\theta}}\left[\int_{\theta}^{\theta_1} D(\tilde{\theta})\, d\tilde{\theta}\right] d\hat{\theta}}{\left[2\displaystyle\int_{\theta_0}^{\theta_1}[\theta - \theta_0]D(\theta)\, d\theta\right]^{1/2}} \tag{61}$$

Parameters $\bar{\theta}$, $\hat{\theta}$, and $\tilde{\theta}$ are simply the integral variables. Equation (60) gives the moisture content profile $\chi_1(\theta)$ directly for any soil once $D(\theta)$ is known. Obviously, the approach can easily be extended to the case where gravity is not negligible (Parlange, 1971b).

In a later article Philip (1973) used this iterative concept by reformulating the F-function as:

$$F(\theta) = \frac{q_x(z,t)}{q_1(t)} = \frac{\displaystyle\int_{\theta_0}^{\theta} \chi_1(\bar{\theta}) \, d\bar{\theta}}{\displaystyle\int_{\theta_0}^{\theta_1} \chi_1(\theta) \, d\theta} \tag{62}$$

where $F(\theta)$ was referred to as the "flux-concentration" relation. Integration of Equation (59) after substitution of Equation (62) yields the well-known equation of $\chi_1(\theta)$:

$$\chi_1(\theta) = \frac{2}{S_1(\theta_1,\theta_0)} \int_{\theta}^{\theta_1} \frac{D(\bar{\theta})}{F(\bar{\theta})} \, d\bar{\theta} \tag{63}$$

The substitution of Equation (63) into Equation (58) allows us to redefine the sorptivity expression as:

$$S_1^2(\theta_1,\theta_0) = 2 \int_{\theta_0}^{\theta_1} \frac{[\theta - \theta_0] D(\theta)}{F(\theta)} \, d\theta \tag{64}$$

Probably the best understanding of the structure of any solution of Equation (55) can be based on the analytical expansion technique of Heaslet-Alksne (1961) which was applied by Parlange et al. (1992). A good compromise for the F-function, which balances accuracy with simplicity (Elrick and Robin, 1981), was given by Parlange (1975):

$$F(\theta) = \frac{2[\theta - \theta_0]}{[\theta_1 + \theta - \theta_0]} \tag{65}$$

resulting in the sorptivity equation used for most studies (e.g., Ross et al., 1996):

$$S_1^2(\theta_1,\theta_0) = \int_{\theta_0}^{\theta_1} [\theta_1 + \theta - 2\theta_0] D(\theta) \, d\theta \tag{66}$$

With gravity effects. For the case where the gravity effects cannot be neglected, the analytical solutions are mostly expressed through the Fokker-Planck equation (Equation [41]). The infiltration solution of Philip (1957a, b) which was originally developed for the surface boundary condition $h_1 = 0$ and $\theta_1 = \theta_S$, is written in terms of a series expansion in powers of $t^{1/2}$ as:

$$z(\theta,t) = \chi_1(\theta) t^{1/2} + \chi_2(\theta) t + \chi_3(\theta) t^{3/2} + \dots \tag{67}$$

where the first term, $\chi_1(\theta)$ embodies the influence of the capillary forces on the flow process, and the following terms, e.g., $\chi_2(\theta)$ and $\chi_3(\theta)$, reflect the gravity effect on infiltration. Integration of Equation (67) gives the cumulative infiltration equation, $I(t)$:

$$I(t) = S_1(\theta_1, \theta_0) t^{1/2} + \left[K_0 + S_2(\theta_1, \theta_0) \right] t + 0\left(t^{3/2}\right) + \dots \tag{68}$$

where the definition of $S_1(\theta_1, \theta_0)$ is given above by Equation (58) and:

$$S_2(\theta_1, \theta_0) = \int_{\theta_0}^{\theta_1} \chi_2(\theta)\, d\theta \tag{69}$$

Like the sorptivity ($S_1(\theta_1, \theta_0)$) the second term ($S_2(\theta_1, \theta_0)$) is specific for the initial and boundary conditions encountered during each particular infiltration event. The third and higher terms of Equation (68) can easily be calculated by transforming the time series expansion into a recurrence series as function of the first and second term (Haverkamp et al., 1990).

In theory, the Philip infiltration equation is physically based which makes it suitable for the simulation (*a posteriori*) of infiltration events once the soil characteristics are fully known. However, in practice, when the solution is used for prediction purposes (*a priori*) some ambiguity exists over the correct interpretation of Equation (68):

1. Though an accurate and simple estimation of sorptivity has been developed by Parlange (1975) in the form of Equation (66), no simple approximation exists for the second integral term. Generally, the value of $S_2(\theta_1, \theta_0)$ is associated to a constant A_2 calculated from curve fitting with:

$$I(t) = A_1 \sqrt{t} + \left[K_0 + A_2 \right] t \tag{70}$$

over experimental cumulative infiltration data where A_2 is confined over the interval (Youngs, 1968; Philip, 1969b; Talsma and Parlange, 1972):

$$\frac{K_1 - K_0}{3} \leq A_2 \leq 2 \frac{K_1 - K_0}{3} \tag{71}$$

As Equation (70) neglects deliberately all higher-order terms ($0(t^{3/2})$) of the original time series expansion (Equation [67]), it is evident that the optimized values of A_1 and A_2 reflect to some extent this truncation effect. Consequently, A_1 and A_2 are just specific *ad hoc* constants which become rather delicate to compare with the physically defined integral parameters $S_1(\theta_1, \theta_0)$ and $S_2(\theta_1, \theta_0)$ (Haverkamp et al., 1988).

2. Then, the definitions of $S_1(\theta_1, \theta_0)$ and $S_2(\theta_1, \theta_0)$ were originally developed for the case of infiltration with negative or zero head surface boundary condition ($h_1 \leq 0$ and $\theta_1 \leq \theta_S$). For the case of positive head infiltration ($h_1 \geq 0$ and $\theta_1 = \theta_S$) the integrals S_1 and S_2 have to be reformulated. This can easily be done for the sorptivity definition of S_1:

$$S_1^2(h_{surf}, h_0) = S_1^2(\theta_s, \theta_0) + 2 K_s h_{surf} \left[\theta_s - \theta_0 \right] \tag{72}$$

where h_0 stands for the initial soil water pressure $h(\theta_0)$ and h_{surf} is the ponded water depth imposed at the soil surface. However, for the second integral term (S_2) it is far more difficult to adjust Equation (69).

3. Finally, the main problem for the application of the Philip (1957a) solution lies in the fact that the time series of Equation (67) becomes divergent for large times no matter how many terms are developed. Thus, the solution is only valid for a limited time range. The time limit is mostly set at t_{grav}:

$$t_{grav} = \left[\frac{S_1(\theta_s, \theta_0)}{K_s - K_0}\right]^2 \tag{73}$$

a time for which the gravity forces are supposed to become predominant over the capillary forces during the infiltration process (Philip, 1969b).

In order to overcome this time limit problem Parlange et al. (1982) developed a quasi-exact infiltration solution of the Fokker-Planck equation (Equation [41]) valid over the entire time range $t \in [0,\infty]$. This work made use of a special double integration procedure of the water transfer equation. In later studies by Parlange et al. (1985) and Haverkamp et al. (1990), this infiltration solution which originally was developed for a zero and/or negative surface head condition, was extended to a positive head boundary condition ($h_1 \geq 0$). For the sake of brevity we will report more details on this solution in Section 5.4.2 dealing with positive head infiltration. To transform the positive head solution to a negative head solution the value of h_{surf} has simply to be set equal to zero and θ_S has to be replaced by θ_1 when necessary.

5.4.2 Constant Positive Pressure Head Condition at the Soil Surface

Before discussing the quasi-exact positive head solution of Parlange et al. (1985), it is preferred to address the somewhat simpler infiltration equation of Green and Ampt (1911); the latter solution which is widely used for hydrologic studies will be used in Section 5.5 of this chapter to illustrate the scaling principles.

The hypotheses underlying the Green and Ampt approach for positive head infiltration assume a nominal wetting front in the form of a step function with a one-point hydraulic conductivity function ($K(\theta) = K_s$) behind the wetting front and a constant driving pressure ($h_{surf} - h_f$) pulling the wetting front downward through the unsaturated zone; h_f is a constant equivalent water pressure head at the wetting front. The corresponding water retention curve takes the form of a step function with $\theta = \theta_0$ for $h \leq h_f$ and $\theta = \theta_S$ for $h_f \leq h \leq 0$. In the original infiltration equation of Green and Ampt (1911), the initial value of hydraulic conductivity (K_0) was considered equal to zero. However, for hydrologic model studies this condition is quite often not verified. Hence, we choose to present here the generalized Green and Ampt infiltration equation with $K_0 \neq 0$ (e.g., Schwartzendruber, 1987; Ross et al., 1996):

$$I(t) = K_s(t) + \frac{[\theta_s - \theta_0][h_{surf} - h_f(\theta_s, \theta_0)]K_s}{[K_s - K_0]} \ln\left\{1 + \frac{[I(t) - K_0 t][K_s - K_0]}{[\theta_s - \theta_0][h_{surf} - h_f(\theta_s, \theta_0)]K_s}\right\} \tag{74}$$

The corresponding infiltration flux (q_1) at the soil surface can be solved by:

$$t = \frac{[\theta_s - \theta_0][h_{surf} - h_f(\theta_s, \theta_0)]K_s}{[K_s - K_0]^2}\left\{\frac{K_s - K_0}{q_1 - K_s} - \ln\left[1 + \frac{K_s - K_0}{q_1 - K_s}\right]\right\} \tag{75}$$

Note that Equations (74) and (75) take the form of the classical Green and Ampt equations given in Chapter 6 (Equations [20] and [21]) when K_0 is set equal to zero. When nonpositive constant head infiltration is involved ($h_1 \leq 0$), Equations (74) and (75) can be applied equally well simply by replacing the values of θ_S and K_S by θ_1 and K_1 and setting h_{surf} equal to zero. Further details on the derivation of the Green and Ampt equation can be found in Chapter 6.

Although Equations (74) and (75) are exact solutions of the Richards equation (Equation [43]) for soils which correspond to the Green and Ampt configuration (hereafter referred to as "Green and Ampt" soils), these conditions are seldom met in the field. For that reason, emphasis is given in the literature to relating the value of h_f in some theoretical way to more realistic soil hydraulic properties (e.g., Bouwer, 1964; Mein and Larson, 1973; Neuman, 1976). Even though the value of h_f changes slightly with time

during infiltration (Haverkamp et al., 1988), a good estimation of h_f can be obtained from its short time limit (Neuman, 1976) derived from the second-order sorptivity definition as given by Parlange (1975):

$$h_f\left(\theta_1, \theta_0\right) = \frac{1}{2} \int_{h(\theta_1)}^{h(\theta_0)} \left[\frac{\theta_1 + \theta(h) - 2\theta_0}{\theta_1 - \theta_0} \right] \frac{K(h)}{K_1} \, dh \tag{76}$$

Parameter $h_f(\theta_s, \theta_0)$ must be seen as the pressure scale parameter of the Green and Ampt water retention curve. Equation (76) clearly shows the dependence of $h_f(\theta_1, \theta_0)$ on initial conditions (θ_0) and boundary conditions (θ_1) encountered during infiltration; hence, all estimations which relate $h_f(\theta_1, \theta_0)$ in a statistical way to purely textural soil data (such as is used by most pedotransfer functions [e.g., Rawls and Brakensiek, 1983]), lead necessarily to erroneous results.

As the integral parameter $h_f(\theta_1, \theta_0)$ is properly defined only for Green and Ampt soils, it is preferable to use the sorptivity concept instead (Equations [66] and [72]) when dealing with general field soils. For a negative head surface boundary condition ($h_1 \leq 0$) the relation between h_f and sorptivity is given by:

$$S_1^2\left(\theta_1, \theta_0\right) = 2K_s\left[\theta_1 - \theta_0\right]h_f\left(\theta_1, \theta_0\right) \tag{77}$$

and for the positive head surface boundary condition ($h_{surf} = h_1 > 0$) by:

$$S_1^2\left(h_{surf}, \theta_0\right) = 2K_s\left[h_{surf} - h_f\left(\theta_s, \theta_0\right)\right]\left[\theta_s - \theta_0\right] \tag{78}$$

Substitution of Equations (77) and (78) into Equations (74) and (75) allows the Green and Ampt infiltration equation to be expressed in terms of sorptivity.

Parlange et al. (1982) developed a quasi-exact infiltration solution of the Fokker-Planck equation (Equation [41]) valid over the entire time range $t \in [0,\infty]$. This work was concerned with a zero and/or negative surface head condition ($h_1 \leq 0$). In later studies by Parlange et al. (1985) and Haverkamp et al. (1990) this infiltration solution was extended to a positive head boundary condition ($h_1 = h_{surf} \geq 0$). Without going into too much detail we will report here a concise description of the approach.

Considering the case of infiltration into a soil at a uniform initial soil water pressure, $h_0 = h(\theta_0)$ (Equation [48]) from a surface boundary source at either a constant negative head ($h_1 \leq 0$) or a constant positive head ($h_1 = h_{surf} \geq 0$), double integration of Equation (43) allows the expression of the water content profiles, $z(\theta, t)$, in the rigorous form:

$$z\left(\theta, t\right) = z_s(t) + \int_\theta^{\theta_S} \frac{D(\bar{\theta})}{F(\bar{\theta}, t)\left[q_1(t) - K_0\right] - \left[K(\bar{\theta}) - K_0\right]} \, d\bar{\theta} \tag{79}$$

where $\bar{\theta}$ is the integration variable; z_s is the depth of soil considered to be saturated when $h_1 \geq 0$; and $F(\theta, t)$ is the flux concentration relation defined for diffusivity and gravity-driven infiltration (see also Equation [62]). Obviously, $z_s = 0$ and $\theta_S = \theta_1$ when $h_1 \leq 0$. Integrating z by parts from θ_0 to θ_S yields then:

$$\left[I(t) - K_0 t\right] = h_{surf} \frac{\left[\theta_s - \theta_0\right]K_s}{q_1(t) - K_s} + \int_{\theta_0}^{\theta_S} \frac{\left[\theta - \theta_0\right]D(\theta)}{F(\theta, t)\left[q_1(t) - K_0\right] - \left[K(\theta) - K_0\right]} \, d\theta \tag{80}$$

For the integration of Equation (80), Parlange et al. (1982) introduced an integral soil parameter β expressed as function of conductivity and diffusivity. The relation was later slightly generalized by Haverkamp et al. (1990) to:

$$\frac{\left[K(\theta)-K_0\right]}{\left[K_s-K_0\right]}=f(\theta)\left[1-\frac{2\beta(\theta_s,\theta_0)}{S_1^2(\theta_s,\theta_0)}\int_\theta^{\theta_S}\frac{\left[\bar{\theta}-\theta_0\right]}{f(\bar{\theta})}D(\bar{\theta})\,d\bar{\theta}\right] \tag{81}$$

where $f(\theta)$ is the purely diffusivity-driven flux concentration relation, $F(\theta,0)$, given by Equation (65). Obviously β is defined over the interval $0 \le \beta \le 1$ and is a function of the soil type and governing initial and boundary conditions. The lower limit $\beta = 0$ corresponds to the Green and Ampt type of soil in which $dK/d\theta$ increases much less rapidly with θ than the diffusivity, whereas the upper limit $\beta = 1$ corresponds to soils in which $dK/d\theta$ and the diffusivity behave similarly. Following Equation (81), parameter β is defined as a function of volumetric water content ($\theta_0 \le \theta \le \theta_S$); as a result, β is only influenced by the surface boundary condition when $\theta_1 \le \theta_S$ or $h_1 \le 0$. Ross et al. (1996) showed that β is slightly affected by changes in the surface boundary condition θ_1, especially when θ_1 stays close to θ_S ($\theta_1 \ge 0.75\ \theta_S$). The effect of the changing initial condition θ_0 is even more straightforward: obviously, β equals 1 for $\theta_0 = 0$ (Equation [81]); it remains close to 1 for $\theta_0 \le 0.25\ \theta_S$ and then decreases substantially at the initial water content $\theta_0 = 0.25\ \theta_S$ (Ross et al., 1996).

The use of Equation (81) allows the solution of the integral term of Equation (80), resulting in the following cumulative infiltration equation (Haverkamp et al., 1990):

$$I(q_1,t)=K_0 t+h_{surf}\left[\theta_s-\theta_0\right]\frac{K_s}{q_1-K_s}+\frac{S_1^2(\theta_s,\theta_0)}{2\left[K_s-K_0\right]}\ln\left[1+\frac{K_s-K_0}{q_1-K_s}\right] \tag{82}$$

However, the cumulative infiltration (I) can only be calculated from Equation (82) once the corresponding surface infiltration flux (q_1) is known. The derivation in time of Equation (82) yields the necessary flux/time equation, $q_1(t)$:

$$2\left[K_s-K_0\right]t=\frac{2h_{\text{surf}}K_s\left[\theta_s-\theta_0\right]}{q_1-K_s}+\frac{S_1^2(\theta_s,\theta_0)}{\beta[1-\beta]\left[K_s-K_0\right]}\ln\left[1+\beta\frac{K_s-K_0}{q_1-K_s}\right]$$

$$-\frac{S_1^2(\theta_s,\theta_0)+2[1-\beta]h_{\text{surf}}K_s\left[\theta_s-\theta_0\right]}{[1-\beta]\left[K_s-K_0\right]}\ln\left[1+\frac{K_s-K_0}{q_1-K_s}\right] \tag{83}$$

The infiltration solution given by the combined use of Equations (82) and (83) is valid over the entire time range $t \in [0,\infty]$ and is applicable for both positive negative head surface boundary conditions. Although this solution is proven to be extremely precise, it is also admittedly cumbersome to apply. For that reason Barry et al. (1995) presented an extension of foregoing infiltration solution which improved the applicability of the approach by transforming the implicit combination of Equations (82) and (83) into an explicit infiltration equation without affecting the precision.

5.5   Scaling Principles

In addition to the obvious lack of theoretical knowledge of the behavior of the flow equation in the unsaturated zone across temporal and spatial scales, the nonlinearity of the flow equations adds also to the complexity of representing spatial variability. There exists an acute disparity between the current theoretical knowledge, characterization, and measurement of water movement in the vadose zone and the characteristic spatial scales for which surface flux models such as SVATs and GCMs are trying to make predictions.

Hence, there is an urgent need to assess the adequacy and usefulness of scale matching between the different flow phenomena involved in the interactions between atmosphere, land surface, vadose zone,

and aquifer. The correct specification of soil hydrological processes at small or large scales is directly conditioned by the possibility of parameterizing soil water fluxes for the scales compatible with the grid size of field scales involved. The efficiency of soil characteristic parameterization at different scales depends on the clear definition of the functional relationships and parameters to be measured. The use of actual technologies to characterize surface fluxes gives only local information and provides only an indication of the underlying spatial variability of the different processes. For these multiple reasons, the input parameters used in different modeling approaches bear little resemblance to those measured. Spatial variability can be incorporated by the determination of effective parameters which are supposed to be valid over a representative elementary volume (e.g., Woods et al., 1995). Some of these methods are based on the use of pedotransfer functions which employ textural soil properties as principal input data. As mentioned earlier (Section 5.2.7), textural properties are characterized by a smaller underlying variability than the structural properties. The soil characteristic shape parameters (such as the water retention shape parameter m and the hydraulic conductivity shape parameter η, Equations [18] and [20]), which are strongly related to soil textural properties, could be determined from pedotransfer functions. However, the description of water flow in the unsaturated zone requires not just the knowledge of the shape parameters; information about the soil hydraulic scale parameters (such as the water content scale parameter θ_S, the water pressure scale parameter h_g, and the hydraulic conductivity scale parameter K_S) is even more important. The scale parameters which imbed the dynamical properties of soils cannot be determined just from soil textural properties. Any attempt to determine the soil scale parameters at different spatial grid scales has to incorporate, in one way or another, the flow behavior of the soil at the chosen grid scale. The nonlinearity between soil water fluxes and soil properties and the extrapolation of theories developed at the local scale to large scales introduce inherent limitations. Hence, methods which simply try to upscale the soil hydraulic characteristic scale parameters and then describe the flow behavior with those aggregated soil characteristics do not take into account these nonlinearities and will generally fail to predict the unsaturated flow behavior in a correct way. Because the knowledge of the unsaturated flow behavior of a soil is of major importance for vadose zone hydrology, rather than the knowledge of the hydraulic soil characteristics (which are nothing more than intermediate mathematical functional relationships used to calculate the flow behavior), it would be more obvious to upscale the transfer equation itself.

Basically, two categories of scaling techniques can be distinguished: physically based techniques such as *dimensional* and *inspectional* analyses, and empirically based techniques such as the *functional normalization* technique. A complete discussion of the different methods is given by Langhaar (1951) and Kline (1965). As stated by Tillitson and Nielsen (1984), the former category allows one "to convert a set of physically interrelated dimensional quantities into a set of nondimensional quantities which conserve the original interrelationship for a system that manifests geometric, kinematic, or dynamic similarity". The scaling factors obtained through such dimensional analysis have definite physical meaning in terms of the system being studied. On the other hand, the second category (i.e., functional normalization analyses) derives scaling factors relating properties of the two systems in some empirical way. Most of these methods (e.g., Miller and Miller 1955a, b and 1956; Reichardt et al., 1972; Simmons et al., 1979; Warrick et al., 1977; Warrick and Nielsen, 1980; Nielsen et al., 1983) attempt to coalesce sets of soil water characteristic curves into one reference curve on the base of simple geometric similarity and, hence, fail *a priori* to scale the vadose zone flow behavior in a general way. For the sake of generality, it is evident that the first category of dimensional and/or inspectional analyses is the most desirable and appropriate for tackling the problem of scaling.

5.5.1   Scaling Infiltration

Dynamical analysis rests on the fundamental postulate that the *invariance* of a physical law under a series of scale transformations implies the invariance of all consequences of the law under the same transformations (Birkhoff, 1960). Because the solution of the cumulative infiltration problem is simply a consequence of the flow equation for a particular set of initial and boundary conditions, this postulate

guarantees that the scaling factors obtained for the cumulative infiltration law apply equally for the flow equation.

The following set of scale transformations is chosen:

$$z = \alpha_z z^*, \qquad t = \alpha_t t^*, \qquad \theta = \alpha_\theta \theta^*$$
$$h = \alpha_h h^*, \qquad K = \alpha_K K^*, \qquad C = \alpha_C C^*$$

(84)

where α_z, α_t, α_θ, α_h, α_K, and $\alpha_C = \alpha_\theta/\alpha_h$ are constant parameters which define, respectively, the nondimensional variables z^*, t^*, θ^*, h^*, K^*, and C^*. Substitution of these scaling factors into Equations (43), (44), (45), and (46) allows us to express the flow equation (e.g., Richards equation) as a nondimensional boundary-value problem. The particular values of the scaling factors α should reflect the effect of the initial and boundary conditions on the values of the independent variable.

The first step of inspectional analysis is to select a quasi-exact solution of the infiltration problem. The generalized positive head infiltration solution (Equations [82] and [83]) of Parlange et al., (1985a) is a quasi-exact one-dimensional analytical solution which is valid for any soil type and over the complete time range: $t \in [0, \infty]$. The obvious choice would be to apply the scaling analysis to this infiltration solution (see Haverkamp et al., 1996). However, for the sake of brevity we choose here to illustrate the principles of the dynamical scaling analysis on the Green and Ampt (1911) solution (Equations [74] and [75]) subject to initial and boundary conditions given by Equations (48) and (49) with $h_1 = h_{surf} > 0$. Even though it is perhaps the least realistic infiltration equation among the many different analytical solutions of one-dimensional infiltration presented in the literature, it is still an exact solution of the flow equation for the particular case where soils are considered to behave like Green and Ampt soils. The scaling factors which will be derived for the Green and Ampt equations (Equations [74] and [75]) apply equally well to the generalized infiltration solution of Parlange et al. (1985a), (Haverkamp et al., 1996).

The second step of inspectional analysis (Sposito and Jury, 1985) consists of trying to define the scaling factors α of Equation (84). To do so, the infiltration solution represented by Equations (74) and (75) is expressed in a dimensionless form through the use of the following dimensionless variables:

$$\left[I - K_0 t \right] = \frac{S_1^2 \left(h_{surf}, \theta_0 \right)}{c_p \left[K_s - K_0 \right]} \Delta I^*$$

(85)

$$t = \frac{S_1^2 \left(h_{surf}, \theta_0 \right)}{c_p \left[K_s - K_0 \right]^2} t^*$$

(86)

where ΔI^* is the dimensionless net cumulative infiltration; and c_p represents a proportionality constant which depends on the functional relationships chosen to describe the soil hydraulic characteristics. For the case of the Green and Ampt configuration where the soil characteristics are simply defined as functions of the soil structure-dependent scale parameters θ_S, h_f, and K_S, the value of c_p is equal to 2. From Equations (85) and (86), the dimensionless infiltration flux at the soil surface is defined as:

$$\Delta q_1^* = \frac{d\Delta I^*}{dt^*} = \frac{q_1 - K_0}{K_s - K_0}$$

(87)

Substitution of Equations (85) through (87) into Equations (74) and (75) allows one to express the Green and Ampt infiltration solution in the following dimensionless form:

$$\Delta I^* = t^* + \frac{c_p}{2} \ln\left[1 + \frac{2}{c_p} \Delta I^* \right] \tag{88}$$

and

$$\frac{2}{c_p} t^* = \frac{1}{\Delta q_1^* - 1} - \ln\left[1 + \frac{1}{\Delta q_1^* - 1} \right] \tag{89}$$

The simultaneous use of Equations (88) and (89) reduces the infiltration problem to a dimensionless boundary value problem valid for all soil types (obviously, still in the context of the Green and Ampt concept), initial and boundary conditions. This dimensionless solution allows the straightforward determination of the specific scaling factors for cumulative infiltration (α_I) and infiltration rate (α_q). Choosing the scaling factors α_I and α_q in a way similar to the general scale transformations defined previously for Equation (84):

$$\left[I - K_0 t \right] = \alpha_I \Delta I^* \qquad \text{and} \qquad \left[q_1 - K_0 \right] = \alpha_q \Delta q_1^* \tag{90}$$

the scaling factors α_I and α_t can directly be calculated from Equations (85) and (86):

$$\alpha_I = \frac{S_1^2\left(h_{\text{surf}}, \theta_0 \right)}{c_p \left[K_s - K_0 \right]} \tag{91}$$

$$\alpha_t = \frac{S_1^2\left(h_{\text{surf}}, \theta_0 \right)}{c_p \left[K_s - K_0 \right]^2} \tag{92}$$

The infiltration rate scaling factor α_q can then be determined from Equation (87):

$$\alpha_q = \frac{\alpha_I}{\alpha_t} = \left[K_s - K_0 \right] \tag{93}$$

The key point illustrated by the foregoing analysis is that for any Green and Ampt soil the infiltration behavior (cumulative and/or infiltration flux) is fully determined by only two scaling factors: α_I (or α_q) and α_t. Both parameters embody the effect of soil type, initial and boundary conditions. The dimensionless invariant cumulative infiltration curve ($\Delta I^*(t^*)$) fixes the only *dynamical similarity class* for which the physical system can be said to be macroscopically similar.

So far, the scaling factors α_I and α_q still exhibit the specific character related to the type of flow problem analyzed (cumulative constant head infiltration). To bypass this specificity it is necessary to go down one step in the hierarchy of the water movement problematic by establishing the set of basic scaling factors which reduce the Richards equation ($\theta(z,t)$) rather than those which reduce the cumulative infiltration equation $I(t)$. This can be pursued by decomposing the specific infiltration scaling factors α_I and α_q into the set of scaling factors α_θ, α_h, α_K, and α_z. The determination of the conductivity scaling factor α_K is straightforward:

$$\alpha_K = \alpha_q = \left[K_s - K_0 \right] \tag{94}$$

The introduction of the scaling factors α_t and α_K into the Richards equation (Equation [43]) gives the identities:

$$\alpha_h = \alpha_z \qquad \text{and} \qquad \alpha_I = \alpha_\theta \alpha_h \tag{95}$$

By choosing the definition for the scaling factor α_θ as:

$$\alpha_\theta = \left[\theta_s - \theta_0 \right] \tag{96}$$

the pressure scaling factor (α_h) and thus the space scaling factor (α_z) are defined by:

$$\alpha_h = \alpha_z = \frac{S_1^2\left(h_{\text{surf}}, \theta_0 \right)}{c_p \left[\theta_s - \theta_0 \right]\left[K_s - K_0 \right]} \tag{97}$$

As the time and space scaling factors (i.e., α_t and α_z) are fully determined by the soil scaling factors α_θ, α_h, and α_K, the soil water transfer equation can be transformed simply into a nondimensional form by the use of only three scaling factors; these factors take into account the effect of soil type, initial and boundary conditions giving the dimensionless form of the Richards equation (Equation [43]) the flexibility to encompass the water flow behavior of diverse physical systems through the adjustment of experimentally controllable parameters.

To make this inspectional analysis of the water transfer and/or infiltration equation useful for practical application, the soil scale factors should be expressed in terms of the classically used soil characteristic parameters contained in the expressions of $h(\theta)$ and $K(\theta)$ such as Equations (18) and (20). To do so, the scaling factors α_θ, α_h, and α_K have to be stripped of initial and boundary condition effects. Letting λ stand for the fraction of α which is purely related to the soil characteristic, the different soil scaling factors can be redefined as functions of the governing surface boundary condition ($h_1 = h_{\text{surf}}$):

$$\alpha_\theta = \lambda_\theta \left\{ \frac{\theta_s - \theta_0}{\theta_s} \right\} \qquad \text{with} \qquad \lambda_\theta = \theta_s \tag{98}$$

$$\alpha_K = \lambda_K \left\{ \frac{K_s - K_0}{K_s} \right\} \qquad \text{with} \qquad \lambda_K = K_s \tag{99}$$

and

$$\alpha_h = \lambda_h \left\{ \frac{S_1^2\left(h_{\text{surf}}, 0 \right)}{S_1^2\left(\theta_s, 0 \right)} \right\} \left\{ \frac{\theta_s}{\theta_s - \theta_0} \frac{K_s}{K_s - K_0} \frac{S_1^2\left(h_{\text{surf}}, \theta_0 \right)}{S_1^2\left(h_{\text{surf}}, 0 \right)} \right\} \tag{100}$$

with

$$\lambda_h = \frac{S_1^2\left(\theta_s, 0 \right)}{c_p \theta_s K_s} \tag{101}$$

For the definitions given above, the residual water content θ_r is chosen equal to zero (Kool et al., 1987). For the case of positive head infiltration, the effect of the surface boundary condition enters only through

the definition of the pressure scaling factor α_h. The second and third factors on the right-hand side of Equation (100) represent, respectively, the correction for the surface boundary condition effects and the initial condition effects.

Although the different sorptivity ratios which enter in the scaling factor α_h (Equation [100]) are well defined by Equations (66) and (72), it is preferable to simplify the definitions of α_h to a maximum by expressing the sorptivity ratios in terms of the scaling factors α_K and α_θ. A precise estimation of the sorptivity ratio $S_1^2(\theta_s, \theta_0)/S_1^2(\theta_s, 0)$ is given by:

$$\frac{S_1^2(\theta_s, \theta_0)}{S_1^2(\theta_s, 0)} \cong \frac{\theta_s - \theta_0}{\theta_s} \frac{K_s - K_0}{K_s} \tag{102}$$

Using Equations (72) and (102), the expression for the positive head pressure scale parameter (α_h) can be simplified considerably:

$$\alpha_h \cong \lambda_h + \frac{2}{c_p} \frac{\lambda_K}{\alpha_K} h_{\text{surf}} \tag{103}$$

The foregoing equations (Equations [98] through [103]) are valid not only for the Green and Ampt (1911) infiltration solution (Equations [74] and [75]), but for the generalized analytical infiltration solution of Parlange et al. (1985), (Equations [82] and [83]) as well. For the particular case of Green and Ampt soils, the scaling factors α_h and λ_h take the form:

$$\alpha_h = \lambda_h + \frac{K_s}{K_s - K_0} h_{\text{surf}} \qquad \text{with} \qquad \lambda_h = -h_f \tag{104}$$

where h_f is taken as a negative pressure head value.

The importance of this example to the Green and Ampt infiltration equation is that there exists a unique dynamical similarity in the behavior of unsaturated soil water movement in Green and Ampt soils when governed by the Richards equation. The infiltration behavior is defined by an invariant curve. However, when the scaling analysis is applied on the generalized constant head infiltration solution (Equations [82] and [83]) there exists, theoretically, no unique dynamical similarity in the soil water behavior (Haverkamp et al. 1996). Instead there is a multitude of dynamical similarity classes depending on the combination of soil type, initial and boundary conditions. In its most general form, the positive and/or negative head infiltration behavior is defined by three infiltration scaling factors embodying the effect of soil type, initial and boundary conditions. However, the nonuniqueness of the different similarity classes was found to be very small, implying that for practical purposes the hypothesis of unique dynamical similarity can be applied once the Green and Ampt equation is adapted to more general field conditions.

The purely soil-related scaling factors λ_θ, λ_h, and λ_K represent the soil-characteristic scale parameters used by the soil hydraulic characteristic curves such as Equations (17), (18), (20), and (21). Considering the water retention curve of Brooks and Corey (1964) substitution of λ_θ and λ_h into Equation (17) gives:

$$\begin{cases} \dfrac{\theta}{\lambda_\theta} = \left[\dfrac{h}{\lambda_h} \right]^{-\lambda} & \text{for} \quad h \le \lambda_h \\ \theta = \lambda_\theta & \text{for} \quad \lambda_h \le h \le 0 \end{cases} \tag{105}$$

and for the van Genuchten (1980) water retention curve:

$$\frac{\theta}{\lambda_\theta} = \left[1 + \left(\frac{h}{\lambda_h}\right)^n\right]^{-m} \tag{106}$$

Substitution of λ_θ and λ_K in the hydraulic conductivity equation (Equation [20]) of Brooks and Corey (1964) yields:

$$\frac{K}{\lambda_K} = \left[\frac{\theta}{\lambda_\theta}\right]^\eta \tag{107}$$

When describing unsaturated water flow subject to a given set of initial and boundary conditions, the water flow behavior of the soil should be independent of the choice of the soil hydraulic functional relationships. This can be guaranteed by applying the sorptivity criterion which imposes the sorptivity, $S_1(\theta_S, 0)$ (Equation [66]) calculated by the use of the redefined water retention and hydraulic conductivity equations such as is given by Equations (105), (106), and (107), and the sorptivity, $S_1(\theta_S, 0)$, expressed as function of λ_h (Equation [101]) to be equal. Establishing this equality yields the correct c_p value which is specific for the functional relationship of $h(\theta)$ chosen and guarantees the sorptivity criterion.

The values of c_p calculated for the Brooks and Corey (1964) and van Genuchten (1980) water retention equations are respectively:

$$c_p = \frac{2\lambda\eta[\lambda\eta - 1] + \lambda[2\lambda\eta - 1]}{[\lambda\eta - 1][\lambda\eta - 1 + \lambda]} \tag{108}$$

and

$$c_p = \Gamma\left(1 + \frac{1}{n}\right)\left\{\frac{\Gamma\left(m\eta - \frac{1}{n}\right)}{\Gamma(m\eta)} + \frac{\Gamma\left(m\eta + m - \frac{1}{n}\right)}{\Gamma(m\eta + m)}\right\} \tag{109}$$

where Γ stands for the usual gamma function. Obviously, for Green and Ampt soils with the sorptivity given by Equation (77), the particular value of c_p is equal to 2. When dealing with general field soils, c_p is defined purely by textural soil parameters (Equations [108] and [109]). The values of m and η or λ and η can be determined by the use of pedotransfer functions (see Section 5.2.6) allowing the values of c_p to be easily calculated.

The use of c_p together with the Green and Ampt equation (Equation [88]) has the advantage of being an invariant dimensionless infiltration curve valid for general field soils without being limited by the hypothesis of the Green and Ampt configuration. The fact that the proportionality constant c_p is pinned down on the sorptivity expression which is valid for realistic soil hydraulic characteristics relations (such as the van Genuchten and/or Brooks and Corey curves), guarantees a realistic description of the infiltration flow behavior as textural soil properties are taken into account. This result is illustrated by the dimensionless cumulative infiltration curves $(\Delta I^*(t^*))$ shown in Figure 5.10. The envelope of possible similarity classes bounded by the two invariant solutions of the classical Green and Ampt equation (c_p = 2) and the Talsma and Parlange (1972) equation (valid for the Gardner [1958] soil) are presented, together with the experimental constant head infiltration data given by Touma et al. (1984). The Green and Ampt (Equation [88]) with c_p calculated from equation Equation (109) gives good agreement with the measured infiltration data.
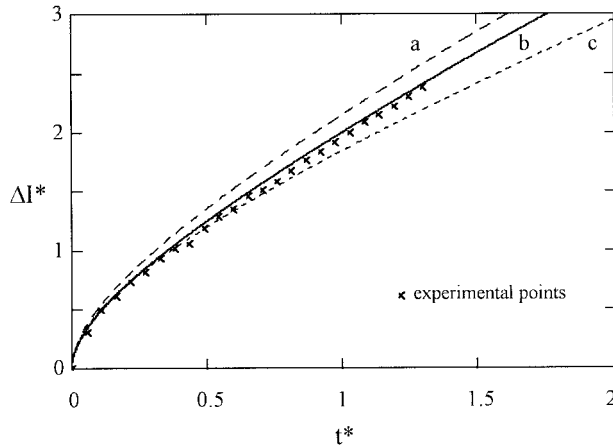
FIGURE 5.10 Experimental (Touma et al., 1984) and predicted dimensionless cumulative infiltration curves: (a) the classical Green and Ampt (1911) equation, (b) the modified Green and Ampt equation, and (c) the Talsma and Parlange (1972) equation.

The experimental cumulative infiltration data together with the surface boundary condition (h_{surf}) are the only data supposed to be known from field measurements. The identification procedure aims to determine the scaling factors α_I and α_t from these input data. Generally speaking, the optimization of two parameters on a monotone increasing function (such as the cumulative infiltration curve) guarantees *a priori* the uniqueness of the solution. Once the values of α_I and α_t are determined, the purely soil-related scaling parameters λ_θ, λ_h, and λ_K can be calculated (details of the identification procedure are given by Haverkamp et al., 1996). Some of the results are illustrated in Figure 5.11. In the context of the European international project on Climate and Hydrological Interactions between the Vegetation, the Atmosphere, and the Land surface (ECHIVAL) an infiltration field experiment was carried out in the desertification-threatened area of La Mancha (Spain). One hundred cumulative infiltration curves were measured over an area of 100 km² with a grid mesh of 1×1 km. The results are reported in Figure 5.11a. For each of the 100 curves the parameters α_I and α_t were determined by the use of the Green and Ampt equation (Equation [88]) with $c_p = 2$ and the Talsma solution (Talsma and Parlange, 1972). The scaled results, $I^*(t^*)$, are given in Figure 5.11b, showing the unique similarity which exists for both infiltration curves. All the 100 points fit each of the two infiltration curves perfectly well. It showed that with a minimum of measurement points for $I(t)$ precise information could be obtained on local soil hydraulic characteristics and their possible variations (upper and lower limits). Large areas were surveyed in short times. For example, the entire cumulative infiltration measurement campaign over the area of 100 km² did not take more than three weeks.

The foregoing scaling concept is valid for any spatial grid scale provided that suitable scaling factors can be measured and the soil is considered equivalent homogeneous for the grid scale involved. The concept can be used for any flow equation. Although the details of the approach were discussed through the infiltration law as flow equation, there are inherent limitations for using the infiltration phenomenon. The characteristic grid size of infiltration is still rather small (in the order of 1 m²). When dealing with hydrologic model studies the ideal choice would be to use the evaporation flux as a flow phenomenon since its characteristic grid size is much larger (in the order of 1 km²).

5.5.2 Scaling Evaporation

An identical approach to that developed previously can be applied for evaporation using Salvucci's (1997) approximate exfiltration equation based on moisture profile similarity:
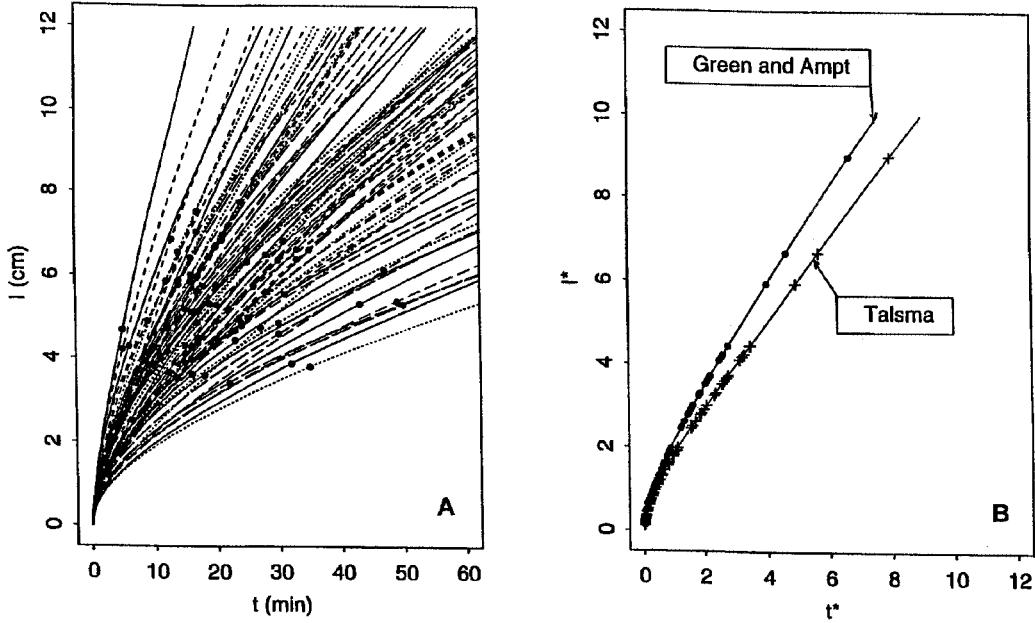
FIGURE 5.11 Fan of 100 different cumulative infiltration curves measured over an area of 100 km² before scaling (A) and after scaling (B). (Adapted from Haverkamp, R., Parlange, J.-Y., Cuenca, R., Ross, P. J., and Steenhuis, T. 1996. On the theory of scaling of the vadose zone water flow equation, in *Scale Invariance and Scale Dependence in Hydrology*, G. Sposito (Ed.). Cambridge University Press, New York, (in press).)

$$\left[K(\theta_0)\right]^2 t = \frac{S_d(\theta_1,\theta_0)}{2}\left[\frac{K(\theta_0)}{f_e} - \ln\left(1 + \frac{K(\theta_0)}{f_e}\right)\right] \tag{110}$$

subject to initial and boundary conditions:

$$
\begin{aligned}
\theta &= \theta_0 \quad && \text{at} \quad t = 0, && \text{and} \quad z \geq 0 \\
\theta &= \theta_1 \quad && \text{at} \quad z = 0, && \text{and} \quad t > 0
\end{aligned}
\tag{111}
$$

where by analogy with sorptivity $S_d(\theta_1,\theta_0)$ is defined as the *desorptivity* (Parlange et al., 1985b); and f_e is the exfiltration rate. Note that for evaporation $\theta_1 < \theta_0$. Even though Equation (110) does not assume a sharp water content front in the drying soil, it is similar to the Green and Ampt infiltration (1911) equation. Considering, for the sake of brevity, $K_0 = 0$ and $c_p = 2$, the following dimensionless variables can be defined:

$$f_e = \frac{K(\theta_0)}{f_e^*} \quad \text{and} \quad t = \frac{S_d^2(\theta_1,\theta_0)}{2\left[K(\theta_0)\right]^2} t^* \tag{112}$$

Equation (110) can then be written in the dimensionless form:

$$f_e^* = t^* + \ln\left[1 + f_e^*\right] \tag{113}$$

Because Equation (113) is identical to the Green and Ampt equation (Equation [88]) with $c_p = 2$, the scaling analysis explained previously can be applied to determine the invariant exfiltration curve. This approach is extremely interesting since there are actually methods, such as remote sensing, that provide areal averages of evaporation. It would then be possible from cumulative exfiltration (by analogy with cumulative infiltration) to evaluate the exfiltration scaling factors. Once the desorptivity is converted to sorptivity, the areal soil hydraulic scale parameters consistent with the flow process can be determined.

As new technologies such as Geographic Information Systems (GIS) and remote sensing are more and more becoming standard tools for evaluating flow processes at various scales, this scaling methodology proposes compatibility rules for combining different parameters and data sets, which is an aspect that is often neglected in GIS operations (Band and Moore, 1995).

5.6   Conclusion and Perspectives on Future Research

This chapter is mainly concerned with the water transfer in the unsaturated or vadose zone. The conceptual aspects of unsaturated soil water flow related to soil physical properties and soil water characteristics have been presented. Different functional relationships to represent soil hydraulic characteristics and different measurement and estimation techniques have been addressed. An overview of the most significant analytical solutions and the most important physical infiltration laws are given. It has been shown that the textural soil properties could give information regarding the shape parameters of the soil hydraulic characteristics, and the scale parameters which imbed the dynamic characteristics of the soil could only be determined if a flow equation is somehow considered. In this perspective, a new scaling approach of the water transfer equation in the unsaturated zone has been given.

This new theory is based on the scaling of the unsaturated flow equation rather than on the scaling of soil hydraulic characteristics. Following the postulate of invariance of a physical law and its consequences under a series of transformations, it is shown that in general there exists no unique dynamical similarity in the behavior of unsaturated soil water movement. A unique dynamical similarity is only found for two particular soil descriptions: the Green and Ampt (1911) soil and the Gardner (1958) soil. Both limit the envelope of possible similarity classes which exist for general field soils. However, when modifying the Green and Ampt configuration by taking into account the effects of soil texture, an invariant infiltration law can be determined which is realistic for general field soils.

This new approach offers a real alternative to classic estimation techniques of soil hydraulic characteristics, which are virtually impossible to perform over a great number of grid locations over a catchment, with a realistic use of resources in terms of time and costs. It is possible through quick infiltration experiments to determine the dynamic characteristics of soils which are far more representative and accurate than any estimate given from textural information such as pedotransfer functions. Furthermore, the effects of preferential flow can be taken into account. The second major advantage of the scaling approach is that it is compatible with the scale of measurement. Since this methodology can be applied to any flow equation such as evaporation, it is possible through the use of remote sensing, which provides realistic aggregated values of water fluxes, to tackle the mismatch that exists between the scale on which current knowledge of vadose zone water transfer processes is based and the scale for which the water budget models are making predictions.

For Further Information

Hillel (1980a, b) provided a comprehensive review on the fundamentals and applications of soil physics. It is well written and readers friendly.

Kutilek and Nielsen (1994) put together a comprehensive review on soil hydrology. The work is an excellent source of updated references.

Klute (1986) edited a monograph which provides a comprehensive review on methods of soil analysis.

Parlange et al. (1996) provided an excellent theoretical overview on infiltration and water movement in unsaturated soils.

Vauclin et al. (1979) provided a comprehensive study of the numerical solution of the infiltration equation.

Sposito (1996) put together all the latest developments about scaling and invariance in hydrology (in press).

References

Ahuja, L. R., Naney, J. W., and Williams, R. D. 1985. Estimating soil water characteristics from simpler properties or limited data. *Soil Sci. Soc. Am. J.* 49: 1100-1105.

Arya, L. M. and Paris, J. F. 1981. A physico-empirical model to predict the soil moisture characteristic from particle-size distribution and bulk density data. *Soil Sci. Soc. Am. J.* 45: 1023-1030.

Babu, D. K. 1976. Infiltration analysis and perturbation methods, 1. Absorption with exponential diffusivity. *Water Resour. Res.* 12(1): 89-93.

Band, L. E. and Moore, I. D. 1995. Scale: landscape attributes and geographical information systems. *Hydrol. Process.*, 9: 401-422.

Barry, D. A., Parlange, J.-Y., Haverkamp, R., and Ross, P. J. 1995. Infiltration under ponded conditions: 4. An explicit predictive infiltration formula. *Soil Sci.* 160: 8-17.

Bear, J. and Bachmat, Y. 1984. Transport phenomena in porous media — basic equations, in *Fundamentals of Transport in Porous Media,* Bear, J. and Corapocioglu, M. Y. (Eds.), Martinus Nijhoff: 3-61.

Birkhoff, G. 1960. *Hydrodynamics.* 2nd ed. Princeton University Press, Princeton, NJ.

Bloemen, G. W. 1980. Calculation of hydraulic conductivities of soils from texture and organic matter content. *Z. Pflanzenernährung und Bodenkunde.* 143(5): 581-615.

Bolt, G. H. 1976. Soil physics terminology. *International Society of Soil Science.* Bull. 49: 16-22.

Boltzmann, L. E. 1894. Zur Integration der Diffusionsgleichung bei variabeln Diffusions-coefficienten. *Ann. Physik* (Leipzig). 53: 959-964.

Bouwer, H. 1964. Unsaturated flow in groundwater hydraulics. *J. Hydraul. Div., Am. Soc. Civ. Eng.* 90HY5: 121-127.

Bouwer, H. 1978. *Groundwater Hydrology.* McGraw-Hill Book Company, New York.

Brooks, R. H. and Corey, C. T. 1964. Hydraulic properties of porous media. *Hydrol. Paper 3.* Colorado State University, Fort Collins.

Broadbridge, P. 1990. Solution of a nonlinear absorption model of mixed saturated-unsaturated flow. *Water Resour. Res.* 26: 2435-2444.

Bruce, R.R. and Klute, A. 1956. The measurement of soil water diffusivity. *Soil Sci. Soc. Am. Proc.* 20: 458-462.

Brustkern, R. L. and Morel-Seytoux, H. J. 1970. Analytical treatment of two-phase infiltration. *J. of the Hydraulics Division A.S.C.E.* 96 (12): 2535-2548.

Brutsaert, W. 1966. Probability laws for pore-size distributions. *Soil Sci.* 101: 85-92.

Buckingham, E. 1907. Studies on the movement of soil moisture. *U.S. Dept. Agr. Bur. Soils. Bull.*: 38.

Burdine, N. T. 1953. Relative permeability calculations from pore size distribution data. *Petr. Trans., Am. Inst. Mining Metall. Eng.* 198: 71-77.

Campbell, G. S. and Gee, G. W. 1986. Water potential: miscellaneous methods, in *Methods of Soil Analysis, Part 1, Physical and Mineralogical Methods*, A. Klute (Ed.), Agronomy Monograph 9, 2nd edition. Madison, WI.

Cassel, D. K. and Klute, A. 1986. Water potential: tensiometry, in *Methods of Soil Analysis, Part 1, Physical and Mineralogical Methods*, A. Klute (Ed.), Agronomy Monograph 9, 2nd edition. Madison, WI.

Childs, E. C. and Collis-George, C. 1950. The permeability of porous materials. *Proc. R. Soc. London.* Ser. A (201): 392-405.

Clapp, R. B. and Hornberger, G. M. 1978. Empirical equations for some hydraulic properties. *Water Resour. Res.* 14: 601-604.

Cosby, B. J., Hornberger, G. M., Clapp, R. B., and Ginn, T. R. 1984. A statistical exploration of the relationship of soil moisture characteristics to the physical properties of soils. *Water Resour. Res.* 20: 682-690.

Crank, J. and Henry, M. E. 1949. Diffusion in media with variable properties, II. *Trans. Faraday Soc.* 45: 1119-1128.

Darcy, H. 1856. Les fontaines publiques de la ville de Dijon. Dalmont. Paris.

Echings, S. O. and Hopmans, J. W. 1993. Optimization of hydraulic functions from transient outflow and soil water pressure data. *Soil Sci. Soc. Am. J.* 57: 1167-1175.

Elrick, D. E. and Robin, M. J. 1981. Estimating sorptivity of soils. *Soil Sci.* 132: 127-133.

Elrick, D. E., Parkin, G. W., Reynolds, W. D., and Fallow, D. J. 1995. Analysis of early time and steady-state single-ring infiltration under falling head conditions. *Water Resour. Res.* 31: 1883-1895.

Fuentes, C., Haverkamp, R., and Parlange, J.-Y. 1992. Parameter constraints on closed form soil water relationships. *J. Hydrol.* 134: 117-142.

Gardner, W. R. 1958. Some steady-state solutions of the unsaturated moisture flow equation with application to evaporation from a water table. *Soil Sci.* 85: 228-232.

Gardner, W. H. 1986. Water content, in *Methods of Soil Analysis, Part 1, Physical and Mineralogical Methods*, A. Klute (Ed.), Agronomy Monograph 9, 2nd edition. Madison, WI.

Gérard-Marchant, P., Angulo-Jaramillo, R., Haverkamp, R., Vauclin, M., Groenevelt, P., and Elrick, D. E. 1997. Estimating the hydraulic conductivity of slowly permeable and swelling materials from single ring experiments. *Water Resour. Res.* 33(6): 1375-1382.

Green, W. H. and Ampt, G. A. 1911. Studies in soil physic: I. The flow of air and water through soils. *J. Agric. Sci.* 4: 1-24.

Groenevelt, P. H. and Parlange, J.-Y. 1974. Thermodynamic stability of swelling soils. *Soil Sci.* 118: 1-5.

Gupta, S. C. and Larson, W. E. 1979. Estimating soil water retention characteristics from particle size distribution, organic matter percent and bulk density. *Water Resour. Res.* 15: 1633-1635.

Hamon, G. 1980. Mise en oeuvre et critique de méthodes de caractérisation hydrodynamique de la zone non-saturée du sol. Applications aux sols de culture du Sénégal. *Thèse Docteur-Ingénieur*, Université Scientifique et Médicale de Grenoble, France.

Haverkamp, R. and Vauclin, M. 1979. A note on estimating finite difference interblock hydraulic conductivity values for transient unsaturated flow problems. *Water Resour. Res.* 15(1): 181-187.

Haverkamp, R. and Vauclin, M. 1981. A comparative study of three forms of the Richards's equation used for predicting one-dimensional infiltration in unsaturated soils. *Soil Sci. Soc. Am. J.* 45: 13-20.

Haverkamp, R., Vauclin, M., and Vachaud, G. 1984. Error analysis in estimating soil water content from neutron probe measurements I. Local standpoint. *Soil Sci.* 137 (2): 78-90.

Haverkamp, R. and Parlange, J.-Y. 1986. Predicting the water retention curve from particle-size distribution: I. Sandy soils without organic matter. *Soil Sci.* 142: 325-339.

Haverkamp, R., Kutilek, M., Parlange, J.-Y., Rendon, L., and Krejca, M. 1988. Infiltration under ponded conditions: 2. Infiltration equations tested for parameter time dependence and predictive use. *Soil Sci.* 145: 317-329.

Haverkamp, R., Parlange, J.-Y., Starr, J. L., Schmitz, G., and Fuentes, C. 1990. Infiltration under ponded conditions: 3. A predictive equation based on physical parameters. *Soil Sci.* 149: 292-300.

Haverkamp, R., Ross, P. J., Smettem, K. R. J., and Parlange, J.-Y. 1994. Three-dimensional analysis of infiltration from the disc infiltrometer: 2. Physically based infiltration equation. *Water Resour. Res.* 30: 2931-2935.

Haverkamp, R., Parlange, J.-Y., Cuenca, R., Ross, P. J., and Steenhuis, T. 1998. On the theory of scaling of the vadose zone water flow equation, in *Scale Invariance and Scale Dependence in Hydrology*, G. Sposito (Ed.). Cambridge University Press, New York.

Haverkamp, R., Arrúe, J. L., Ross, P. J., Dalton, F. N., Connell, L. D., Lopez, M. V., Bouraoui, F., and Gracia, R. 1997a. Influence of stones and porous coarse fragments on water content and bulk density measurements at field scale. *Soil Sci.* (submitted).

Haverkamp, R., Zammit, C., Bouraoui, F., Rajkai, K., Arrúe, J. L., and Heckmann, N. 1997b. GRIZZLY, Grenoble Soil Catalogue: Soil survey of field data and description of particle size, soil water retention and hydraulic conductivity functions. Laboratoire d'Etude des Transferts en Hydrologie et Environnement (LTHE), Grenoble Cedex 9, France.

Haverkamp, R., Zammit, C., Bouraoui, F., Fuentes, C., Saâdi, Z., and Sander, G. 1997c. A physically-based approach for estimating water retention curve parameters. *Soil Sci. Soc. Am. J.* (submitted).

Heaslet, M. A. and Alksne, A. 1961. Diffusion from a fixed surface with a concentration-dependent coefficient. *J. Soc. Ind. Appl. Math.* 9: 584-596.

Hillel, D. 1980a. *Fundamentals of Soil Physics.* Academic Press, New York.

Hillel, D. 1980b. *Applications of Soil Physics.* Academic Press, New York.

Hogarth, W. L., Hopmans, J., Parlange, J.-Y., and Haverkamp, R. 1988. Application of a simple soil water hysteresis model. *J. Hydrol.* 98: 21-29.

Kabat, P., Hutjes, R. W. A., and Feddes, R. A. 1997. The scaling characteristics of soil parameters: from plotscale heterogeneity to subgrid parametrization. *J. Hydrol.* 190: 363-396.

Kline, S. J. 1965. *Similitude and Approximation Theory.* McGraw-Hill, New York.

Klute, A. 1986. *Methods of Soil Analysis. Part I — Physical and Mineralogical Methods.* Agronomy Monograph 9, 2nd edition. Madison, WI.

Knight, J. H. 1992. Sensitivity of time domain reflectometry measurements to lateral variations in soil water content. *Water Resour. Res.* 28: 2345-2352.

Kool, J. B., Parker, J. C., and van Genuchten, M. Th. 1987. Parameter estimation for unsaturated flow and transport models — A review. *J. Hydrol.* 91: 255-293.

Kool, J. B. and Parker, J. C. 1988. Analysis of the inverse problem for transient flow. *Water Resour. Res.* 24: 817-830.

Kutilek, M. and Nielsen, D. R. 1994. *Soil Hydrology.* Catena Verlag, Cremlingen-Desteds, Germany.

Lambe, T. W. and Whitman, R. V. 1979. *Soil Mechanics, SI Version.* John Wiley & Sons, New York.

Langhaar, H. L. 1951. Dimensional Analysis and Theory of Models. John Wiley & Sons Inc., New York.

Leij, F. J., Alves, W. J., van Genuchten, M. Th., and Williams, J. R. 1996. The UNSODA — Unsaturated Soil Hydraulic Database — User's Manual Version 1.0. Report EPA/600/R-96/095. National Risk Management Research Laboratory, Office of Research Development, U.S. Environmental Protection Agency, Cincinnati, Ohio: 1-103.

Liu, Y., Parlange, J.-Y., and Steenhuis, T. S. 1995. A soil water hysteresis model for fingered flow data. *Water Resour. Res.* 31(9): 2263-2266.

McCuen, R. H., Rawls, W. J., and Brakensiek, D. L. 1981. Statistical analysis of the Brooks and Corey and Green and Ampt parameters across soil texture. *Water Resour. Res.* 17: 1005-1013.

Mein, R. G. and Larson, C. L. 1973. Modeling infiltration during a steady rain. *Water Resour. Res.* 9: 384-394.

Miller, E. E. and Miller, R. D. 1955a. Theory of capillary flow: I. Practical implications. *Soil Sci. Soc. Am. J.* 19: 267-271.

Miller, E. E. and Miller, R. D. 1955b. Theory of capillary flow: II. Experimental information. *Soil Sci. Soc. Am. J.* 19: 271-275.

Miller, E. E. and Miller, R. D. 1956. Physical theory for capillary flow phenomena. *J. Appl. Phys.* 27: 324-332.

Millington, R. J. and Quirk, J. P. 1961. Permeability of porous solids. *Trans. Faraday Soc.* 57: 1200-1206.

Mishra, S. J. and Parker, J. C. 1990. On the relation between saturated hydraulic conductivity and capillary retention characteristics. *Ground Water.* 28: 775-777.

Mualem, Y. 1974. A conceptual model of hysteresis. *Water Resour. Res.* 10: 514-520.

Mualem, Y. 1976. A new model for predicting the hydraulic conductivity of unsaturated porous media. *Water Resour. Res.* 12: 513-522.

Neuman, S. P. 1976. Wetting front pressure head in the infiltration model of Green and Ampt. *Water Resour. Res.* 12(3): 564-566.

Nielsen, D. R., Tillotson, P. M., and Vieira, S. R. 1983. Analyzing field-measured soil-water properties. *Agric. Water Manage.* 6: 93-109.

Parlange, J.-Y. 1971a. Theory of water movement in soils: 1. One-dimensional absorption. *Soil Sci.* 111(2): 134-137.

Parlange, J.-Y. 1971b. Theory of water movement in soils: 2. One-dimensional infiltration. *Soil Sci.* 111(3): 170-174.

Parlange, J.-Y. 1975. On solving the flow equation in unsaturated soils by optimization: Horizontal infiltration. *Soil Sci. Soc. Am. Proc.* 39: 415-418.

Parlange, J.-Y. 1976. Capillary hysteresis and relationship between drying and wetting curves. *Water Resour. Res.* 12(2): 224-228.

Parlange, J.-Y. 1980. Water transport in soils. *Ann. Rev. Fluid Mech.* 12: 77-102.

Parlange, J.-Y., Lisle, I., Braddock, R. D., and Smith, R. E. 1982. The three parameter infiltration equation. *Soil Sci.* 133: 337-341.

Parlange, J.-Y., Haverkamp, R., and Touma, J. 1985. Infiltration under ponded conditions: 1. Optimal analytical solution and comparison with experimental observations. *Soil Sci.* 139: 305-311.

Parlange, J.-Y., Vauclin, M., Haverkamp, R., and Lisle, I. 1985. Note: the relation between desorptivity and soil-water diffusivity. *Soil Sci.* 139(5): 458-461.

Parlange, J.-Y., Steenhuis, T. S., Haverkamp, R., Barry, D. A., Culligan, P. J., Hogarth, W. L., Parlange, M. B., and Ross, P. 1998. Soil properties and water movement, in *Vadose Zone Hydrology, Cutting Across Disciplines*, M. B. Parlange and J. Hopmans (Eds.), Oxford University Press.

Parlange, M. B., Prasad, S. N., Parlange, J.-Y., and Römkens, M. J. M. 1992. Extension of the Heaslet-Alksne technique to arbitrary soil water diffusivities. *Water Resour. Res.* 28: 2793-2797.

Peck, A. J. 1965a. Moisture profile development and air compression during water uptake by bounded bodies. 2. Horizontal columns. *Soil Sci.* 99: 327-334.

Peck, A. J. 1965b. Moisture profile development and air compression during water uptake by bounded bodies. 3. Vertical columns. *Soil Sci.* 100: 44-51.

Perroux, K. M. and White, I. 1988. Design of disc permeameters. *Soil Sci. Soc. Am. J.* 52: 1205-1215.

Philip, J. R. 1955. Numerical solution of equations of the diffusion type with diffusivity concentration-dependent: 1. *Trans. Faraday Soc.* 51(7): 885-892.

Philip, J. R. 1957a. The theory of infiltration: 1. The infiltration equation and its solution. *Soil Sci.* 83: 345-357.

Philip, J. R. 1957b. The theory of infiltration: 4. Sorptivity and algebraic infiltration equations. *Soil Sci.* 84: 257-264.

Philip, J. R. 1969a. Hydrostatics and hydrodynamics in swelling soils. *Water Resour. Res.* 5: 1070-1077.

Philip, J. R. 1969b. The theory of infiltration. *Adv. Hydrosci.* 5: 215-305.

Philip, J. R. 1973. On solving the unsaturated flow equation: 1. The flux-concentration relation. *Soil Sci.* 116: 328-335.

Philip, J. R. and Knight, J. H. 1974. On solving the unsaturated flow equation: 3. New quasianalytical technique. *Soil Sci.* 117(1): 1-13.

Phuc Le Van and Morel-Seytoux, H. J. 1972. Effect of soil air movement and compressibility on infiltration rates. *Soil Sci. Soc. Am. Proc.* 36: 237-241.

Poulovassilis, A. 1962. Hysteresis in pore water and application of the concept of independent domains. *Soil Sci.* 93: 405-412.

Rawlins, S. L. and Campbell, G. S. 1986. Water potential: thermocouple psychrometry, in *Methods of Soil Analysis, Part 1, Physical and Mineralogical Methods.* A. Klute (Ed.), Agronomy Monograph 9, 2nd edition. Madison, WI.

Rawls, W. J. and Brakensiek, D. L. 1982. Estimating soil water retention from soil properties. *J. Irrig. and Drain. Div., ASCE.* 108 (IR2): 166-171.

Rawls, W. J. and Brakensiek, D. L. 1985. Prediction of soil water properties for hydrologic modeling. Watershed Management in the Eighties. *Proc. Irrig. and Drain. Div.*, ASCE, Denver, Colorado.

Rawls, W. J. and Brakensiek, D. L. 1989. Estimating soil hydraulic properties, in *Unsaturated Flow in Hydrologic Modeling*, H. J. Morel-Seytoux (Ed.), Kluwer Academic Publishers, published in cooperation with NATO Scientific Affairs Division, Boston.

Reichardt, K., Nielsen, D. R., and Biggar, J. W. 1972. Scaling of horizontal infiltration into homogeneous soils. *Soil Sci. Soc. Am. Proc.* 36: 241-245.

Richards, L. A. 1931. Capillary conduction of liquids through porous media. *Physics.* 1: 318-333.

Rijtema, P. E. 1965. An analysis of actual evapotranspiration. *Agric. Res. Rep.* 659, Center for Agricultural Publications and Documentation, Wageningen, the Netherlands.

Rogowski, A. S. 1971. Watershed physics: Model of soil moisture characteristics. *Water Resour. Res.* 7(6): 1575-1582.

Rose, C. W. 1966. *Agricultural Physics.* Pergamon Press, New York.

Ross, P. J., Haverkamp, R., and Parlange, J.-Y. 1996. Calculating parameters of infiltration equations from soil hydraulic functions. *Transport in Porous Media.* 24: 315-339.

Salvucci, G. D. 1997. Soil and moisture independent estimation of stage-two evaporation from potential evaporation and albedo or surface temperature. *Water Resour. Res.* 33(1): 111-122.

Simmons, C. R., Nielsen, D. R., and Biggar, J. W. 1979. Scaling of field-measured soil-water properties: I. Methodology. *Hilgardia.* 47: 77-102.

Smettem, K. R. J., Parlange, J.-Y., Ross, P. J., and Haverkamp, R. 1994. Three-dimensional analysis of infiltration from the disc infiltrometer: 1. A capillary-based theory. *Water Resour. Res.* 30: 2925-2929.

Smiles, D. E. and Rosenthal, M. J. 1968. The movement of water in swelling material. *Aust. J. Soil Res.* 6: 237-248.

Smiles, D. E. 1974. Infiltration into a swelling material. *Soil Sci.* 117: 140-147.

Smiles, D. E. 1995. Liquid flow in swelling soils. *Soil Sci. Soc. Am. J.* 59: 313-318.

Soil Survey Laboratory Staff. 1992. Soil survey laboratory methods manual. *Soil Surv. Invest. Rep. 42 Version 2.0.* USDA-SCS U.S. Gov. Print. Office, Washington, D.C.

Sposito, G. 1973. Volume changes in swelling clays. *Soil Sci.* 115: 315-320.

Sposito, G. 1975a. Steady vertical flows in swelling soils. *Water Resour. Res.* 11: 461-464.

Sposito, G. 1975b. A thermodynamic integral equation for the equilibrium moisture profile in swelling soils. *Water Resour. Res.* 11: 490-500.

Sposito, G. 1975c. On the differential equation for the equilibrium moisture profile in swelling soils. *Soil Sci. Soc. Am. Proc.* 39: 1053-1056.

Sposito, G. and Jury, W. A. 1985. Inspectional analysis in the theory of water flow through unsaturated soil. *Soil Sci. Soc. Am. J.* 49: 791-798.

Sposito, G. 1996. *Scale Invariance and Scale Dependence in Hydrology.* Cambridge University Press, New York. (in press).

Starr, J. L., de Roo, H. C., Frink, C. R., and Parlange, J.-Y. 1978. Leaching characteristics of a layered field soil. *Soil Sci. Soc. Am. J.* 42: 386-391.

Swartzendruber, D. 1987. Rigorous derivation and interpretation of the Green and Ampt equation, in *Infiltration Development and Application*, Y.-S. Fok (Ed.), Water Resources Research Center, University of Hawaii, Honolulu, Hawaii.

Talsma, T. and Parlange, J.-Y. 1972. One-dimensional vertical infiltration. *Austr. J. Soil Res.* 10: 143-150.

Tietje, O. and Tapkenhinrichs, M. 1993. Evaluation of pedo-transfer functions. *Soil Sci. Soc. Am. J.* 57: 1088-1095.

Tillotson, P. M. and Nielsen, D. R. 1984. Scale factors in soil science. *Soil Sci. Soc. Am. J.* 48: 953-959.

Topp, G.C. 1971. Soil-water hysteresis: the domain theory extended to pore interaction conditions. *Soil Sci. Soc. Am. Proc.* 35: 219-225.

Topp, G. C., Davis, J. L., and Annam, A. P. 1980. Electromagnetic determination of soil water content: measurement in coaxial transmission lines. *Water Resour. Res.* 16: 574-582.

Touma, J., Vachaud, G., and Parlange, J.-Y. 1984. Air and water flow in a sealed ponded vertical soil column. *Soil Sci.* 137: 181-187.

Vachaud, G. and Thony, J. L. 1971. Hysteresis during infiltration and redistribution in a soil column at different initial water contents. *Water Resour. Res.* 7: 111-127.

Vachaud, G., Vauclin, M., Khanji, D., and Wakil, N. 1973. Effects of air pressure on water flow in an unsaturated stratified vertical column of band. *Water Resour. Res.* 9: 160-173.

Vachaud, G., Gaudet, J. P., and Kuraz, V. 1974. Air and water flow during ponded infiltration in a vertical bounded column of soil. *J. Hydrol.* 22: 89-108.

van Bavel, C. H. M. 1969. The three phase domain in hydrology, in *Water in the Unsaturated Zone*. Rijtema, P. E. and H. Wassink (Eds.), UNESCO, Paris, France, Vol. 1: 23-32.

van Dam, J. C., Stricker, J. N. M., and Droogers, P. 1994. Inverse method to determine soil hydraulic functions from multistep outflow experiments. *Soil Sci. Soc. Am. J.* 58: 647-652.

van Genuchten, M. Th. 1980. A closed form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Sci. Soc. Am. J.* 44: 892-898.

Vauclin, M., Haverkamp, R., and Vachaud, G. 1979. *Résolution numérique d'une équation de diffusion non linéaire. Application à l'infiltration de l'eau dans les sols non saturés*. Presses Universitaires de Grenoble, Grenoble, France.

Vauclin, M., Haverkamp, R., and Vachaud, G. 1984. Error analysis in estimating soil water content from neutron probe measurements II. Spatial standpoint. *Soil Sci.* 137 (3): 141-148.

Vereecken, H., Maes, J., Feyen, J., and Darius, P. 1989. Estimating the soil moisture retention characteristic from texture, bulk density and carbon content. *Soil Sci.* 148: 389-403.

Vereecken, H., Diels, J., van Orshoven, J., Feyen, J., and Bouma, J. 1992. Functional evaluation of pedotransfer functions for the estimation of soil hydraulic properties. *Soil Sci. Soc. Am. J.* 56: 1371-1378.

Warrick, A. W., Mullen, G. J., and Nielsen, D. R. 1977. Scaling field-measured soil hydraulic properties using a similar media concept. *Water Resour. Res.* 13: 355-362.

Warrick, A. W. and Nielsen, D. R. 1980. Spatial variability of soil physical properties in the field, in *Applications of Soil Physics*, D. Hillel (Ed.), Academic Press, New York, 319-344.

Warrick, A. W., Lomen, D. O., and Yates, S. R. 1985. A generalized solution to infiltration. *Soil Sci. Soc. Am. J.* 49: 34-38.

Wösten, J. H. M. and van Genuchten, M. Th. 1988. Using texture and other soil properties to predict the unsaturated soil hydraulic functions. *Soil Sci. Soc. Am. J.* 52: 1762-1770.

Wooding, R. A. 1968. Steady infiltration from a shallow circular pond. *Water Resour. Res.* 4: 1259-1273.

Woods, R. A., Sivapalan, M., and Duncan, M. 1995. Investigating the representative area concept: an approach based on field data. *Hydrol. Process.* 9:291-312.

Yates, S. R., van Genuchten, M. Th., Warrick, A. W., and Leij, F. J. 1992. Analysis of measured, predicted, and estimated hydraulic conductivity using the RETC computer program. *Soil Sci. Soc. Am. J.* 56: 347-354.

Yong, R. N. and Warkentin, B. P. 1966. *Introduction to Soil Behavior*. Macmillan, New York.

Youngs, E. G. 1968. An estimation of sorptivity for infiltration studies from moisture moment considerations. *Soil Sci.* 106: 157-163.

Glossary and Parameter Notation

A_1 first term of Philip's truncated infiltration equation $(m/\sqrt{s})$

A_2 second term of Philip's truncated infiltration equation (m/s)

$c_1,\ldots,_n$ regression coefficients

c   light velocity (m/s)

C   specific capacity (1/m)

C_e dielectric constant

Cl   clay fraction (%)

D   diffusivity (m²/s)

Dg   particle size scale parameter (m)

D_p	particle radius of soil particle (m)
D_{p_i}	particle radius of size class "i" of the Arya and Paris (1981) model (m)
e	void ratio (m³/m³)
f_e	exfiltration rate (m/s)
f_e^*	dimensionless exfiltration rate
F	flux concentration relation
g	gravitational acceleration, 9.82 m/sec²
h	soil water pressure head (m)
h_0	soil water pressure head at initial water content (m)
h_1	surface soil water pressure head (m)
h_{bc}	Brooks and Corey (1964) soil water retention scale parameter (m)
h_g	van Genuchten (1980) soil water retention scale parameter (m)
h_f	Green and Ampt wetting front pressure head value (m)
h_{surf}	ponded water depth at soil surface (m)
h^*	dimensionless soil water pressure head
H	hydraulic head (m)
I	one-dimensional cumulative infiltration (m³/m²)
k	intrinsic permeability (m²)
k_m	parameter relating van Genuchten (1980) shape parameters m and n
k_M	parameter relating particle size shape parameters M and N
k_{rw}	relative water permeability
K	hydraulic conductivity (m/s)
K_S	hydraulic conductivity at natural saturation (m/s)
K_0	hydraulic conductivity at initial water content θ_0, (m/s)
K_1	hydraulic conductivity at water content θ_1, (m/s)
K^*	dimensionless hydraulic conductivity
l	rod length of TDR probe (m)
m	van Genuchten (1980) water retention shape parameter
M	particle size shape parameter
M_g	mass of gas phase of REV (kg)
M_l	mass of liquid phase of REV (kg)
M_{mol}	molecular weight
M_S	mass of solid phase of REV (kg)
M_t	total mass of REV (kg)
n	van Genuchten (1980) water retention shape parameter
N	particle size shape parameter
nb_i	number of particles in size class "i" of the Arya and Paris (1981) model
p	water vapor pressure
p_0	saturated vapor pressure
$p\beta$	tortuosity factor
q	soil water flux (m/s)
q_1	one-dimensional infiltration flux at the soil surface (m/s)
r	radius pore opening (m)
R	pore radius (m)
R_i	pore radius of size class "i" of the Arya and Paris (1981) model (m)
R_g	universal gas constant
REV	representative elementary soil volume (m³)
S	degree of saturation (m³/m³)
Sa	sand fraction (%)
Si	silt fraction (%)
S_1	sorptivity (m/$\sqrt{s}$)

S_2	second integral term of Philip's analytical infiltration equation (m/s)
t	time (s)
t_{tr}	travel time of electromagnetic TDR wave (s)
t^*	dimensionless time
T_l	temperature liquid phase (K)
V_g	volume gas phase of REV (m³)
V_l	volume liquid phase of REV (m³)
V_S	volume solid phase of REV (m³)
V_t	total volume of REV (m³)
w	gravimetric soil water content (kg/kg)
W_g	weight of gas phase of REV (kg)
W_l	weight of liquid phase of REV (kg)
W_S	weight of solid phase of REV (kg)
W_t	total weight of REV (kg)
x	horizontal space coordinate (m)
z	depth, positive downward (m)
z_m	reference depth of instantaneous profile method (m)
z_0	zero flux plan (m)
z^*	dimensionless depth
z_f	depth of infiltration wetting front (m)
α	tortuosity factor ($\alpha = 1.38$) of the Arya and Paris (1981) model
α_C	specific capacity scaling factor (1/m)
α_h	pressure head scaling factor (m)
α_I	cumulative infiltration scaling factor (m)
α_K	hydraulic conductivity scaling factor (m/s)
α_q	flux scaling factor (m/s)
α_t	time scaling factor (s)
α_z	depth scaling factor (m)
α_θ	water content scaling factor (m³/m³)
β	Parlange et al. (1982) conductivity shape parameter
γ	packing coefficient of the Haverkamp and Parlange (1986) model
ΔI^*	dimensionless net cumulative infiltration
Δq_1^*	dimensionless net infiltration flux
ΔS	change in water storage (m)
ε	soil porosity (m³/m³)
ε_{ap}	whole soil porosity (m³/m³)
ε_c	porosity of porous coarse fragments (m³/m³)
λ	Brooks and Corey (1964) shape parameter
λ_h	purely soil-related pressure head scaling factor (m)
λ_K	purely soil-related hydraulic conductivity scaling factor (m/s)
λ_θ	purely soil-related water content scaling factor (m³/m³)
η	Brooks and Corey (1964) hydraulic conductivity shape parameter
θ	volumetric soil water content (m³/m³)
θ_r	residual volumetric soil water content (m³/m³)
θ_S	volumetric soil water content at natural saturation (m³/m³)
θ_0	initial volumetric soil water content (m³/m³)
θ_1	surface volumetric soil water content (m³/m³)
θ^*	dimensionless volumetric soil water content: θ/θ_S (m³/m³)
μ_w	water viscosity (kg/m/s)
ρ_{ap}	whole soil dry bulk density (kg/m³)
ρ_c	specific weight of porous coarse fragments (kg/m³)

ρ_d dry bulk density (kg/m^3)
ρ_s specific weight of soil particles: 2.65 (g/cm^3)
ρ_w soil water specific weight (kg/m^3)
σ_{LV} liquid–vapor surface tension of water (kg/m/s^2)
ϕ_a air potential (joule/m^3)
ϕ_c capillary potential (joule/m^3)
ϕ_e envelope potential (joule/m^3)
ϕ_g gravitational potential (joule/m^3)
ϕ_o osmotic potential (joule/m^3)
Φ total soil water potential (joule/m^3)
χ_1 Boltzmann similarity variable (m/s$^{1/2}$)
χ_2 second term of Philip's time series solution (m/s)
χ_3 third term of Philip's time series solution (m/s$^{3/2}$)
ω contact angle between liquid and solid phase (rad)

6
Infiltration

R. David Espinoza
GeoSyntec Consultants

6.1   Introduction

This section deals with the physical process of water entering the soil from its surface. This process, known as infiltration, is one of the most important components of the hydrologic cycle. The amount of water that infiltrates into the soil is linked to the water redistribution process in the unsaturated zone, and this distribution depends upon the soil moisture conditions, water pressure, and unsaturated permeability. The amount of storm water infiltration can be used for water resource management (e.g., crop irrigation, groundwater recharge), design of drainage structures (e.g., drainage tunnels in slopes, drainage pipes for roads, leachate collection systems for landfills). An additional motivation for studying the infiltration process comes from the need to predict the amount of surface water runoff. Calculations of storm water runoff can be used for the design of diversion channels, culvert pipes, and erosion control systems. In addition, the growing public concern with pollutant transport through soils makes the quantification of infiltration a critical parameter for the design of appropriate remediation and pollution prevention techniques.

Several semiempirical and physically based infiltration models are discussed in this section, and specific examples for these models are presented. A summary of the most commonly used models is also presented.

6.2 Definitions

6.2.1 Basic Soil Properties

The purpose of this section is to introduce a number of fundamental relations between the constituent parts of a soil mass. These relationships can be used for soil classification. Soil properties have been found to correlate quite well for a given soil deposit. Thus, given the soil classification, it is possible to provide an initial estimate of the soil properties. The most widely accepted systems for classifying soils are the Unified Soil Classification System (USCS) and the Classification System of the United States Department of Agriculture (USDA). The USCS is commonly used by civil engineers, whereas the USDA system is more commonly used by agricultural engineers and soil physicists.

In general, soil is composed of solid particles and void spaces. The void spaces may be filled with air or water. The total volume of the soil sample is designated by V, and the volume of the solid particles, water, and air are defined by V_s, V_w, and V_a, respectively. The volume of voids (V_v) is the sum of volumes of air and water (i.e., $V_v = V_a + V_w$). The following ratios are widely used for soil description:

Porosity (*n*) represents the percentage of volume of voids with respect to the total volume of the soil mass. It is expressed as:

$$n = \frac{V_v}{V} \tag{1}$$

Degree of Saturation (*s*) represents the percentage of the total volume of voids that are filled with water, and it is expressed as:

$$s = \frac{V_w}{V_v} \tag{2}$$

Water Content (*w*), commonly used by soil engineers, is the ratio between the weight of water (W_w) and the weight of solids (W_s). It is usually expressed in percentage.

$$w = \frac{W_w}{W_s} \tag{3}$$

Volumetric Moisture Content (θ), commonly used in the description of infiltration, is a volumetric expression for the amount of moisture in a soil. It is defined as the ratio between volume of water and total volume.

$$\theta = \frac{V_w}{V} \tag{4}$$

The value of the volumetric moisture content at saturation (θ_s) is equivalent to the soil porosity (*n*).

Wilting Point (θ_w) is the lowest volumetric water content at which plant transpiration can take place. Figure 6.1 shows typical values of the wilting point for different soil types. The wilting point is also defined as the value of volumetric moisture content measured when a vacuum pressure (suction) of 15 bars is applied to the soil.

Residual Volumetric Moisture Content (θ_r) represents the amount of water that remains in a soil mass even when a high vacuum pressure (i.e., higher than 15 bars) is applied to the soil (Figure 6.1). In most practical cases, it is sufficient to assume θ_r equal to the wilting point (van Genuchten, 1980).
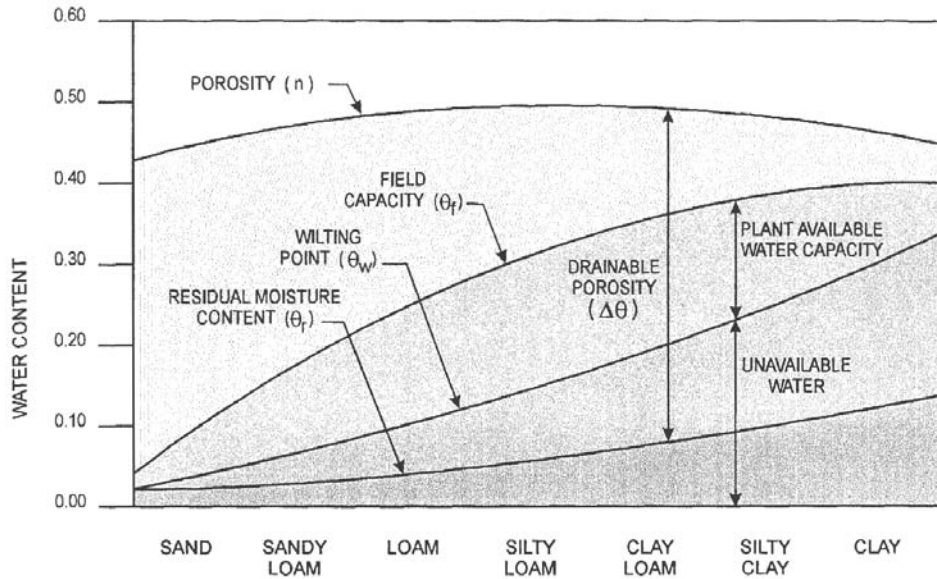
FIGURE 6.1 Relation among moisture retention parameters and soil texture class. (Adapted from Schroeder, P.R., Dozier, T.S. Zappi, P.A., McEnroe, B.M., Sjostrom, J.W., and Peyton, R.L. 1994. *The Hydrologic Evaluation of Landfill Performance HELP Model: Engineering Documentation for Version 3.* EPA/600/9-94/168b, U.S. Environmental Protection Agency Risk Reduction Engineering Laboratory. Cincinnati, Ohio.)

Effective Porosity ($\Delta\theta$), also known as drainable porosity, represents the total amount of water that can be stored in the soil (Figure 6.1). It is defined as the difference between the saturated and residual volumetric moisture content:

$$\Delta\theta = \theta_s - \theta_r \tag{5}$$

Moisture Deficit ($\Delta\theta_i$) represents the amount of water required to saturate the soil mass and is defined as the difference between the maximum and initial volumetric moisture content

$$\Delta\theta_i = \theta_s - \theta_i \tag{6}$$

Field Capacity (θ_f) is the value of volumetric moisture content that remains after a prolonged period of gravity drainage without any additional water supply (Figure 6.1). It is also defined as the volumetric moisture content that corresponds to a vacuum pressure of 0.33 bars.

Effective Degree of Saturation (S_e) represents the percentage of the effective porosity that is filled with water, and it is expressed as:

$$S_e = \frac{\left(\theta - \theta_r\right)}{\left(\theta_s - \theta_r\right)} = \frac{\left(\theta - \theta_r\right)}{\Delta\theta} = \frac{\left(s - s_r\right)}{\left(1 - s_r\right)} \tag{7}$$

where s_r is the residual degree of saturation corresponding to the value of residual moisture content θ_r, and s is the degree of saturation that corresponds to the value of volumetric moisture content, θ.

Hydraulic Conductivity (K) is a soil property that measures the ease with which water is able to travel through its pores. The hydraulic conductivity was defined by Henry Darcy in 1856 when he established

that the unidirectional flux of water through a sample of uniform sandy soil can be calculated by means of the following empirical relationship:

$$Q = KA \frac{\Delta\phi}{L} = KAj \tag{8}$$

where L and A are the thickness and cross-sectional area of the soil sample, respectively, K is the soil hydraulic conductivity, $\Delta\phi$ is the difference in piezometric head between the the ends of the soil sample (points 1 and 2 in Figure 6.2, respectively), and j is the hydraulic gradient. The hydraulic conductivity (K) depends on the viscosity and density of the fluid and the properties of the porous medium (e.g., pore size, shape, tortuosity, specific surface, and porosity). Because soil deposits are not generally isotropic media, soil hydraulic conductivity is not the same in every direction.
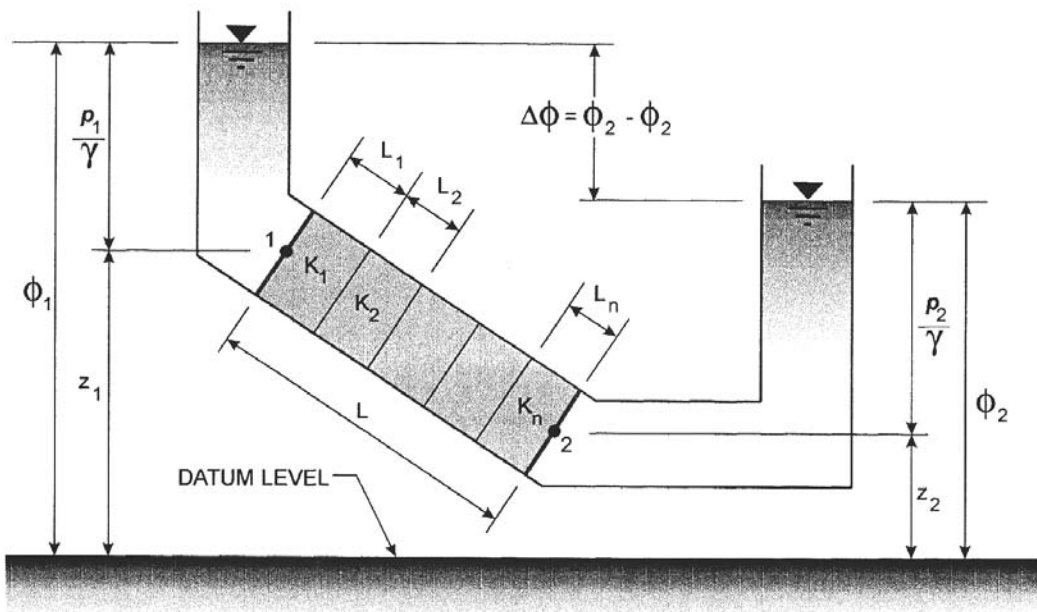
FIGURE 6.2 Schematic representation of Darcy's experiments.

There are many situations in which analysis of flow through a series of parallel layers is appropriate (e.g., a liner system for landfills). For such cases, if the hydraulic conductivity (K_i) and thickness (L_i) of each layer is known (Figure 6.2), the equivalent hydraulic conductivity (K) of the system with n layers can be computed as:

$$K = \frac{\sum_{i=1}^{n} L_i}{\sum_{i=1}^{n} \frac{L_i}{K_i}} \tag{9}$$

Typical values for some of the soil properties described above for various soil types, grouped according to USCS and USDA classifications systems, are provided in Table 6.1.

TABLE 6.1 Typical Soil Properties for Different Soil Textures

Soil Density	Soil Texture USDA	Soil Texture USCS	Soil Porosity (n)	Field Capacity (θ_f)	Wilting Point (θ_w)	Saturated Hydraulic Conductivity
Low	CoS	SP	0.417	0.045	0 018	1.0×10^{-2}
	S	SW	0.437	0.062	0.024	5.8×10^{-3}
	FS	SW	0.457	0.083	0.033	3.1×10^{-3}
	LS	SM	0.437	0.105	0.047	1.7×10^{-3}
	LFS	SM	0.457	0.131	0.058	1.0×10^{-3}
	SL	SM	0.453	0.190	0.085	7.2×10^{-4}
	FSL	SM	0.473	0.222	0.104	5.2×10^{-4}
	L	ML	0.463	0.232	0.116	3.7×10^{-4}
	SiL	ML	0.501	0.284	0.135	1.9×10^{-4}
	SCL	SC	0.398	0.244	0.136	1.2×10^{-4}
	CL	CL	0.464	0.310	0.187	6.4×10^{-5}
	SiCL	CL	0.471	0.342	0.210	4.2×10^{-5}
	SC	SC	0.430	0.321	0.221	3.3×10^{-5}
	SiC	CH	0.479	0.371	0.251	2.5×10^{-5}
	C	CH	0.475	0.378	0.251	2.5×10^{-5}
	G	GP	0.397	0.032	0.013	3.0×10^{-1}
Moderate	L	ML	0.419	0.307	0.180	1.9×10^{-5}
	SiL	ML	0.461	0.360	0.203	9.0×10^{-6}
	SCL	SC	0.365	0.305	0.202	2.7×10^{-6}
	CL	CL	0.437	0.373	0.266	3.6×10^{-6}
	SiCL	CL	0.445	0.393	0.277	1.9×10^{-6}
	SC	SC	0.400	0.366	0.288	7.8×10^{-7}
	SiC	CH	0.452	0.411	0.311	1.2×10^{-6}
	C	CH	0.451	0.419	0.332	6.8×10^{-7}
High	Liner Soil		0.427	0.418	0.367	1.0×10^{-7}
	Bentonite		0.750	0.747	0.400	3.0×10^{-9}

Soil Texture Abbreviations:
USDA: G= gravel; S= sand; Si= silt; C= clay; L= loam; Co= coarse; F= fine
USCS: G= gravel; S= sand; M= silt; C= clay; H= high plasticity; L= low plasticity;
        W= well graded; P= poorly graded
From Schroeder, P.R., Dozier, T.S. Zappi, P.A., McEnroe, B.M., Sjostrom, J.W., and Peyton, R.L. 1994. *The Hydrologic Evaluation of Landfill Performance HELP Model: Engineering Documentation for Version 3.* EPA/600/9-94/168b, U.S. Environmental Protection Agency Risk Reduction Engineering Laboratory. Cincinnati, Ohio. With permission.

6.2.2 Infiltration Definitions

In this section, a number of definitions commonly associated with the description of infiltration are described. A rainfall event is commonly described by its *intensity* (*i*) and *cumulative rainfall* (*I*). *Rainfall intensity* is the rate at which precipitation is applied on the soil surface. The rainfall intensity is commonly represented by a hyetograph which is a bar diagram that shows average rates of rainfall for a range of time intervals. An example of a hyetograph is presented in Figure 6.3. *Cumulative rainfall* (*I*) is the rainfall volume per unit area of soil surface, and it corresponds to the area of the hyetograph.

The actual *infiltration rate* at which water available at the surface infiltrates into the soil at a given time ranges from zero to a maximum value equal to the *infiltration capacity* (*f*). The soil's infiltration capacity is the maximum rate at which a given soil can absorb water. The infiltration capacity (*f*) varies with time, and it decreases approaching a minimum value (which is approximately equal to the saturated hydraulic conductivity, K_s) as the infiltration process continues (Figure 6.3a). The area under the infiltration capacity curve is the soil's *cumulative infiltration capacity* (*F*). The cumulative infiltration capacity represents the maximum volume of water per unit area that a given soil can absorb over a specific time period.
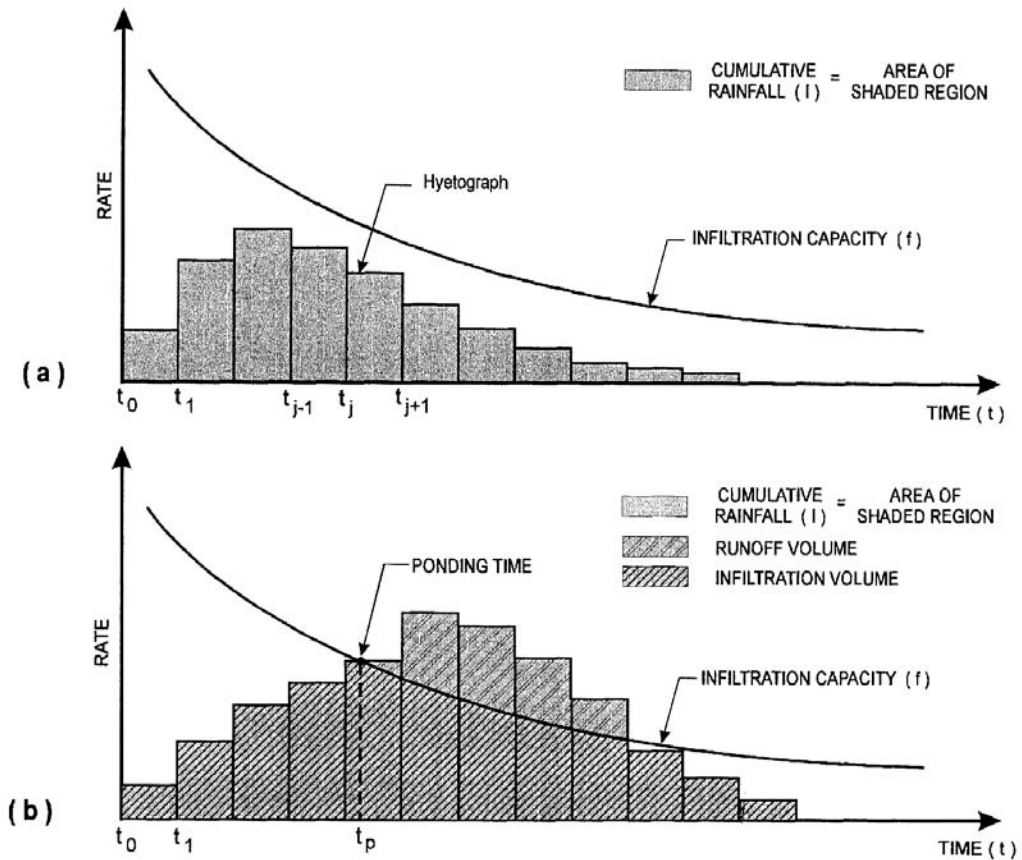
FIGURE 6.3 Infiltration diagram, (a) without ponding, (b) with ponding.

At the beginning of a rainfall event, all rainfall infiltrates into the soil. As rainfall continues, the soil infiltration capacity decreases with time. If the applied rainfall intensity (i) is smaller than the soil infiltration capacity (f), then all the rainfall will infiltrate into the soil (Figure 6.3a) over the entire rainfall period. If the applied rainfall intensity (i) is such that it is larger than the soil infiltration capacity (f), surface runoff will begin (Figure 6.3b). The amount of surface runoff can be estimated as the difference between the cumulative rainfall and cumulative infiltration capacity of the soil (Figure 6.3b). The particular point in time when surface runoff begins is called the *time of ponding* (t_p).

Figure 6.4 shows the three typical infiltration scenarios described by Mein and Larson (1973) for a constant rainfall intensity (i), and a homogeneous soil with a saturated hydraulic conductivity (K_s). When the rainfall intensity is less than the minimum infiltration capacity of the soil (i.e., $i < K_s \leq f$), independent of the rainfall duration, all rainfall is able to infiltrate and no runoff takes place. This scenario is shown as Case A in Figure 6.4. When the rainfall intensity is larger than the minimum infiltration capacity of the soil and the rainfall intensity (i) does not exceed the soil infiltration capacity (i.e., $K_s < i \leq f$), no runoff takes place. This scenario is shown as Case B in Figure 6.4. When ponding occurs, runoff starts and the soil infiltration capacity continuous to decrease until the minimum f is reached. This scenario is shown as Case C in Figure 6.4 (i.e., $K_s < f \leq i$).

6.2.3 Unsaturated Hydraulic Properties

Because infiltration can only take place into unsaturated soils, a brief description of unsaturated soil properties is presented. There are two important concepts associated with flow into unsaturated soils,
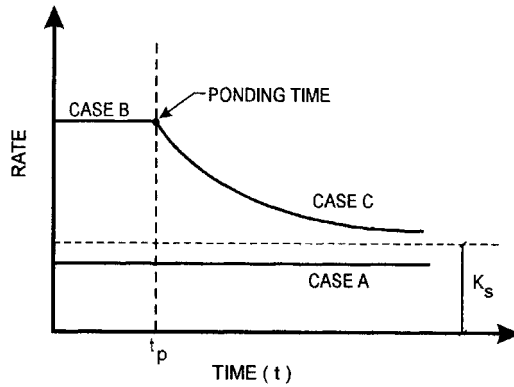
FIGURE 6.4 Typical scenario during infiltration.

namely (1) pressure and (2) relative permeability. The water pressure (u) in unsaturated soils is negative (i.e., less than atmospheric). Negative pressure is commonly referred to in the literature as suction or tension, and its associated suction head is defined as:

$$\psi = -\frac{u}{\gamma_w} \tag{10}$$

Above the phreatic level, water is held within the soil grains by capillary forces (surface tension). A similar concept is given by the rise above the water surface of a column of water inside a capillary tube (Figure 6.5a). For a vertical capillary tube (Figure 6.5a) with a constant radius (r), the value of the capillary rise (ψ_b) is given by:

$$\psi_b = \frac{c}{r} \leq h \tag{11}$$

where c is a constant for a given surface and liquid (c = 0.15 cm² for water on glass). The capillary rise cannot be higher than the height of the tube (h) above the water surface (Figures 6.5a and 6.5b). Equation (11) is obtained from force equilibrium in the vertical direction for a tube of uniform cross section, and
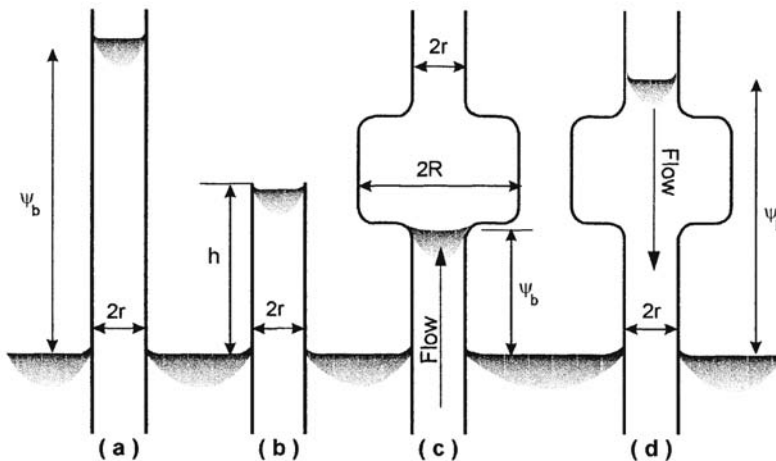
FIGURE 6.5 Influence of the tube cross-section variation on the height of capillary rise.

it applies regardless of whether the tube is filled from the bottom, a process called imbibition, or from the top, a process called drainage. On the contrary, the capillary rise in a tube of variable cross section depends on whether the tube is filled from the bottom (Figure 6.5c), or from the top (Figure 6.5d). Hence, to predict the capillary rise in these situations, it is necessary to know the wetting history. A porous medium can be regarded as a series of interconnected tubes so that a uniform location of the capillary rise cannot be estimated by means of a simple expression such as Equation (11).

Suction can be deemed as the pressure required to withdraw water from a soil sample. Several studies have shown that suction can be related to volumetric moisture content (or effective degree of saturation) of the soil. The variation of suction with volumetric moisture content (or effective degree of saturation) is called the water retention curve. A number of empirical and semiempirical functions have been proposed in the past to represent the water retention curves, and a summary of the most commonly used functions and the appropriate references are shown in Table 6.2. Figure 6.6a shows a typical variation of the water retention curve for infiltration and drainage. As the degree of saturation above the capillary rise decreases, the associated pore pressure also decreases (i.e., it becomes more negative).

TABLE 6.2 Summary of Water Retention Functions

Relationship	Reference
$S_e = \left(\dfrac{\psi}{\psi_b} \right)^{-\lambda}$ for $\psi \geq \psi_b$	Brooks and Corey (1964)
$S_e = \dfrac{1}{(1 + (\alpha\psi)^\beta)^{(1-1/\beta)}}$	van Genuchten (1980)
$S_e = \dfrac{B}{(B + \psi^\beta)}$ for $\psi > 0$	Brutsaert (1966)
$S_e = \dfrac{\alpha}{(\alpha + (\ln|\psi|)^\beta)}$ for $\psi > 1$	Vauclin et al. (1979)
$S_e = \dfrac{1}{\left[1 + \exp\left[(\psi - \psi_b)/\beta \right] \right]}$	Bumb et. al. (1992)

The unsaturated hydraulic conductivity varies with the soil water content, and it increases from a very low value for dry soils to its saturated value as soil water content increases to saturation. The variation of the unsaturated hydraulic conductivity is commonly represented by the *relative permeability* (k_r), which represents a ratio of the unsaturated hydraulic conductivity $K(S_e)$ to the saturated hydraulic conductivity (K_s):

$$k_r\left(S_e\right) = \frac{K\left(S_e\right)}{K_s} \tag{12}$$

The relative permeability varies from 0 to 1, which corresponds to values of water content varying from θ_r to θ_s, respectively.

Figure 6.6b shows the variation of a typical relative permeability curve (S_e vs. k_r). As shown in Figure 6.6b, the hydraulic conductivity is not a constant, but it varies with the effective degree of saturation, i.e., the hydraulic conductivity decreases as the amount of air in the porous medium increases. In Table 6.3, a summary of the most commonly used empirical and semiempirical relationships for unsaturated relative permeability are summarized.
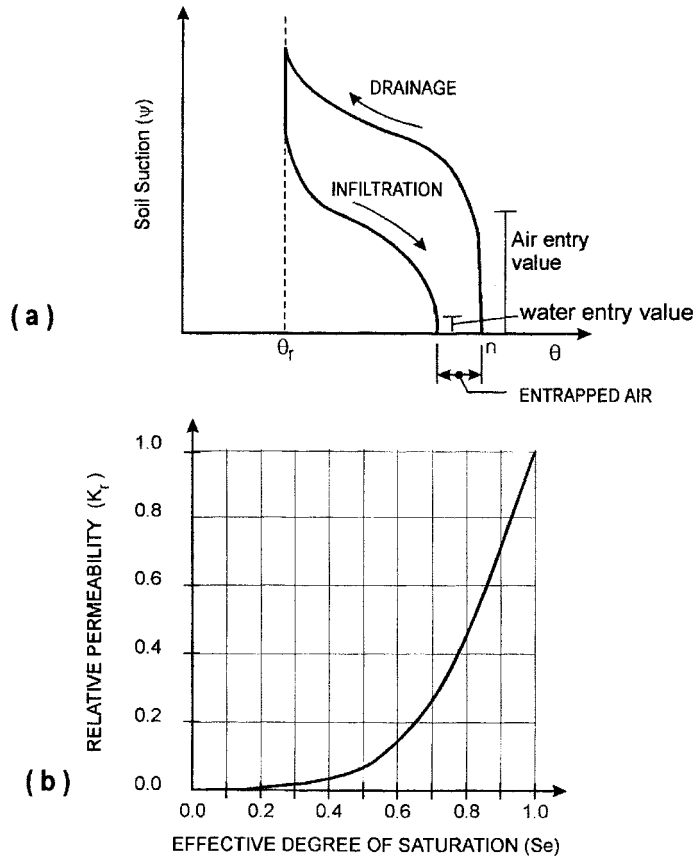
FIGURE 6.6 (a) Water retention curves, (b) relative permeability.

TABLE 6.3 Summary of Unsaturated Relative Permeability Functions

Relationship	Reference
$k_r(S_e) = S_e^\eta$	Brooks and Corey (1964)
$k_r(S_e) = S_e^{1/2}(1-(1-S_e^{1/\gamma})^\gamma)^2 \qquad (\gamma = 1-1/\beta)$	van Genuchten (1980)
$k_r(S_e) = S_e^2(1-(1-S_e^{1/\gamma})^\gamma) \qquad (\gamma = 1-2/\beta)$	van Genuchten (1980)
$k_r(\psi) = \dfrac{B}{B+\psi^\beta} \qquad \text{for } \psi \geq 0$	Gardner (1956)
$k_r(\psi) = \exp(\alpha\psi)$	Warrick (1991)
$k_r(\psi) = \psi^\alpha$	El-Kadi (1985)

6.3   General Infiltration Model

6.3.1   Introduction

Over the past 80 years, numerous analytical and semiempirical equations for one-dimensional horizontal and vertical infiltration through homogeneous soil with specific simplified initial conditions have been developed (e.g., Green and Ampt, 1911; Kostiakov, 1932; Horton, 1940; Philip, 1957a; Mein and Larson,
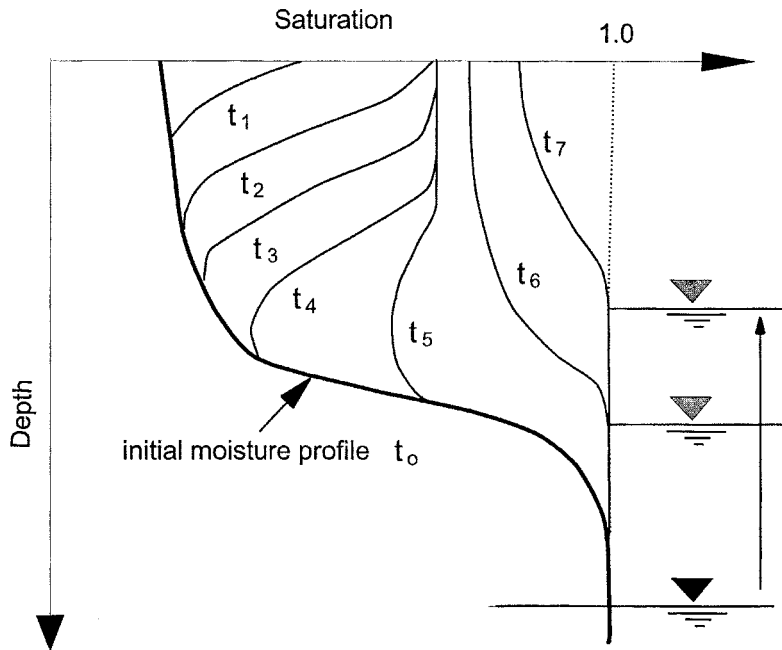
FIGURE 6.7 Moisture variation vs. depth at different times.

1973; Philip, 1984; Kao and Hunt, 1996). In addition, most vertical infiltration models assume that the rainfall applied at the surface can infiltrate for a long period of time (i.e., the groundwater table is very deep or soil has low permeability).

The amount of water that can infiltrate at its surface is directly related to the moisture characteristics of the unsaturated soil. The moisture variation, soil heterogeneities, and degree of saturation are the most influential properties. Although most models assume uniform moisture content profiles, soil deposits with varying moisture content profiles are found in many practical cases. A typical example of moisture variation in a homogeneous soil is shown in Figure 6.7 where soil is saturated at the groundwater surface and degree of saturation decreases in the upward vertical direction. The moisture distribution above the groundwater table at equilibrium conditions (i.e., no flow through the soil takes place) is similar to the water retention curves (see Figure 6.6a). Figure 6.7 also shows the moisture evolution at different times for a precipitation rate applied at the soil surface such that no surface ponding occurs.

Infiltration can only take place into unsaturated soils. Moreover, infiltration at the surface can be seen as a particular case of flow into unsaturated porous media in which the solution at the surface boundary is described. Early expressions of infiltration capacity (e.g., Green and Ampt, 1911) were later discussed within the more general framework of flow in unsaturated porous media (e.g., Philip, 1969; Mein and Larson, 1973; Morel-Seytoux and Khanji, 1974). To facilitate the discussion of some of the infiltration models, a brief description of the theory of unsaturated flow is presented in this section.

6.3.2 General Infiltration Equation

A complete description of the infiltration process entails the characterization of the variation of the moisture and pressure profile with depth at different times. The general second-order partial differential equation that describes the water flow through unsaturated or saturated anisotropic media, sometimes referred to as the Richards' equation (Mein and Larson, 1973) is written as:

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial x}\left[K_x(\psi)\frac{\partial \phi}{\partial x} \right] + \frac{\partial}{\partial y}\left[K_y(\psi)\frac{\partial \phi}{\partial y} \right] + \frac{\partial}{\partial z}\left[K_z(\psi)\frac{\partial \phi}{\partial z} \right] \qquad (13)$$

where *x, y, z* are the three space coordinates, ϕ is the piezometric head, and $\psi (= z - \phi)$ is the suction head. Equation was derived following the classical theory of mechanics of continuum media where the balance equation takes the form of a partial differential equation that describes the change per unit time of the considered quantity in a unit volume of porous media. In the derivation of Equation (13), the coordinate system is aligned to the principal directions of the hydraulic conductivity (i.e., $k_{xz} = k_{xz} = k_{xy} = k_{yx} = k_{yz} = k_{zy} = 0$). The volumetric moisture content on the left-hand side of Equation (13) can be written in terms of the piezometric head (ϕ) as:

$$C_w \frac{\partial \phi}{\partial t} = \nabla \left(K(\psi)\nabla \phi \right) \qquad (14)$$

or in terms of suction (ψ) as

$$C_w \frac{\partial \psi}{\partial t} = \nabla \left(K\nabla(z - \psi) \right) = \frac{\partial K}{\partial z} + \nabla \left(K\nabla \psi \right) \qquad (15)$$

where C_w is the specific water capacity and is given by:

$$C_w = -\frac{\partial \theta_w}{\partial \psi} \qquad (16)$$

As discussed in the previous section, the unsaturated soil hydraulic conductivity depends upon the moisture content. The variation of θ with z is the solution of the differential equation. Hence, the partial differential equations for unsaturated flow are nonlinear, and closed form solutions can be obtained only for some particular cases.

6.3.3   Moisture Diffusion Equation

Another useful representation of the unsaturated flow equation is in terms of volumetric moisture content. Because suction is related to the volumetric moisture content, Equation (13) can be rewritten as follows:

$$\frac{\partial \theta}{\partial t} = \frac{\partial K}{\partial \theta}\frac{\partial \theta}{\partial z} + \nabla \left(D(\theta)\nabla \theta \right) \qquad (17)$$

where

$$D(\theta) = K(\theta)\frac{d\psi}{d\theta} \qquad (18)$$

In Equation (17), θ is the volumetric moisture content and $D(\theta)$ is the coefficient of water diffusion. The first term on the right-hand side is the contribution of gravity, whereas the second term is the contribution of capillarity. If the flow is produced mainly by capillary forces (e.g., horizontal flow), then the first term on the right-hand of Equation (17) can be neglected. In such a case, the simplified formulation is equivalent to a diffusion equation, and the techniques available for deriving closed-form solutions for certain diffusion problems could be used for solving unsaturated flow problems (Philip,

1969). It is noted that Equation is not valid at saturation because the change in moisture content is zero and, hence, the diffusivity term, Equation (18), is not applicable.

In general, when dealing with general geometric, boundary, and/or initial conditions, finite difference and finite element numerical models are often used to solve Equation (17), or its equivalent, Equation (13). Despite the increasing popularity of numerical models, closed-form and semiempirical solutions are commonly used for predictions of infiltration and runoff volumes (Van Mullen, 1991; Rawls and Brakensiek, 1986). In what follows, several expressions available for vertical infiltration are described.

6.4 Physically Based Infiltration Capacity Equations

6.4.1 Introduction

A number of physically based procedures have been derived that evaluate the volume of infiltrated water into a uniform soil (e.g., Green and Ampt, 1911; Philip, 1957a; Parlange, 1975; Babu, 1976; Tolikas et al., 1983). Most of these expressions assume that the rainfall rate is constant and that ponding occurs at the beginning of the infiltration process. The Green and Ampt model and Philip's model are the physically based models most commonly used to calculate the soil infiltration capacity. A description of these models is presented in the following sections.

6.4.2 Green and Ampt Solution

6.4.2.1 General

Green and Ampt (1911) made one of the first attempts at solving for the infiltration capacity using an analytical method. They proposed an approximate solution based on one-dimensional vertical flow. In their model for vertical infiltration, the infiltrating wetting front is defined by a constant suction head at the wetting front (ψ_f), and in the soil above the wetting front, the moisture deficit ($\Delta\theta_i$) is uniform, the coefficient of hydraulic conductivity (K) is constant, and the hydraulic head at inlet (h_o) is also constant throughout the infiltration process. A diagram of this model is presented in Figure 6.8a. The main assumption in Green and Ampt's model is the distribution of the moisture profile during infiltration. It is assumed that the soil above the wetting front is fully saturated, while the soil below the wetting front is at the initial degree of saturation. This process is depicted in Figure 6.8b where at time $t = t_1$ the soil is at saturation ($S_e = 1.0$) from the surface to the depth of the wetting front (Z_f). By applying Darcy's law and continuity equation between the surface and the wetting front, Green and Ampt (1911) proposed the following expression for the infiltration capacity (f):

$$f = K \frac{\left(h_o + Z_f + \psi_f\right)}{Z_f} = \frac{d\left(Z_f \, \Delta\theta_i\right)}{dt}$$ (19)

The infiltration rate (f) given by Equation (19) decreases as t increases, approaching the value of (K) at large times. This behavior is characteristic, and it agrees well with experimental evidence. It is also noted that Equation (19) predicts that at small times the infiltration capacity is high, which is appropriate. However, the rate of infiltration at $t = 0$ is infinite. Integration of Equation (19) yields an expression for the amount of cumulative infiltration capacity (F) as a function of time (t):

$$F = Z_f \, \Delta\theta_i = Kt + \Delta\theta_i \left(\psi_f + h_o\right) Ln\left[1 + \frac{F}{\Delta\theta_i\left(\psi_f + h_o\right)}\right]$$ (20)

Other researchers (e.g., Philip, 1957d; Swartzendruber, 1987; Parlange, 1975) arrived at an expression similar to the Green and Ampt model using the theory of unsaturated flow. Similarly, using a two-phase
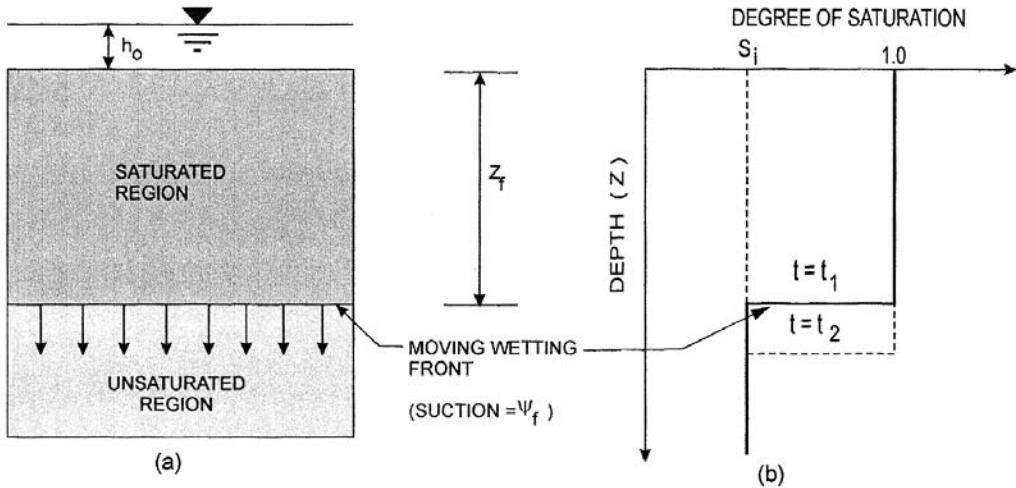
FIGURE 6.8 Green and Ampt flow model.

flow (air-water) model, Morel-Seytoux and Khanji (1974) arrived at an expression similar to Equation (20). The main advantage of the Green and Ampt model is that despite its simplicity, the model seems to capture the basic features of the vertical infiltration process. In addition, all Green and Ampt model parameters have measurable physical soil properties. The step function assumed in the Green and Ampt model provides a fair representation for relatively dry and coarse soils (Philip, 1957d).

It is noted that the cumulative infiltration in Equation (20) is not expressed explicitly as a function of time and, hence, an iterative procedure is needed to compute F for a given time t. The Green and Ampt equation is commonly expressed in dimensionless form as:

$$T = U - Ln\left[1 + U\right] \qquad (21)$$

where U and T are the normalized cumulative infiltration and normalized time, respectively, and are given by:

$$U = \frac{F}{\Delta\theta_i\left(\psi_f + h_o\right)} \qquad (22)$$

$$T = \frac{Kt}{\Delta\theta_i\left(\psi_f + h_o\right)} \qquad (23)$$

A plot of normalized infiltration U vs. T is shown in Figure 6.9. This plot may be used to facilitate an approximation of the infiltration rate for a given time t without an iterative procedure.

Example: Given a silty loamy soil with a suction value (ψ_f) of 16.7 cm, ponding head (h_o) of zero, hydraulic conductivity (K) of 0.65 cm/hr, and an effective porosity ($\Delta\theta$) of 0.34, find the cumulative infiltration capacity (F) after 1 hour.

Solution: From Equation (23), the normalized time is T = (0.65)(1)/(16.7)/(0.34) = 0.115. Using Figure 6.9, the normalized infiltration capacity corresponding to T = 0.12 is U = 0.57. Then, from Equation (22), F = (0.34)(0.57)(16.7 + 0) = 3.18 cm.
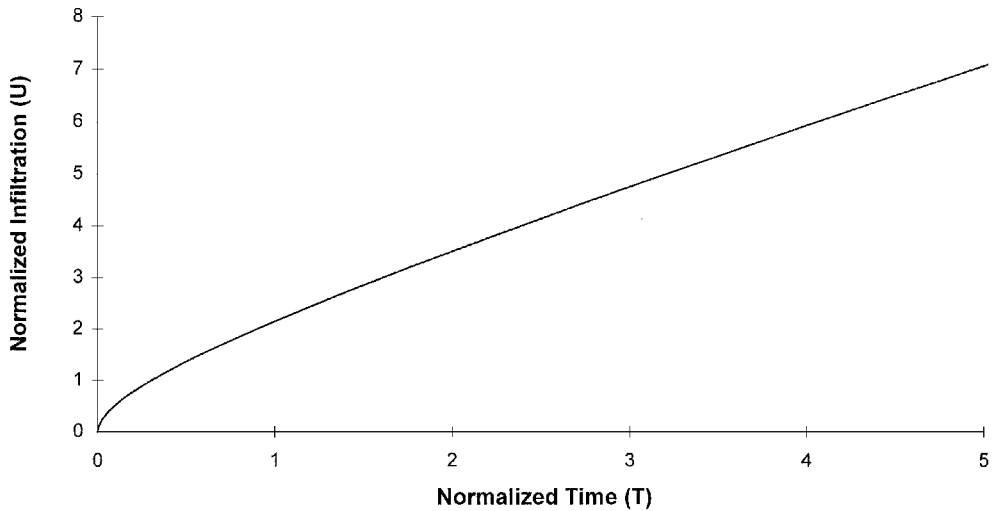
FIGURE 6.9 Dimensionless curve of cumulative infiltration vs. time.

6.4.2.2 Time of Ponding

During the early stages of infiltration, all the rainfall percolates into the soil and no ponding above the surface occurs. During this period of time, soil is unsaturated at the surface, thus water pore pressure is negative. For a uniform rain, when the infiltration rate is equal to the infiltration capacity and the pressure head at the surface becomes zero, ponding occurs. Assuming the moisture profile as shown in Figure 6.8b, the depth of percolated water can be approximated as $Z_f = F/\Delta\theta_i$ and replacing $f = i$ and $h_o = 0$ in Equation (19), the cumulative infiltration at ponding (F_p) can be computed as (Mein and Larson, 1973):

$$F_p = \frac{K\,\Delta\theta_i\,\psi_f}{(i-K)} \qquad (24)$$

The corresponding time of ponding can be computed as:

$$t_p = \frac{F_p}{i} \qquad (25)$$

Example: Given a silty loam with an suction value (ψ_f) of 16.7 cm, ponding head (h_o) equal to 0, hydraulic conductivity of 0.65 cm/hr, an effective porosity ($\Delta\theta$) of 0.486, an initial effective degree of saturation of 0.3, and a rainfall intensity (i) of 5 cm/hr, find the cumulative infiltration at ponding (F_p) and the corresponding ponding time (t_p).

Solution: The initial moisture deficit can be estimated as $\Delta\theta_i = (1 - 0.3)0.486 = 0.34$. Using Equations (24) and (25), the cumulative infiltration is then $F_p = (0.65)(0.34)(16.7)/(5 - 0.65) = 0.85$ cm, and the corresponding time of ponding is $t_p = 0.85/5 = 0.17$ hr.

6.4.2.3 Parameter Determination

Good agreement between computed and measured data has contributed to the widespread use of the Green and Ampt model (Bouwer, 1969; Whisler and Bouwer, 1970; Childs and Bybordi, 1969; Van Mullen, 1991; Rawls and Brakensiek, 1982). This good agreement depends largely on the appropriate selection of the parameters $\Delta\theta_i$, K, and ψ_f The moisture deficit parameter ($\Delta\theta_i$) can be measured in the field.

Because of entrapped air in the soil, the hydraulic conductivity of the soil above the wetting front (K) is less than the saturated hydraulic conductivity (K_s). Available data indicate that the value of saturated permeability in the wetted zone during infiltration may range from $0.4K_s$ to $0.6K_s$, and for many applications, an average value of $0.5K_s$ may be used (Bouwer, 1966).

Based on theoretical and experimental studies, a number of expressions have been proposed to estimate the parameter ψ_f. Results obtained using Equation (20) and values of suction at the wetting front (ψ_f) that corresponds to the initial volumetric water content do not agree with experimental data (Whisler and Bouwer, 1970). Experimental studies performed by Bouwer (1966) suggest that the use of the water entry value of the soil (Figure 6.6a) as the value of ψ_f provides a good agreement between experimental and theoretical results. Based upon a comprehensive compilation of published soil moisture characteristics data (approximately 5000 soil samples), Rawls et al. (1983b) provided infiltration parameters (effective porosity, wetting front, capillary pressure, and hydraulic conductivity) for typical soils (Table 6.4).

Using Richard's equation for one-dimensional vertical infiltration, Mein and Farrel (1974) and Swartzendruber (1987) proposed a theoretical expression for ψ_f by relating the area of suction versus relative hydraulic conductivity as:

$$\psi_f = \int_0^{\psi_i} k_r(\psi)d\psi \tag{26}$$

where ψ_i corresponds to the suction value at the initial moisture content. Experimental results indicate that Equation (26) provides a good approximation for the parameter ψ_f (Morel-Seytoux and Khanji, 1974; Mein and Farrel, 1974; Brakensiek, 1977; Swartzendruber, 1987). Similarly, based on two-phase flow theory (air and water), Morel-Seytoux and Khanji (1974) derived an expression for suction at the wetting front (ψ_f) similar to Equation (26). Previously, Bouwer (1966) proposed an empirical expression for ψ_f similar to (26) where the upper limit of integration was taken as very large or infinite, thus neglecting the effect of the initial conditions. Also, Mein and Larson (1973) defined ψ_f as the area under the curve ψ vs. k_r. Table 6.3 shows different expressions of k_r written explicitly in terms of soil suction (ψ) that can be integrated analytically or numerically to compute ψ_f using Equation (26).

Rawls et al. (1983a) expressed the Green and Ampt ψ_f parameter in terms of the Brooks and Corey (1964) water retention parameters (ψ_b and λ) as:

$$\psi_f = \frac{2+3\lambda}{1+3\lambda}\left(\frac{\psi_b}{2}\right) \tag{27}$$

where ψ_b is the water entry value, and λ is the pore index, as defined by Brooks and Corey (1964). An experimental study performed by Brakensiek (1977) showed that the suction at the wetting front can be initially approximated as $0.76\psi_b$.

The parameters $\Delta\theta_i$, K, and ψ_f can also be determined from infiltration tests by fitting Equation (20) to the data. Although this approach is less complicated than the laboratory tests required to determine ψ vs. k_r curves, the parameters obtained from infiltration tests cannot be used for different initial and rainfall conditions. On the contrary, the use of Equation (26) allows for the representation of Green and Ampt model parameters in terms of fundamental soil unsaturated parameters. The specific procedure used to evaluate the parameter ψ_f will depend on the type of information available.

6.4.2.4 Flow in Nonuniform Soils

Soil deposits with varying moisture content are found in many practical cases. For instance, surface evapotranspiration and vertical drainage cause water content to increase with depth (i.e decreasing moisture deficit). The main advantage of the Green and Ampt model, Equation (20), is that it depends on measurable physical parameters such as permeability and moisture content, and it is mathematically

TABLE 6.4 Green and and Ampt Parameters

Soil Type	Horizon	Total Porosity	Effective Porosity	ψ_f (cm)	K (cm/h)
Sand[1]		0.437	0.417	4.95	11.7
	A	0.452	0.431	5.34	
	B	0.440	0.421	6.38	
	C	0.424	0.408	2.07	
Loamy sand		0.437	0.401	6.13	2.99
	A	0.457	0.424	6.01	
	B	0.447	0.412	4.21	
	C	0.424	0.385	5.16	
Sandy loam		0.453	0.412	11.01	1.09
	A	0.505	0.469	15.24	
	B	0.466	0.428	8.89	
	C	0.418	0.389	6.79	
Loam		0.463	0.434	8.89	0.34
	A	0.512	0.476	10.01	
	B	0.512	0.498	6.40	
	C	0.412	0.382	9.27	
Silt loam		0.501	0.486	16.68	0.65
	A	0.527	0.514	10.91	
	B	0.533	0.515	7.21	
	C	0.470	0.460	12.62	
Sandy clay loam		0.398	0.330	21.85	0.15
	B	0.393	0.330	26.10	
	C	0.407	0.332	23.90	
Clay loam		0.464	0.309	20.88	0.10
	A	0.497	0.430	27.00	
	B	0.451	0.397	18.52	
	C	0.452	0.400	15.21	
Silty clay loam		0.471	0.432	27.30	0.10
	A	0.509	0.477	13.97	
	B	0.469	0.441	18.56	
	C	0.475	0.451	21.54	
Sandy clay		0.430	0.321	23.90	0.06
	B	0.435	0.335	36.74	
Silty clay		0.479	0.423	29.22	0.05
	B	0.476	0.424	30.66	
	C	0.464	0.416	45.65	
		0.475	0.385	31.63	
Clay	B	0.470	0.412	27.72	0.03
	C	0.483	0.419	54.65	

[1] Values for the texture class

Adapted from Rawls, W.J., Brakensiek, D.L., and Miller, N. 1983b. Agricultural management effects on soil water processes. Part I: Soil water retention and Green and Ampt infiltration parameters. *Trans ASAE.* 2(66), 1747-1752.

simple. This characteristic allows Equation (20) to model conditions of varying moisture content and increasing permeability with depth (Bouwer, 1969 and 1976; Childs and Bybordy, 1969). In what follows, the procedures outlined by Bouwer (1969) for handling variable moisture content and variable hydraulic conductivity are described.

Variable moisture content. To find the relationship between cumulative infiltration capacity (F) and time (t) for a soil with variable moisture content, the following procedure is used. First, the soil column is divided into horizontal layers with a uniform moisture content within each layer. Then, the traveling time (Δt_m) required for the wetting front to travel through the mth soil layer, i.e., from the top (Z_m) to the bottom (Z_{m+1}), is computed. The traveling time is computed as $\Delta t_m = t'_m - t'_{m+1}$, where t'_m and t'_{m+1} are "fictitious" times evaluated using Equation (20), assuming that the soil column has a uniform

moisture content $\Delta\psi_m$ corresponding to the mth layer. The actual time required for the wetting front to travel from the soil surface to the bottom of the mth is computed by adding the traveling times for each soil layer. The procedure outlined above is shown in the example problem taken from Bouwer (1969).

Example: For a 100-cm soil column with a variable moisture deficit shown in column 2 of Table 6.5, $\psi_f + h_o = 10$ cm and hydraulic conductivity (K) of 10 cm/day, find the variation of F vs. t for the different depths shown in column 1 of Table 6.5.

TABLE 6.5 F vs. t for a Soil with Varying Moisture Content

Z_f	$\Delta\theta_i$	t_{m+1}	t_m	Δt	t	F
0	0.000	0.000			0.000	0.00
10	0.222	0.030	0	0.030	0.030	2.22
20	0.206	0.096	0.028	0.068	0.099	4.28
30	0.190	0.175	0.089	0.086	0.185	6.18
40	0.174	0.254	0.160	0.094	0.278	7.92
50	0.158	0.325	0.230	0.095	0.373	9.50
60	0.142	0.384	0.292	0.092	0.465	10.92
70	0.126	0.427	0.341	0.086	0.551	12.18
80	0.110	0.451	0.373	0.079	0.629	13.28
90	0.094	0.455	0.386	0.069	0.699	14.22
100	0.078	0.437	0.378	0.059	0.758	15.00

Solution: The following computation shows how the values for the third layer (4th row), i.e., when the wetting front is at 30 cm, are obtained. Replacing $\Delta\theta_i = 0.190$, $\psi_f + h_o = 10$ cm and K = 10 cm/day in Equation (20), the values of $t_1 = 0.175$ (column 3) and $t_2 = 0.089$ (column 4) are computed for $Z_3 = 30$ cm and $Z_2 = 20$ cm, respectively. The time increment (column 5) is $\Delta t_3 = 0.175 - 0.089 = 0.086$ and the corresponding total time is $t_3 = 0.99 + 0.086 = 0.185$. The value of cumulative infiltration $F_3 = 4.28 + (0.190)(30 - 20) = 6.18$. The values of $t_2 = 0.99$ and $F_2 = 4.28$ were previously computed using a similar procedure.

Variable hydraulic conductivity. To find the relationship between F and t for a soil with variable hydraulic conductivity, a procedure similar to the one described in the preceding section is used. Because of the nonuniform hydraulic conductivity, two additional steps are required: (1) an equivalent hydraulic conductivity, K_m, for the soil column above the mth layer at each depth is computed using Equation (9) and (2) an average hydraulic conductivity, $K_{avg} = [K_m + K_{m+1}]/2$, for each layer is computed. The fictitious time t'_m required for the wetting front to reach the top of the mth layer (Z_m) is computed using Equation (26) with the value of K_m. The value of t'_{m+1} required for the wetting front to reach the bottom of the mth layer (Z_{m+1}) is computed assuming that the layer has an average permeability K_m and that the entire soil column has a uniform moisture content $\Delta\theta_m$. The procedure outlined above is shown in the example below.

Example: For a 100-cm soil column with a variable moisture deficit and hydraulic conductivity (K) shown in columns 2 and 3 of Table 6.6, respectively, and $\psi_f + h_o = 10$ cm, find the variation of F vs. t for the different depths shown in column 1 of Table 6.6.

Solution: The following computation shows how the values for the third layer (4th row of Table 6.6) are obtained. The first step is to compute the average hydraulic conductivity for the layer of interest using Equation (9) . Thus $(Z_3 - Z_2)/K_3 = (30 - 20)/6 = 1.67$, $K_3 = 30/(1.67 + 1.25 + 1) = 7.66$, and $K_{avg3} = (8.89 + 7.66)/2 = 8.27$. The next step is to calculate t_3 using the values of $\Delta\psi_{i3} = 0.190$ and $K_{avg3} = 8.27$ cm/day, a value of $t_3 = 0.211$ days is obtained (column 7). Similarly, using $\Delta\theta_{i3} = 0.190$ and $K_2 = 8.89$ cm/day, a value of $t_2 = 0.100$ days is obtained (column 8). The time increment (column 9) is $\Delta t_3 = 0.211 - 0.100 = 0.111$, and the corresponding total time is $t_3 = 0.104 + 0.111 = 0.216$.

TABLE 6.6   F vs. t for a Soil with Varying Permeability

Z_f	$\Delta\theta_i$	K	$\Delta Z/K$	K_{eq}	K_{avg}	t_{m+1}	t_m	Δt	t	F
0	0.000			10.00	10.00	0.000			0.000	0.00
10	0.222	10	1.00	10.00	10.00	0.030	0	0.030	0.030	2.22
20	0.206	8	1.25	8.89	9.44	0.102	0.028	0.074	0.104	4.28
30	0.190	6	1.67	7.66	8.27	0.211	0.100	0.111	0.216	6.18
40	0.174	4	2.50	6.23	6.95	0.365	0.209	0.156	0.372	7.92
50	0.158	4	2.50	5.61	5.92	0.549	0.370	0.180	0.551	9.50
60	0.142	3	3.33	4.90	5.25	0.731	0.521	0.210	0.761	10.92
70	0.126	3	3.33	4.49	4.69	0.909	0.696	0.214	0.975	12.18
80	0.110	2	5.00	3.89	4.19	1.077	0.830	0.247	1.222	13.28
90	0.094	2	5.00	3.52	3.70	1.229	0.992	0.237	1.459	14.22
100	0.078	2	5.00	3.27	3.39	1.287	1.073	0.214	1.673	15.00

The value of cumulative infiltration $F_3 = 4.28 + (0.190)(30 - 20) = 6.18$. The values of $t_2 = 0.99$, $F_2 = 4.28$, and K_{eq2} have been previously computed using a similar procedure.

6.4.2.5   Nonuniform Rainfall Conditions

Nonuniform rainfall events are represented by hydrographs. Assuming that ponding occurs in a time interval Δt_j (i.e., $t_{j-1} \le t_p < t_j$), the total cumulative infiltration (F_p) at the time of ponding (t_p), obtained from water balance, is given by:

$$F_p = \sum_{m=1}^{j-1} \Delta t_m i_m + \left(t_p - t_{j-1}\right)i_j = F_{j-1} + \left(t_p - t_{j-1}\right)i_j \tag{28}$$

From Equation (28), assuming that before ponding the cumulative infiltration equals the cumulative rainfall (i.e., $F_{j-1} = I_{j-1}$), a general expression for time of ponding for non-uniform conditions can be obtained as:

$$t_p = t_{j-1} - \frac{I_{j-1}}{i_j} + \frac{F_p}{i_j} \tag{29}$$

Replacing Equation (24) for cumulative infiltration at ponding into Equation (28), it can be shown that t_p for the Green and Ampt model is given by:

$$t_p = t_{j-1} - \frac{I_{j-1}}{i_j} + \frac{K \Delta\theta_i \psi_f}{i_j\left(i_j - K\right)} \tag{30}$$

Equation (30) is evaluated for $j = 1, 2, \ldots$ until $t_j > t_p$. It can be easily shown that for a constant rainfall (i), Equation (30) reduces to Equation (24). The time of ponding may take place at the beginning of the interval (i.e., $t_p = t_{j-1}$) or within the interval (i.e., $t_{j-1} < t_p < t_j$). The latter case can be reduced to the first case by dividing the time step into two time steps with the computed value of t_p corresponding to the end of the first interval and the beginning second interval, i.e., $[t_{j-1}, t_p]$ and $[t_p, t_j]$. This facilitates the computer implementation of the evaluation of the cumulative infiltration for nonuniform conditions.

The values of f_t (infiltration capacity at time t) are computed from known values of cumulative infiltration (F_t) according to the following equation:

$$f_t = K\left[\frac{\psi_f \Delta\theta}{F_t} + 1\right] \tag{31}$$

The computed infiltration capacity (f_t) is then compared to the rainfall intensity (i_j) for the same time interval. No runoff will take place if the infiltration capacity is greater than the rainfall intensity (i.e., $f_t > i_j$) in which case the cumulative infiltration capacity ($F_{t+\Delta t}$) is computed as:

$$F_{t+\Delta t} = F_t + i_j \, \Delta t_j \tag{32}$$

where the subscripts t and t + Δt indicate the quantities computed at the previous and present time steps, respectively. If the infiltration capacity is smaller than the rainfall intensity $f_t \leq i_j$, then there is ponded water throughout the interval and the cumulative infiltration at the end of the interval is given by:

$$F_{t+\Delta t} = F_t + K \, \Delta t + \Delta \theta_i \, \psi_f \, Ln\left[\frac{F_{t+\Delta t} + \Delta \theta_i \psi_f}{F_t + \Delta \theta_i \psi_f} \right] \tag{33}$$

Example: Given a sandy loam soil with an suction value (ψ_f) of 10 cm, ponding head $h_o = 0$, hydraulic conductivity of 1.1 cm/hr, an initial effective porosity of 0.3, and a cumulative rainfall measured every 10 minutes and shown in column 3 of Table 6.7. Find the excess rainfall runoff.

TABLE 6.7   Variable Rainfall Conditions

Time	Rainfall Characteristics				Infiltration Capacity		Runoff
(min)	ΔI (cm)	I (cm)	i (cm/h)	t_p (min)	F (cm)	f (cm/h)	(cm)
0	0.18	0	1.08		0	19.43	
10	0.21	0.18	1.26	26.23	0.18	19.43	
20	0.26	0.39	1.56	24.35	0.39	9.56	
30	0.32	0.65	1.92	31.76	0.65	6.18	
40	0.37	0.97	2.22	40.89	0.97	4.50	
50	0.43	1.34	2.58	50.34	1.34	3.56	
60	0.64	1.77	3.84	59.85	1.77	2.96	0.00
70	1.14	2.41	6.84	69.73	2.23	2.58	0.18
80	3.18	3.55	19.08	79.82	2.64	2.35	0.91
90	1.65	6.73	9.9	89.36	3.02	2.19	3.71
100	0.81	8.38	4.86	98.46	3.37	2.08	5.01
110	0.52	9.19	3.12	107.58	3.71	1.99	5.48
120	0.42	9.71	2.52	117.07	4.04	1.92	5.67
130	0.36	10.13	2.16	126.75	4.35	1.86	5.78
140	0.28	10.49	1.68	137.14	4.66	1.81	5.83
150	0.24	10.77	1.44	149.26	4.94	1.77	
160	0.19	11.01	1.14	222.71	5.18	1.74	
170	0.17	11.20	1.02	118.58	5.37	1.71	
180		11.37			5.54	1.70	

Solution: Values of cumulative infiltration (I in column 3 of Table 6.7) are computed by cumulatively adding the values of rainfall shown in column 2. The values of rainfall intensity (4th column) are computed by dividing ΔI_j by the time increment Δt_j. For instance, the value of i for the 1st row is $i_1 = \Delta I_1/\Delta t_1 = (0.18)/(10 - 0) = 0.018$ cm/min $= 1.08$ cm/hr. The hyetograph shown in Figure 6.10 corresponds to these values. The 5th column of Table 6.7 is computed using Equation (29). Comparing the values of the 1st and 5th column, the time of ponding corresponds to the first value in the 5th column that is smaller than the value of the 1st column. In our example, the time of ponding is $t_p = 59.85$ min (assume $t_p = 60$ min).
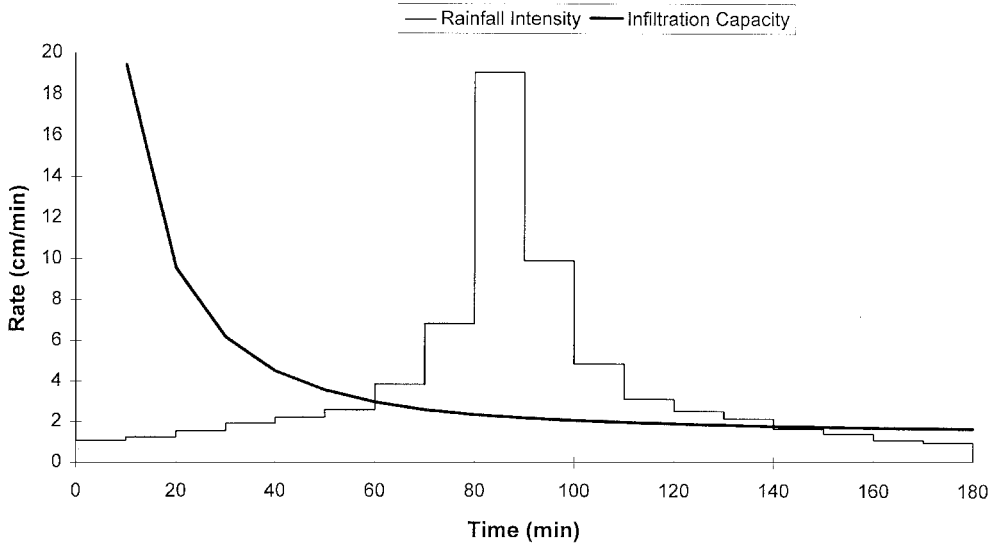
FIGURE 6.10 Example of nonuniform rainfall conditions.

The following computation shows how the values for the cumulative infiltration (F) before and after ponding are computed. Before ponding (i.e., i < f), F is computed using Equation (32). For example, F_5 = I_5 = 0.65 + 0.32 = 0.97 cm. After ponding two cases may occur: (1) the rainfall intensity is greater than the infiltration capacity (i.e., i > f); and (2) the rainfall intensity is smaller than the infiltration capacity (i.e., i < f). If i > f, then F is computed using Equation (33). For instance, the 8th column of Table 6.7 corresponds to this case, and the cumulative infiltration capacity is computed as F_8 = 1.77 + 1.1(10) + 0.3(10) ln[(F_8 + 3)/(1.77 + 3)] = 2.23 cm. If i < f then the infiltration capacity is computed using Equation (32).

Once the cumulative infiltration capacity is calculated, the column corresponding to the infiltration capacity (f) is computed using Equation (31). For instance, the value of the 5th row is computed as f_5 = (1.1)(0.3(10)/0.97 + 1) = 4.5. The resulting infiltration capacity curve is also shown in Figure 6.10.

6.4.2.6 Generalized Green and Ampt Solution

In the derivation of the Green and Ampt model, it is implicitly assumed that ponding occurs at the beginning of the infiltration process (i.e., $t = 0$). Assuming that a negligibly thin layer of water ponds at the surface after ponding time (i.e., $h_o = 0$ in Equation [19]), Swartzendruber (1987) derived a generalized expression of the Green and Ampt equation that is valid for ponding time greater than or equal to zero, as follows.

$$T - T_o = U - U_o - Ln\left[1 + N\left(U - U_o\right)\right] \quad \text{for} \quad T \geq T_o \qquad (34)$$

where U and T are the normalized expressions given by Equations (22) and (23), respectively (assuming $h_o = 0$); the values of U_o and T_o are the normalized values corresponding to the cumulative infiltration and time at ponding (i.e., I_p and t_p, respectively), and the parameter N is given by:

$$N = 1 - K/i \qquad (35)$$

Equation (35) is valid for $0 < N \leq 1$.

Using Equation (35), the dimensionless parameters T_o and U_o can be expressed as a function of N as:

$$T_o = \left(1 - N\right)^2 \big/ N \tag{36}$$

$$U_o = \left(1 - N\right) \big/ N \tag{37}$$

It can be shown that for $N = 1$ (low permeability and/or large i), Equation (34) yields the familiar Green and Ampt equation.

6.4.3 Philip Equation

6.4.3.1 General

Philip (1957a) presented a closed-form solution for the soil infiltration capacity. Starting from the diffusion-type unsaturated flow equation, Equation (17), and assuming a homogeneous soil with an initial uniform moisture content (θ_i) and an applied moisture content at the surface (θ_s), Philip (1957a) derived a series expansion solution for the cumulative infiltration (F) as a function of time as:

$$F(t) = \sum_{n=1}^{\infty} A_n\left(\theta\right) t^{n/2} + K_i t \tag{38}$$

where the coefficients A_n are functions of the volumetric moisture content. These coefficients are the solution of ordinary partial differential equations, and K_i is the permeability corresponding to the initial volumetric moisture content. Philip (1957b) showed that, for large times, the series solution fails to converge; Philip (1969) provided a practical upper limit for convergence (t_c, also known as characteristic time) as:

$$t_c = \frac{S^2}{4\left(K - A\right)^2} \tag{39}$$

Moreover, for small and intermediate times, Philip (1957b) showed that only two terms of the series expansion were needed to approximate most practical cases. Accordingly, the following two-parameter infiltration equation, valid for times smaller than t_c, was proposed:

$$F = St^{1/2} + At \qquad \text{for} \quad 0 \le t \le t_c \tag{40}$$

where S is called the water sorptivity, and A is a parameter determined by numerical integration. The sorptivity is a measure of the capacity of the medium to absorb and desorb water by capillarity. The parameters S and A, as defined by Philip (1957a), depend upon the soil moisture content (θ) and the initial water content (θ_1), and they are given by:

$$S = \int_{\theta_i}^{\theta_s} \phi_1\left(\theta\right) d\theta \tag{41}$$

$$A = \int_{\theta_i}^{\theta_s} \phi_2\left(\theta\right) d\theta + K_i \tag{42}$$

where $\phi_1(\theta)$ and $\phi_2(\theta)$ are functions of moisture content to be determined by a numerical procedure outlined by Philip (1957a). Hence, the parameters S and A are not soil constants but functions of moisture

content. By taking the derivative of Equation (40) with respect to time, the cumulative infiltration capacity can be expressed as (Philip, 1957a):

$$f = \tfrac{1}{2} S t^{-1/2} + A \qquad \text{for} \quad 0 \le t \le t_c \tag{43}$$

For large times, Philip (1957b) derived a closed-form solution for the wetting front at infinite time and showed that the rate of infiltration is given by:

$$f = K \qquad \text{for} \quad t > t_c \tag{44}$$

The hydraulic conductivity (K) in Equation (44) is less than the saturated hydraulic permeability (K_s). This is because surface infiltration traps air within the soil, reducing the net area available for water flow. Upon integration of Equation (44), the cumulative infiltration capacity for large times can be expressed as (Philip, 1987):

$$F = Kt + t_c \big(K - A \big) \qquad \text{for} \quad t > t_c \tag{45}$$

The second term in the right-hand side of Equation (45) is such that the cumulative infiltration evaluated at $t = t_c$ using Equation (45) is the same as the one computed using Equation (40). Hence, a continuous representation of the cumulative infiltration (F) for all times is obtained. The infiltration capacity is computed using Equations (43) and (44), and the corresponding cumulative infiltration capacity is given by Equations (40) and (45), respectively. Thus, the description of infiltration capacity using Philip's model requires the parameters S, A, and K.

Numerous experimental tests and numerical simulations have been performed to validate Philip's series expansion model (e.g., Youngs, 1957; Philip, 1957c; Brutsaert, 1968; Talsma and Parlange, 1972; Warrick et al., 1985; Davidoff and Selim, 1986). The main advantage of Philip's model is that it does not require any assumptions about the moisture profile during the infiltration process.

6.4.3.2 Time of Ponding

For a rainfall event with a constant rate i, ponding occurs when the infiltration capacity equals the rainfall intensity (i.e., $f = i$). Following the definition for ponding time given by Mein and Larson (1973) (i.e., $t_p = F_p/i$), an expression for time of ponding for Philip's model can be obtained. The first step is to obtain an expression of the cumulative infiltration (F) as a function of the infiltration capacity (f). Using Equation (43) to express t as a function of f and placing the resulting equation into Equation (40), an expression for F as a function of the infiltration capacity (f) can be obtained as:

$$F = \frac{S^2 \big(f - A/2 \big)}{2 \big(f - A \big)^2} \qquad \text{for} \quad f \ge A \tag{46}$$

Replacing $f = i$ in Equation (46), the time of ponding is:

$$t_p = \frac{S^2 \big(i - A/2 \big)}{2i \big(i - A \big)^2} \qquad \text{for} \quad i \ge A \tag{47}$$

6.4.3.3 Parameter Determination

Experimental data have confirmed the validity of Philip's theoretical model. However, the numerical procedure outlined by Philip (1957a) to obtain the parameters S and A is too cumbersome for practical

applications. In many cases, the values of S and A are obtained by curve-fitting Philip's model to infiltration data. Although this procedure is valid and expedient, the parameters obtained in this manner are only applicable to specific rainfall rate and initial moisture content.

Significant research has been performed to evaluate the parameters S and A as a function of fundamental soil hydraulic parameters (e.g., Brutsaert, 1968; Parlange, 1975; Brutsaert, 1976; Parlange et al. 1982; Warrick et al., 1985; Philip, 1987). Accordingly, a number of closed-form solutions relating the parameter S to available functional relationships for relative permeability (or diffusivity) and soil suction were derived. For instance, Brutsaert (1968) proposed the following closed-form solution for sorptivity (S) for a relatively dry porous medium:

$$S = \left[2BK_s\left(\Delta\theta_i\right)\big/\eta \right]^{-1/2} \tag{48}$$

where B is the parameter corresponding to Brutsaert's (1966) water retention model (Table 6.2), and the parameter η is the exponent of the Brooks and Corey relative permeability function (Table 6.3). Similarly, Parlange (1975) proposed an expression for sorptivity as:

$$S = \left[\int_0^{\psi_i} \left(\theta_s + \theta(\psi) - 2\theta_i\right) K_s k_r(\psi)\, d\psi \right]^{1/2} \tag{49}$$

For a nearly rectangular moisture profile, Equation (49) may be simplified as:

$$S^2 = 2\Delta\theta_i K_s \int_0^{\psi_i} k_r(\psi)\, d\psi = 2\Delta\theta_i K_s \psi_f \tag{50}$$

See Brutsaert (1976) for a summary of various sorptivity expressions available in the literature.

Using an interpolation function and the general equation of flow into unsaturated porous media, Parlange et al. (1982) derived a new infiltration capacity expression from which the parameter A may be approximated by:

$$A = \left(K - K_i\right)\left(2 - \beta\right)\big/3 \tag{51}$$

where K_i is the initial permeability and β is given by:

$$\beta = \frac{1}{\theta_s - \theta_i} \int_{\theta_i}^{\theta_s} \frac{K_s - K(\theta)}{\left(K_s - K_i\right)}\, d\theta \tag{52}$$

Parlange et al. (1982) showed that β may take values from 0 to 1. Therefore, from Equation (51), it follows that A may take values ranging from $0.33(K - K_i)$ to $0.66(K - K_i)$. To determine the parameter β, the variation of the unsaturated permeability as a function of volumetric moisture content, $k_r(\theta)$, is needed. A summary of the most commonly used relationships for $k_r(\theta)$ is presented in Table 6.3.

Philip (1969) showed that A ranges from $0.38K$ up to $0.66K$ for different moisture profile conditions. The results of numerical simulations as well as of closed-form solutions suggest that the approximation $A = 0.363K$ may be appropriate for soils with a relatively low initial moisture content (Philip, 1987).

It is a common practice to present Equation (43) as Philip's infiltration model without any restriction on time. As experimental evidence indicates that the infiltration capacity f equals K for prolonged rainfall events, it follows from Equation (43) that the parameter A is equal to K for large times. However, it is

emphasized that the series solutions, Equation (43), developed by Philip is valid only for small times ($t < t_c$). Hence, it is not valid to interpret A as the hydraulic conductivity K at saturation.

The use of the generalized expressions to calculate Philip's model parameters (e.g., Equation (49)) is far more complicated than conducting an infiltration test and fitting Equation (43) to the data because it requires the determination of soil relative permeability curves (ψ vs. k_r) and soil water retention curves (ψ vs. S_e). However, the added complexity is balanced by the generality of the S and A parameters obtained.

6.4.3.4 Nonuniform Rainfall Conditions

Philip's model can also be applied to nonuniform rainfall conditions. Substituting Equation (46) into Equation (29), an expression for the time of ponding for a nonuniform rainfall using Philip's model is given by:

$$t_p = t_{j-1} - \frac{I_{j-1}}{i_j} + \frac{S^2\left(i - A/2\right)}{2i\left(i - A\right)^2} \quad \text{with} \quad t_p \le t_c \tag{53}$$

The cumulative infiltration (F) and infiltration capacity (f) equations expressed in incremental form are given by:

$$F_{t+\Delta t} = \begin{cases} F_t + A\,\Delta t - \dfrac{S^2}{2\left(f_t - A\right)} + S\left[\Delta t + \dfrac{S^2}{4\left(f_t - A\right)^2}\right]^{1/2} & \left(t < t_c\right) \\[4mm] F_t + K\,\Delta t & \left(t \ge t_c\right) \end{cases} \tag{54}$$

$$f_{t+\Delta t} = \begin{cases} A + S\left(\dfrac{S + \sqrt{s^2 + 4AF_{t+\Delta t}}}{4F_{t+\Delta t}}\right) & \left(t < t_c\right) \\[4mm] K & \left(t \ge t_c\right) \end{cases} \tag{55}$$

Equations (53), (54), (55) and for Philip's model are equivalent to Equations (29), (31), (33) and for the Green and Ampt model, respectively. Hence, the procedure for nonuniform rainfall conditions described in Section 6.4.2.5 can be used with Philip's infiltration capacity equations.

6.4.4 Parlange's Solution

Assuming that the diffusivity term in Equation (17) approaches a delta function, Parlange (1975a) derived two expressions for cumulative infiltration capacity expressed in terms of the parameters from Philip's model:

$$F = K_1 t + \frac{S^2}{2K_1} Ln\left[1 + \frac{2K_1 F}{S^2}\right] \tag{56}$$

and

$$F = K_1 t - \frac{S^2}{2K_1}\left[\exp\left(-\frac{2K_1 F}{S^2}\right) - 1\right] \tag{57}$$

where the parameter S is the soil sorptivity defined by Philip (1957); K_1 is the difference between the saturated and initial hydraulic conductivity (i.e., $K_1 = K_s - K_i$). Note that Equation (2) is similar in form to the Green and Ampt model.

In the derivation of Equation (56), the unsaturated hydraulic conductivity was assumed to be a constant, whereas in the derivation of Equation (57), the term $\partial K/\partial \theta$ in Equation (17) was assumed to be proportional to the diffusivity term. For other variations of the diffusivity term and the unsaturated hydraulic conductivity, Parlange et al. (1982) derived an implicit cumulative infiltration capacity equation (F) as:

$$\frac{2K_1^2 t}{S^2} = \frac{1}{1-\beta}\left[\frac{2K_1 F}{S^2} - \ln\left[\frac{1}{\beta}\left(\exp\left(\frac{2\beta K_1 F}{S^2} \right) + \beta - 1 \right) \right] \right]$$ (58)

where β is a lumping parameter that is a function of water content and hydraulic conductivity, as defined in Equation (52). This parameter takes into account the moisture distribution with depth. In the derivation of Equations (56), (57), and (58), it is assumed that a negligibly thin layer of water accumulates at the surface after ponding time (i.e., $h_o = 0$).

Based on numerical examples, Smith and Parlange (1978) showed that Equations (56) and (57) are upper and lower bound for most of the closed-form solutions. Moreover, in some cases, the results obtained from Equations (56) and (57) were similar to each other. Using an interpolation parameter ß, Parlange (1978) interpreted Equation (58) as an interpolation between Equations (56) and (57). Equations (56) and (57) can be obtained from Equation (58) by setting the interpolation parameter ß equal to 0 and equal to 1, respectively.

Taking the derivative of F with respect to time, an infiltration capacity (f) is obtained from Equation (58) as:

$$f = K_1 + \beta K_1 \left(1 - \exp\left(\frac{2F\beta K_1}{S^2} \right) \right)^{-1}$$ (59)

It is interesting to note that for extreme values of t (i.e., small and large times), Equation (58) can be approximated by

$$F = St^{1/2} + \tfrac{1}{3}(2-\beta)K_1 t$$ (60)

$$F = K_1 t + \frac{S^2}{2K_1(1-\beta)}\ln\left(\frac{1}{\beta} \right)$$ (61)

which is equivalent to Philip's infiltration capacity model for small and large times, respectively. As shown above, Equation (58) is physically consistent and, for some particular cases, it agrees with both the Green and Ampt model [Equation (57)] and Philip's model [Equation (60)]. Parlange et al. (1982) suggested a value of ß = 0.85 for cases in which $k_r(\theta)$ was not available. Parlange (1985) proposed an extension of Equation (58) to take into consideration ponded conditions with $h_o > 0$ (i.e., h_o cannot be assumed to be negligibly small).

6.4.5 Swartzendruber's Solution

Swartzendruber (1987) derived an expression for the infiltration capacity (f) similar to the Green and Ampt equation, Equation (19), as:

$$f = \frac{K}{Z_f}\left(\frac{K_s}{K}h_o + Z_f + \psi_f\right) = \Delta\theta_i \frac{dZ_f}{dt} + K_n \tag{62}$$

where ψ_f is given by Equation (26), K_s is the hydraulic conductivity at saturation, and K is the resaturated hydraulic permeability ($K < K_s$). The derivation of Equation (62) does not require that we assume a moisture profile represented by a step function as it is in the case of the Green and Ampt model. Integration of Equation (62) yields a cumulative infiltration equation (F) as:

$$F = Z_f\,\Delta\theta_i = Kt + \frac{\Delta\theta_i}{(K-K_i)}\left(K\psi_f + K_s h_o\right)Ln\left[1 + \frac{(F - K_i t)(K - K_i)}{\Delta\theta_i(K\psi_f + K_s h_o)}\right] \tag{63}$$

As Equation (63) is not explicit in either t or F, an iterative procedure is required to solve this infiltration equation. Swartzendruber (1987) provided an alternative semianalytical infiltration capacity equation valid for all times ($0 \le t \le \theta$):

$$F = \frac{S}{A_o}\left[1 - \exp\left(-A_o t^{1/2}\right) + Kt\right] \tag{64}$$

where S is the sorptivity and A_o is a fitting parameter that depends upon the water content at the surface. Swartzendruber and Hogarth (1991) included the effect of ponded water in the sorptivity parameter (S) in Equation (64) as:

$$S^2 = S_p^2 + 2Kh_o\,\Delta\theta_i \tag{65}$$

where S_p is the soil sorptivity as defined by Philip (i.e., for $h_o = 0$).

6.5 Algebraic Infiltration Equations

6.5.1 Introduction

A number of empirical equations have been proposed to describe the infiltration process (e.g., Kostiakov, 1932; Horton, 1940; Smith, 1972) and have shown good agreement with measured infiltration data. The coefficients used in these equations are usually obtained by curve-fitting infiltration test data to the empirical equations. The major drawback of these infiltration equations is that the coefficients are not universal parameters because they depend not only on soil type, but also on other factors such as rainfall infiltration and initial moisture content. In the following paragraphs, the most common empirical infiltration equations are described.

6.5.2 Horton's Equation

6.5.2.1 Introduction

Horton (1940) proposed one of the earliest semiempirical expressions for infiltration capacity (f) as a function of time (t) using an exponential model, as follows:

$$f = f_c + \left(f_o - f_c\right)e^{-\beta t} \tag{66}$$

where f_c is the infiltration rate at saturation and it can be assumed to be equal to the resaturated permeability (K), f_o is the initial infiltration capacity, and $ß$ is an empirical coefficient that can be related to diffusivity. If the parameter f_o is not the initial infiltration capacity, then the value of t must be measured from the time this value occurs. The parameter $ß$ corresponds to the slope of the line f vs. t plotted on semilogarithm paper. In practice, these parameters must be determined experimentally. Although Horton's infiltration equation was originally introduced as an empirical equation, Eagleson (1970) showed that the same expression can be obtained by omitting the gravity term in Equation (17) and assuming the diffusivity term (D) is a constant. The integration of Equation (66) yields an equation for cumulative infiltration as a function of time:

$$F = f_c t + \frac{(f_o - f_c)}{\beta}\left(1 - e^{-\beta t}\right) \tag{67}$$

6.5.2.2 Time of Ponding

For a rainfall event with a constant rate i, ponding occurs when the infiltration capacity equals the rainfall intensity (i.e., $f = i$). As in the case of the Green and Ampt model, an expression for the time of ponding for Hortons's model can also be obtained. The first step is to solve Equation (66) for t and to combine the resulting equation with Equation (67). Then, an expression for the cumulative infiltration as a function of f can be derived as:

$$F = \frac{1}{\beta}\left[f_o - f + f_c \ln\left(\frac{f_o - f_c}{f - f_c} \right) \right] \tag{68}$$

Following the definition given by Mein and Larson (1973) for time of ponding (i.e., $t_p = F/i$), the time of ponding associated with Horton's equation can be computed by setting $f = i$ in Equation (68).

$$t_p = \frac{1}{i\beta}\left[f_o - f + f_c \ln\left(\frac{f_o - f_c}{f - f_c} \right) \right] \tag{69}$$

6.5.2.3 Nonuniform Rainfall Conditions

Horton's model for infiltration capacity can be applied to nonuniform rainfall conditions. By substituting the expression for cumulative infiltration, Equation (68), into the expression for time of ponding for nonuniform rainfall conditions, Equation (29), an expression for time t_p at which ponding begins for a nonuniform rainfall is given by:

$$t_p = t_{j-1} - \frac{I_{j-1}}{i_j} + \frac{1}{i\beta}\left[f_o - i + f_c \ln\left(\frac{f_o - f_c}{i - f_c} \right) \right] \tag{70}$$

It can be shown that the cumulative infiltration (F) and infiltration capacity (f) equations expressed in incremental form are given by:

$$F_{t+\Delta t} = F_t + f_c\, \Delta t + \left(f_t - f_c \right)\left(1 - \exp\left(\beta \Delta t \right) \right) \big/ \beta \tag{71}$$

$$f_{t+\Delta t} = f_t - \beta\left(F_{t+\Delta t} - F_t - f_c\, \Delta t \right) \tag{72}$$

Equations (70), (71), (72) and presented above are equivalent to Equations (29), (31), and (33), respectively, presented in Section 6.4.2.5. The procedure described in detail in Section 6.4.2.5 can be used to model nonuniform rainfall conditions using Horton's model.

6.5.3   Kostiakov's Equation

Kostiakov (1932) proposed the following empirical equation to estimate the soil infiltration capacity:

$$f = at^{-\alpha} \tag{73}$$

where a and α are constants to be determined empirically. Although this equation is successful at modeling laboratory and field data over a considerable range (e.g., Davidoff and Selim, 1986), Equation (73) predicts physically a zero infiltration capacity for prolonged rainfall events. Nevertheless, this empirical model is widely used because of its simplicity.

6.5.4   Smith's Formula

Using a numerical simulation to solve the partial differential equation for unsaturated flow, Equation (13), for various soil types and rainfall patterns, Smith (1972) found that the infiltration capacity curves can be approximated by the following empirical equation:

$$f = f_c + A\left(t - t_o\right)^{-\alpha} \qquad \text{for} \quad t \geq t_p \tag{74}$$

where f_c is the steady-state infiltration rate, α, t_o, and A are empirical coefficients that depend on the soil type, initial moisture condition and rainfall rate, and t_p is the time at which ponding begins. The parameter t_o is the vertical asymptote of the infiltration decay function $(0 \leq t_o \leq t_p)$. Figure 6.11 depicts the meaning of Smith's parameters. The upper limit for Equation (74) is the infiltration rate (i.e., $f \leq i$). The parameters A, α, and f_c are found by curve-fitting experimental data to Equation (74). Smith (1972) normalized the infiltration capacity (f) and rainfall rate (i) with respect to the steady-state infiltration rate, f_c, and the time variables (i.e., t, t_p, t_o) with respect to a time parameter T_o which is defined as:

$$T_o = \left[\frac{A}{\left(1 - \alpha\right)f_c}\right]^{1/\alpha} \tag{75}$$
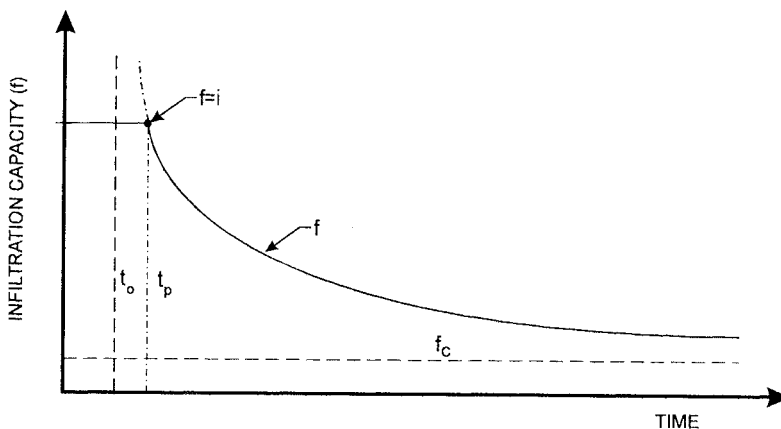
FIGURE 6.11 Smith's infiltration model.

This normalization resulted in similarities between the infiltration capacity equations for different soil types, rainfall rates, and initial water contents. The normalizing parameters T_o and f_c can also be found graphically from measured accumulated infiltration tests (Figure 6.12). The parameter f_c corresponds to the slope of the asymptote of the cumulative infiltration curve, and the parameter T_o corresponds to the value of time when $F = 2f_cT_o$. Smith (1972) expressed Equation (74) in the following dimensionless form:

$$f_* = 1 + (1 - \alpha)(t_* - t_{o*}) \tag{76}$$
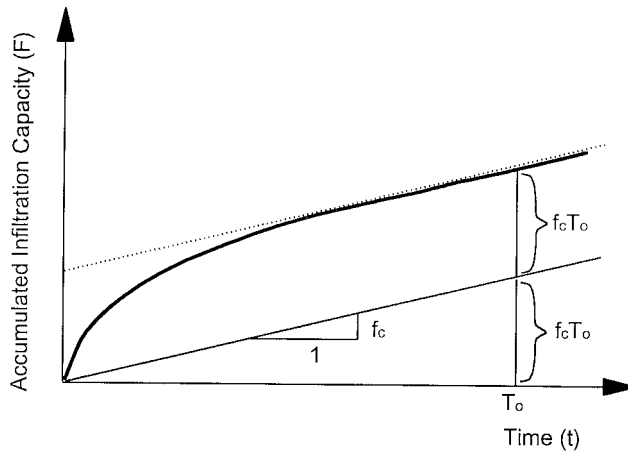
FIGURE 6.12 Graphical determination of normalizing parameters.

where $f_* = f/f_c$, $t_* = t/T_o$, and $t_{o*} = t_o/T_o$. From numerical simulations, Smith (1972) found that the normalized time of ponding, $t_{p*} = t_p/T_o$, can be represented by a power decay curve as:

$$t_{p*} = Bi_*^{-\beta} \qquad \text{with} \quad \beta = 1.62 + 0.407B - 0.027B^2 \tag{77}$$

where $i_* = i/f_c$ and B and β are empirical coefficients. Values of B and β for six different soil types are presented in Table 6.8. Smith (1972) also found that the normalizing time factor, T_o, was linearly related to the effective degree of saturation as:

$$T_o = C\left[S_e(\theta_s) - S_e(\theta_i)\right] \tag{78}$$

where C is a constant to be determined experimentally for each soil.

TABLE 6.8 Parameters for Predicting Dimensionless Time of Ponding

Soil Type	K_s(cm/min)	Porosity	S_r	B	β
Sand	0.2330	0.487	0.12	4.160	2.838
Gravelly sandy loam	0.0381	0.423	0.08	2.420	2.437
Silty clay loam	0.0167	0.520	0.00	1.350	2.125
Colby silt loam	0.0085	0.460	0.15	0.674	1.876
Colby silt loam*	0.0085	0.460	0.15	0.959	1.995
Muren clay	0.0106	0.438	0.00	0.467	1.801

(*) Swelling soil

Adapted from Smith, R.E. 1972. The infiltration envelope: results from a theoretical infiltrometer. *J. of Hydrology.* 17, 1-21.

By expressing the model parameters in terms of soil type and initial moisture conditions, Smith (1972) provided an empirical infiltration equation which can be used for a wide range of rainfall and initial conditions. Model parameters for different soils can be obtained from infiltrometer tests.

Example: An infiltrometer test is performed on a silty loam soil. The initial effective degree of saturation of the silty loam is 60% and the maximum effective degree of saturation is 95%. Assuming that the values of 0.05 mm/min, 0.802, and 0.59 for f_c, A, and α, respectively, were obtained by curve-fitting Equation (74) to the infiltrometer test data, the corresponding infiltration equation is:

$$f = 0.05 + 0.802(t - 4.6)^{0.59} \qquad \text{for} \quad t > 9.35$$

Find the infiltration capacity equation for an initial effective degree of saturation (S_e) of 0.5 subject to the storm shown in columns 1 and 2 of Table 6.9.

TABLE 6.9 Example of Smith's Model

Rainfall (i) (mm/min)	Duration (min)	Time (t) (min)	i_* [-]	t_* [-]	$I_* = i_* t_*$ [-]	$F'_{p*} = t_{p*}i$ [-]
0.2	5	5	4	0.00775	0.0310	0.236
0.5	5	10	10	0.01550	0.1085	0.099
1.0	10	20	20	0.03100	0.4181	0.050
0.5	5	25	10	0.03875	0.4960	0.099

Solution: The first step is to find the normalizing time parameter. From Equation (75), $T_o = [0.802/(0.41)(0.05)]^{0.59} = 500$ min. Knowing that the initial and maximum effective degrees of saturation are 60% and 95%, respectively, the parameter C in Equation (78) can be estimated as C = 500/(0.95 – 0.6) = 1430. Using this value of C and Equation (78), the value of T_o for a 50% of initial effective degree of saturation can be estimated. Thus, $T_o = 1430(0.95 - 0.5) = 645$ min.

The parameter B can be estimated as 0.95 from Table 6.8. By expressing the infiltrometer data in dimensionless form (columns 4, 5, 6, and 7 of Table 6.9), the dimensionless cumulative infiltration at ponding (F_{p*}, column 8 in Table 6.9) for the different dimensionless infiltration rates (i_*) are computed. Ponding takes place whenever $F_{p*} \leq I_*$. In this example, it occurs during the second step (Table 6.9). Thus, by simple interpolation between the values of the first and second row, one may compute the normalized ponding time $t_{p*} = 0.0145$.

Noting that f_* equals i_{p*} for t = t_{p*}, and solving Equation (76) for t_{o*}, the resulting equation yields $t_{o*} = 0.0145 - [(1 - 0.59)/(10 - 1)]^{1/0.59} = 0.0092$. By multiplying by T_o (645 min), the corresponding dimensional quantities of t_p and t_o are 9.35 min and 5.84 min, respectively. Solving Equation (75) for A and replacing the values of α, T_o, and f_c in the resulting expression, yields $A = (635)^{0.59}(1 - 0.59)(0.05) = 0.932$. Thus one can write the infiltration equation as:

$$f_* = 1 + 0.41(t_* - 0.0092)^{0.59} \qquad \text{for} \quad t_* > 0.0145$$

or in dimensional form as:

$$f = 0.05 + 0.932(t - 5.93)^{0.59} \qquad \text{for} \quad t > 9.35$$

6.5.5   Holtan's Solution

Holtan (1961) proposed an empirical function for the evaluation of the soil infiltration capacity (*f*) as:

$$f = f_c + G_i a V_t^b \qquad (79)$$

where G_i is the plant growth index ($0 < G_i \leq 1$), f_c is the percolation rate, a is an index of surface-connected porosity, V_t is the volume of unused moisture storage that is available in the control zone at time t, and the exponent b is usually assumed to be 1.4. The available storage at any time (V_t) can be computed from a water balance equation as:

$$V_t = V_{t-1} - \left(f - f_c - E \right) \Delta t \qquad (80)$$

where V_{t-1} is the volume storage at a previous time ($t-1$), E is the rate of evapotranspiration, and f is the water infiltration. The storage volume V_{t-1} is increased by percolation out of the control zone at a rate f_c and by evapotranspiration at a rate E and is decreased by water infiltration at a rate f. The advantage of Holtan's empirical equation is that the parameters are described in terms of general descriptions of soil type and crop conditions. The main drawback of this expression is the identification of the control depth. The available volume storage, V_t, is based upon the assumed control depth.

6.5.6   The Soil Conservation Service (SCS) Method

An alternative method commonly used in the evaluation of the infiltration capacity of soils is the one developed by the U.S. Soil Conservation Service (SCS). This procedure, derived from a water balance equation, relates the cumulative rainfall (I) to the rainfall runoff volume (R) by the following empirical equation:

$$R = \frac{\left(I - I_A S' \right)^2}{I + \left(1 - I_A \right) S'} \qquad \text{for} \quad I \geq I_A S' \qquad (81)$$

where S' is the maximum volume storage in inches, I_A is the portion of S' that is already occupied, and it is expressed as a fraction of S' (typically 0.2). Values of S' can be determined by the following empirical expression:

$$S' = 1000 / CN - 10 \qquad (82)$$

where C_N is a parameter called the soil curve number. In Equation (82), S' is in inches. For $C_N = 100$, the storage capacity is zero, and the volume of runoff is equal to the cumulative infiltration. Because the *SCS* has classified and assigned values of *CN* to most soils in the United States, the use of this model is widespread. In the absence of evapotranspiration or any other source of abstraction, the amount of water infiltrated can be estimated as the cumulative infiltration minus the volume of runoff (i.e., $I - R$). Figure 6.13 shows the relation between C_N and soil texture for various vegetative covers. Table 6.1 provides typical soil properties for different soil textures.

6.6   Summary

Several closed-form solutions describing the infiltration process are available in the literature. These models can be classified within two main groups, namely physically based models and empirical models. In general, these models do not describe the moisture redistribution process within the unsaturated zone, and more complete solutions or numerical analysis are required to gain a better understanding of the infiltration process. Nonetheless, the equations summarized in this section provide powerful tools for estimating the amount of water infiltration.
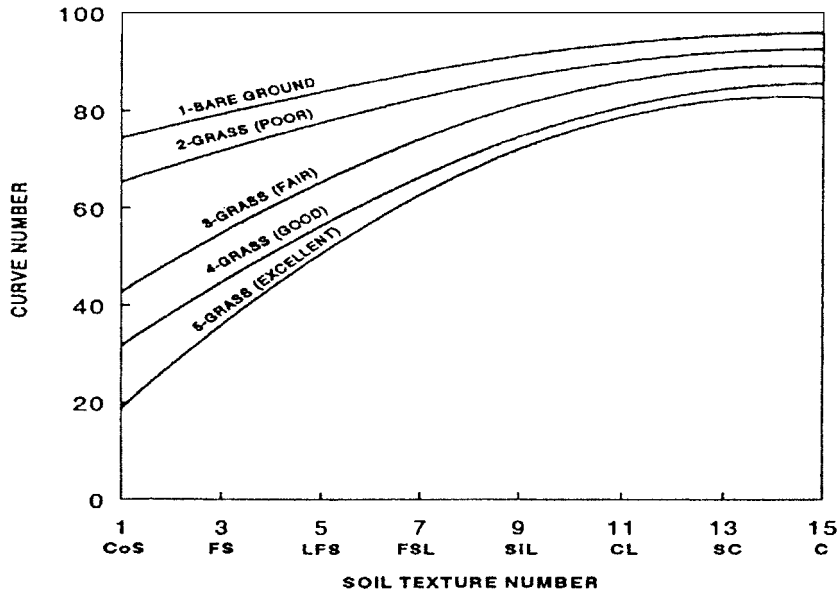
FIGURE 6.13 Relation between SCS curve number and soil texture for various levels of vegetation.

Physically based (e.g., Green and Ampt model) and empirically based models (e.g., Horton's model) have been used successfully to predict the volume of infiltration and runoff. Because spatial variability plays an important role in the accurate prediction of the volume of infiltration, the choice of a particular model depends upon the size of the watershed and the available data. For smaller areas where soil properties such as permeability and initial moisture content can be reasonably estimated, the use of physical-based models becomes more appealing. For larger watersheds, where interpolating, averaging, and weighing may be required to determine the model parameters, empirically based models may be more convenient.

For Further Information

Bear, J. (1979) provides an in-depth analytical study of the influence of infiltration on groundwater recharge. Numerical solutions are also provided.

Delleur (1986) presents a comprehensive summary of commonly used infiltration procedures.

Dingman (1994) provides a general treatment of infiltration. The most commonly used infiltration models as well as concepts of unsaturated flow are described in Chapter 6.

Morel-Seytoux (1973) provides a discussion and analytical solutions for infiltration, taking into consideration two-phase flow theory.

Philip (1969) provides analytical solutions for one-, two-, and three-dimensional infiltration problems. An extensive discussion and solution schemes are presented.

Vauclin et al. (1979b) provides a comprehensive study of the numerical resolution of the infiltration equation.

See the References section for complete citations of these and other works.

References

Babu, D.K. 1976. Infiltration analysis and perturbation methods, 3: Vertical infiltration. *Water Resources Research*, 12, 5, 1019-1024.

Bear, J. 1972. *Dynamics of Fluids in Porous Media*. American Elsevier, New York.

Bear, J. 1979. *Hydraulics of Groundwater*, McGraw Hill Inc., New York.

Bear, J. and Verruijt, A. 1990. *Modelling Groundwater Flow and Pollution*. D. Reidel Publishing Company, Dordrecht, The Netherlands.

Bouwer, H. 1969. Infiltration of water into non-uniform soil. *J. of Irrigation and Drainage Div.*, Proc. of the Am. Soc. of Civil Engineers, 95, IR4, 451-462.

Bouwer, H. 1976. Infiltration into increasingly permeable soils. *J. of Irrigation and Drainage Div.*, Am. Soc. of Civil Engineers, 102, IR1, 127-136.

Bouwer, H. 1966. Rapid field measurements of air entry value and hydraulic conductivity of soil as significant parameters in flow system analysis. *Water Resources Research*, 2, 4, 729-738.

Brakensiek, D.L. 1977. Estimating the effective capillary pressure in the Green and Ampt infiltration equation. *Water Resources Research*. 13, 3, 680-682.

Brooks, R.H. and Corey A.T. 1964. Hydraulic properties of porous media. Hydrology paper No. 3, Civil Eng. Department, Colorado State University, Fort Collins.

Brutsaert, W. 1966. Probability laws for pore-size distribution. *Soil Sci.*, 101, 85-92.

Brutsaert, W. 1968. A solution for vertical infiltration into a dry porous medium. *Water Resources Research*. 4, 1031-1038.

Bumb, A.C., Murphy, C.L., and Everett L.G. 1992. A comparison of three functional forms for representing soil moisture characteristics. *Ground Water*. 30, 2, 177-185.

Childs, E.C. and Bybordi, M. 1969. The vertical movement of water in stratified porous material. 1: Infiltration. *Water Resources Research*. 5, 2, 446-459.

Chow, V.T. 1964. Hydrology and its development, in *Handbook of Hydrology*, Ed. V.T. Chow, Sect. 1, p.4, McGraw Hill Inc., New York.

Davidoff, B. and Selim, H.M. 1986. Goodness of fit for eight water infiltration models. *Soil Sci. Soc. of Amer. J.*, 50, 759-762.

Delleur, J. 1986. Rainfall abstraction and infiltration in nonpoint source pollution, in *Agricultural Nonpoint Source Pollution: Model and Application*. Developments in Environmental Modelling 10, Ed. Giorgini, A. and Zingales, F. Elsevier, Amsterdam, 55-77.

Dingman, L.S. 1994. *Physical Hydrology*. Prentice Hall, Englewood Cliffs, New Jersey.

Eagleson, P.S. 1970. *Dynamic Hydrology*. McGraw Hill Inc., New York.

El-Kadi, Aly. 1985. On estimating the hydraulic properties of soil, Part 2. A new empirical equation for estimating hydraulic conductivity for sands. *Advance Water Resources*, 8, 148-153.

Gardner, W.R. 1956. Calculation of capillary conductivity from pressure plate outflow data. *Soil Sci. Soc. Am. J.* 3, 317-320.

Green, W.H. and Ampt, G.A. 1911. Studies on soil physics. The flow of air and water through soils. *J. Agric. Sc.* 4, 1-24.

Holtan, H.N. 1961. Concept for infiltration estimates in watershed engineering. U.S. Department of Agriculture, Ag. Res. Serv. Public, 41-51.

Horton, R.E. 1940. Approach toward a physical interpretation of infiltration capacity. *Soil Sci. Soc. of Amer. Proc.* 1939 5, 399-417.

Kao, C.S. and Hunt, J.R. 1996. Prediction of wetting front movement during one-dimensional infiltration into soils. *Water Resources Research*. 9, 2, 384-395.

Kostiakov, A.N. 1932. On the dynamics of the coefficient of water percolation in soils and on the necessity for studying it from a dynamic point of view for purposes of amelioration. *Trans. 6th Com. Intern. Soc. Soil Sci. Russian Part A.* 17-21.

Mein, R.G. and Larson, C.L. 1973. Modeling infiltration during a steady rain. *Water Resources Research*. 9, 2, 384-395.

Mein, R.G. and Farrel, D.A. 1974. Determination of wetting front suction in the Green-Ampt equation. *Soil Sci. Soc. Amer. Proc.* 38, 872-876.

Morel-Seytoux, H.J. 1973. Two-phase flows in porous media. in *Advances in Hydroscience.* V.T. Chow, Ed. 9, 119-202.

Morel-Seytoux, H.J. and Khanji, J. 1974. Derivation of an equation of infiltration. *Water Resources Research*. 10, 4, 795-800.

Morel-Seytoux, H.J. 1983. Infiltration affected by air, seal, crust, ice and various sources of heterogeneity special problems. *Proc. Symposium on Advances in Infiltration*. 34-47.

Parlange, J.-Y. 1975. On solving the flow equation in unsaturated soils by optimization: horizontal infiltration. *Soil Sci. Soc. of Amer. Proc.* (1975). 5, 415-418.

Parlange J.-Y., Haverkamp R., and Touma J. 1985. Infiltration under ponded conditions: 1 Optimal analytical solution and comparison with experimental observations. *Soil Sci. Soc. of Amer. Proc.* 5, 139, 4, 415-418.

Parlange J.-Y., Lisle I., Braddock R.D., and Smith R.E. 1982. The three-parameter infiltration equation. *Soil Sci. Soc. of Amer. Proc.* 5, 133, 6, 337-341.

Philip, J.R. 1957a. The theory of infiltration: 1. The infiltration equation and its solution. *Soil Science.* 83, 345-357.

Philip, J.R. 1957b. The theory of infiltration: 2. The profile of infinity. *Soil Science.* 83, 435-448.

Philip, J.R. 1957c. The theory of infiltration: 3. Moisture profile and relation to experiment. *Soil Science.* 84, 163-178.

Philip, J.R. 1957d. The theory of infiltration: 4. Sorptivity and algebraic infiltration equations. *Soil Science.* 84, 257-264.

Philip, J.R. 1969. *Theory of Infiltration. Advances in Hydroscience*, Vol. 5, Academic Press, New York.

Philip, J.R. 1984. Steady Infiltration form circular cylindrical cavities. *Soil Science Society of America Journal*, 48, 270-278.

Philip J.R. 1987. The infiltration joining problem. *Water Resources Research.* 23, 12, 2239-2245.

Rawls, W.J. and Brakensiek, D.L. 1982. Estimating soil water retention from soil porosities. *J. Irrig. and Drain. Engrg.* ASCE, 1082, 166-171.

Rawls, W.J., Brakensiek, D.L., and Soni, B. 1983a. Agricultural management effects on soil water processes. Part I: Soil water retention and Green and Ampt infiltration parameters. *Trans ASAE.* 266, 1747-1752.

Rawls, W.J., Brakensiek, D.L., and Miller N. 1983b. Predicting Green and Ampt infiltration parameters from soil data. *J. Hydr. Engrg.* ASCE, 1091, 62-70.

Rawls, W.J. and Brakensiek, D.L. 1986. Comparison between Green and Ampt infiltration parameters and curve number runoff predictions. *Trans ASAE.* 296, 1597-1599.

Schroeder, P.R., Dozier, T.S. Zappi, P.A., McEnroe, B.M., Sjostrom, J.W., and Peyton, R.L. 1994. *The Hydrologic Evaluation of Landfill Performance HELP Model: Engineering Documentation for Version 3.* EPA/600/9-94/168b, U.S. Environmental Protection Agency Risk Reduction Engineering Laboratory. Cincinnati, Ohio.

Smith, R.E. 1972. The infiltration envelope: results from a theoretical infiltrometer. *J. of Hydrology.* 17, 1-21.

Smith, R.E. and Parlange, J.-Y. 1978. A parameter efficient hydrologic infiltration model. *Water Resources Research.* 14, 533-538.

Swartzendruber, D. 1987. Rigorous derivation and interpretation of the Green and AMPT equation, *Proceedings Intl. Conf. on Infiltration Development and Application*, Y-K Fok, Ed. Water Resources Center, University of Hawai, Honolulu, 28-37.

Swartzendruber, D. and Hogarth, W.L. 1991. Water infiltration into soil response to ponded-water head. *Soil Science Society of America Journal.* 55, 6, 1511-1515.

Talsma, T. and Parlange, J.-Y. 1972. One-dimensional vertical infiltration. *Aust. J. of Soil Res.* 10, 143-150.

Taylor, D.W. 1948. *Fundamentals of Soils Mechanics.* John Wiley & Sons, New York.

Tolikas, P.K., Tolikas, D.K., and Tzimopolus, C.D. 1983. Vertical infiltration of water into unsaturated soil with variable diffusivity and conductivity. *J. of Hydrology.* 624,321-332.

Vachaud, G., Vauclin, M., and Khanji, D. 1973. Effects of air pressure on water flow in an unsaturated stratified vertical column of sand. *Water Resources Research.* 9, 1, 160-173.

Van Genuchten, M.T. 1980. A close form solution for predicting the hydraulic conductivity of unsaturated soil. *Soil. Sci. Soc. Am. J.* 44, 892-898.

Van Mullen, J.A. 1991. Runoff and peak discharge using Green-Ampt infiltration model. *Journal of Hydraulic Engineering*, 117, 3, 354-370.

Vauclin, M., Khanji, D., and Vachaud, G. 1979a. Experimental and numerical study of a transient two dimensional unsaturated-saturated water table recharge problem. *Water Resources Research*. 15, 5, 1089-1101.

Vauclin. M., Haverkamp, R., and Vachaud, G. 1979b. *Etude de la Résolution Numerique de l'Equation de l'Infiltration d'Eau en Milieu non Saturé*. Presses Universitaires de Grenoble, France, 165 pp.

Warrick A.W., Lomen D.O. and Yates S.R. 1985. A generalized solution to infiltration. *Soil Science Society of America Journal*. 49, 34-38.

Warrick, A.W. 1991. Numerical approximations of darcian flow through unsaturated soil. *Water Resources Research*. 27, 6, 1215-1222.

Whisler, F.D. and Bouwer, H. 1970. Comparisons of methods for calculating vertical drainage and infiltration in soils. *J. of Hydrology*. 101,1-19.

Youngs, E.G., 1957. Moisture profiles during vertical infiltration. *Soil Science*. 4, 84, 283-290.

Glossary

Diffusion   Spreading of a contaminant in groundwater due to molecular diffusion.

Hydraulic Gradient   Difference in piezometric head divided by the flow length between two points.

Piezometric Head   Sum of the pressure head and the elevation head.

Pressure head is the water pressure divided by γ_w (unit weight of water), and elevation head is the vertical distance from a reference elevation.

Resaturated Hydraulic Conductivity   Maximum value of hydraulic conductivity during an infiltration process. When infiltration takes place, air is trapped within the soil reducing the net area for water flow.

Suction   Water pressure in the unsaturated zone. Suction is negative (i.e., less than atmospheric pressure).

Tortuosity   Factor that takes into consideration the flow tortuous path of a porous media (see Bear [1979] for a more detailed description).

7
Modeling the Movement of Water and Solute Through Preferential Flow Paths

Rony Wallach
The Hebrew University of Jerusalem

Tammo S. Steenhuis
and J.-Yves Parlange
Cornell University

Abstract

Mathematical models are widely used in soil physics and hydrology for predicting water percolation and water-aided transport of solutes and contaminants through the unsaturated zone. Most of these models are based on combining Darcy's law with a dispersive/diffusion equation (convective-dispersive equation), resulting in transport of the bulk of the solutes at *one* average velocity. Several studies have detected high concentrations of pesticides in groundwater shortly after the first rainfall as a consequence of preferential flow paths carrying the pesticides and other solutes directly from the surface to the groundwater. Modeling this fast moisture and solute transport through the unsaturated zone is complicated because of different forms and types of preferential flow and the complicated networks of interconnected pathways in the soil which can transmit water and its solutes at varying velocities (and, therefore, cannot be simulated with the convective transport equation). Preferred pathways in soils may result from biological and/or geological activity (e.g., macropores, earthworm burrows, channels consisting of highly conductive media) or from farm management practice (e.g., conservation tillage). Preferential flow paths are also found in homogeneous and layered sandy soils due to the instability of the wetting front. Such preferred paths may transmit water and its solutes at higher velocities than those predicted by Darcy's theory.

   A better understanding of the factors affecting the percolation of water and the mobility, extent, and nature of chemical transport are required to improve modeling efforts and management practice. In this chapter, the nature and impact of preferential flow on moisture and solute transport is discussed, with emphasis on structured soils.

7.1 Introduction

Recent experimental research in soil structure and non-ideal chemical transport is forcing scientists to refine their analysis of the movement and fate of contaminants in the soil and consider processes which have become important when small concentrations of highly toxic chemicals must be accurately accounted for in order to protect groundwater reserves from contamination (Stagnitti et al., 1995). The mobility of the chemicals and contaminants in well-structured soils is affected by the continuity as well as the size of the pores. Networks of interconnected, highly-conductive pathways which result from biological and geological activity, such as subsurface erosion, faults and fractures, shrink-swell cracks, animal burrows, wormholes, and decaying roots may be responsible for transmitting moisture, solutes, and contaminants to groundwater at faster velocities than those predicted by theory based on the convective-dispersive equation in which the water velocity is represented by the local average values (Bear, 1972; Sposito et al., 1986). In most cases, the volume of water within the preferential paths is much lower than the volume of the stagnant fluid, and only a small part of the solution (but very significant when the solute is toxic at low concentrations) may be moving within well-defined preferential paths ahead of the main flow. For example, Smettem and Collis-George (1985) found that a single continuous macropore can conduct more water than a surrounding soil sample 10 cm in diameter. In structured soils, the convective-dispersive model provides unsatisfactory results under field conditions, and the interchange of fluids between the matrix and the fracture or large pore must be modeled explicitly (Steenhuis et al., 1990; Stagnitti et al., 1995; Parlange et al., 1996, 1998).

Contrary to widely held belief, even in relatively uniform sandy or water-repellent soil profiles, water does not necessarily take an average flow path and flow can take place through fingers caused by dynamic instabilities of the wetting front as was first documented by Hill and Parlange (1972). Further research has led to an understanding of many of the mechanisms involved in this phenomenon (Parlange and Hill, 1976; Raats, 1973; Philip, 1975; Hillel and Baker, 1988; Glass et al., 1989a,b,c, 1991; Selker et al., 1992a,b, 1993; Liu et al., 1993, 1994a,b). The phenomena are not only restricted to the laboratory but can also be observed under field conditions. Fingered flow in a field was first noted in the Connecticut valley (Starr et al., 1978) and later by Glass et al. (1989a) on Long Island. Hendrickx and Dekker (1991), using dye-tracing techniques, found that only 10 to 20% of the vadose zone was actually involved in the transport of the dye for a water-repellent soil in Ouddorp (Netherlands). Also, Ritsema and Dekker (1993a,b) and Dekker and Ritsema (1994) found that preferential flow was more the rule than the exception during the late summer and early fall in the Netherlands. Finally, Rice et al. (1991) in the midwestern U.S. found that solutes and herbicides travel at velocities 1.6 to 2.5 times faster than traditional water balance models of piston flow in sandy soils with little or no structure.

Finger properties can be derived from soil physical theory. Analysis of the wetting front instability suggests that the finger diameter, d, is given by

$$d = \frac{\kappa S_F^2}{K_F(\theta_F - \theta_i)} \left[\frac{1}{1 - \dfrac{Q}{K_F}} \right] \tag{1}$$

where θ_i is the water content ahead of the finger and θ_F, K_F, and S_F are the water content, conductivity, and sorptivity at the wettest spot in the finger, respectively, usually close to the tip. Q is the imposed flux at the surface and κ is a constant, which is equal to π in two dimensions (Parlange and Hill, 1976) or 4.8 in three dimensions (Glass et al., 1991). Hillel and Baker (1988) and Baker and Hillel (1990) suggest that at the wettest spot in the tip of the finger, the soil-water potential is at its water entry value. Indeed this seems to be a general result when the soil is initially dry. However, in subsequent infiltration events (using the finger paths created earlier), hysteresis is becoming a dominant factor and the finger becomes dryer (Liu et al., 1994b). The form of Equation (1) is physically quite intuitive. $S_F^2/(\theta_F - \theta_i)$ represents

the sorptive properties of the soil, and the larger this term is, the wider the finger. K_F, on the other hand, represents the gravity effect, which drives the instability. The larger it is, the thinner the finger. Finally, increasing the imposed flux results in wider fingers. However, the main effect of the imposed flux is on the density of the fingers. The larger the Q, the closer the fingers are spaced. Hence, if Q approaches K_F, the whole space is necessary to carry the water entering the soil. At that point all the fingers merge so that $d \to \infty$.

Another mechanism for initiation of fingers in field soils is interbedded coarse layers in fine sand. Coarse layers concentrate the water from a large area into a finger-like structure. Kung (1990a,b) coined the term *funnel flow* for this phenomenon. He found that water and solutes flowed through less than 10% of the total soil matrix between depths of 3.0 and 3.6 m and less than 1% between 5.6 and 6 m. Funnel flow was also observed in Delaware (Boll et al., 1997) and in Massachusetts along the Connecticut River where the water flowed over coarse layers under low flux conditions and then broke through at higher fluxes (Weiler et al., 1997; Steenhuis et al., 1998).

In contradiction to the fingers or funnel flow — formed by the dynamic instability in sandy soils, and where the shape, size, and distribution of the fingers depend on the imposed flux — the preferential flow paths in structured loamy or clayey soils are intrinsically defined as part of the soil profile, with shape, size, and distribution independent of the fluids involved. In this chapter, we will concentrate on the structural defined flow paths. One of the main difficulties in modeling preferential flow in the "structural" large-pore system is the interaction of the preferentially moving chemicals with the soil matrix. To predict the solute distribution in structured soils, one ideally should know the exact spatial distribution, shape, size, and connectiveness of the preferential flow paths. This information is only available for a few soil cores (Heijs et al., 1996), and not for field soils. Consequently, simplified models that overcome this lack of information are being used to interpret the vast amount of solute concentration data from field and laboratory experiments (Flühler et al., 1996).

In this chapter, we will give a survey of the various mathematical approaches for predicting solute breakthrough curves in structured soils. Preferential flow in sandy and layered soils have been presented elsewhere (Glass and Nichol, 1996; Ritsema et al., 1996). Dual-porosity models with and without a distribution layer are discussed first followed by multi-pore group models.

7.2 Dual-Porosity Models — Soil Structure Similar Throughout the Profile

Flow through cracks and fissures was initially studied in relation to the exploitation of nonhomogeneous groundwater and petroleum reservoirs. The naturally fractured reservoir can be taken as a collection of porous rock blocks, which are usually called the matrix, and an interconnected system of fracture planes. Most of the fluid resides in the matrix, where it moves very slowly, but the small amount of fluid rapidly flowing in the fractures can have a profound effect on the overall solute flow. A common approach to modeling this system mathematically is the dual-porosity concept, first presented by Barenblatt et al. (1960) and Warren and Root (1963). It is based on the assumption that naturally fractured reservoirs behave as two porous structures rather than one. The difference between the dual-porosity media and the usual porous media is that the "solid part" is in itself permeable and both matrix and fracture flow are defined at each point of the matrix. Thus, fluid flow is a combination of general macroscopic motions and takes place within the matrix blocks and also around them. The dual-porosity models have also been used to describe flow and transport in structured porous media where less permeable soil aggregates are analogous to the matrix (Passioura, 1971).

An intensive use in the dual-porosity concept has been made to model the non-ideal breakthrough curves (BTCs) measured at the outlet of soil columns in the laboratory and soil cores from the field. Non-ideal BTCs are asymmetric and characterized by early breakthrough and an extensive tail which indicates that some part of the fluid is moving ahead of a front calculated by the convective-dispersive equation (CDE) (Biggar and Nielsen, 1962; Krupp and Elrick, 1969; Starr and Parlange, 1977; van

Genuchten and Wierenga, 1977; Rao et al., 1980; Nkedi-Kizza et al., 1983; Seyfried and Rao, 1987). The intra-aggregate porosity of aggregated soil and the dead-end pores in structured soils are considered as the low conductivity matrix, while the inter-aggregate porosity consists of the pathways where preferential flow takes place. The advection within the regions with the relatively low hydraulic conductivity is usually assumed zero, therefore, these domains act as sink/source components and the dual-porosity, dual-velocity model becomes a mobile–immobile model.

The dissolved chemical transport in a single porosity model is usually described by the CDE (Biggar and Nielsen, 1967)

$$\frac{\partial(\theta_m c_m)}{\partial t} = \frac{\partial}{\partial z}\left[D_e \frac{\partial c_m}{\partial z}\right] - \frac{\partial(J_w c_m)}{\partial z} \tag{2}$$

where the subscript m stands for mobile, z is the direction of flow, and c is the solute concentration. D_e is the effective dispersion coefficient, combining the influence of diffusion in the dissolved phase with the long-time dispersion representation of the effect of small-scale convection on the mixing of solute about the center of motion (Jury et al., 1991). J_w is the water flux in the mobile paths. When water flow is steady and there is no chemical exchange with immobile porosity, Equation (2) reduces to

$$\frac{\partial c_m}{\partial t} = D\frac{\partial^2 c_m}{\partial z^2} - v\frac{\partial c_m}{\partial z} \tag{3}$$

where $D = D_e/\theta_m$ is the effective diffusion-dispersion coefficient and $v = J_w/\theta_m$ is the pore water velocity. Analytical solutions of Equation (3) for many initial and boundary conditions can be found in van Genuchten and Alves (1982).

In the mobile–immobile model, exchange takes place between the two regions and the chemical flux to/from the immobile porosity appears as a sink term in Equation (2). For the convenience of mathematical presentation, the equations which follow will be written for saturated mobile and immobile porosities, and Equation (3) becomes

$$\frac{\partial c_m}{\partial t} + \frac{\theta_{im}}{\theta_m}\frac{\partial \bar{c}_{im}}{\partial t} = D\frac{\partial^2 c_m}{\partial z^2} - v\frac{\partial c_m}{\partial z} \tag{4}$$

where the subscript im stands for immobile, $\theta_{im} = \theta - \theta_m$ is the immobile water content, and $\bar{c}_{im}$ is the uniformly distributed solute concentration in the immobile porosity. The dissolved chemical exchange between the mobile and immobile porosities can be represented by two conceptually different mechanisms: a local equilibrium and a rate-limited exchange. The choice between the two is dictated by how fast the rate of exchange is compared with the transport rate in the mobile porosity. The exchange has been usually modeled by a diffusion equation based on Fick's law or by employing a first-order mass transfer approximation which, for particular cases, can be derived directly from a diffusion-based model (Rao et al., 1980; van Genuchten and Dalton, 1986; Parker and Valocchi, 1986). The rate-limited first-order exchange between the mobile and immobile porosities may be written as

$$\theta_{im}\frac{\partial \bar{c}_{im}}{\partial t} = k\left[c_m - c_{im}\right] \tag{5}$$

where k is the rate coefficient describing the mass transfer process between the mobile and immobile regions. A brief review of mass transfer models and their characteristics is presented in Brusseau and Rao (1989) and Brusseau et al. (1994).

Two time scales are involved in the mobile–immobile problem: the convective-dispersive transport in the mobile pores and the rate-limited dissolved chemical exchange between the two regions. For cases where the microscopic processes are rapid enough in relation to the fluids flow rate in the mobile pores, the local equilibrium assumption (LEA) was suggested to replace the rate-limited exchange (Equation [5]) (Rubin, 1983). This replacement provides conceptual and mathematical simplifications. Models based on LEA have proven to be useful in a variety of solute transport applications, but there is an increasing body of experimental evidence indicating that the LEA may not be applicable under many conditions of interest in the study of solute transport (Valocchi, 1985; Jennings and Kirkner, 1984; Bahr and Rubin, 1987). These studies aimed to express conditions that are based on the flow and transport parameters under which the LEA is valid and provide small deviations from the rate-limited model. The identification of these parameter values was based on comparing concentration profiles, which were calculated by the LEA model and solutions to the rate-limited model, obtained by different mathematical methods, for different scenarios and conditions. Wallach (1998) took advantage of the different time scales for the two rate-limited processes and defined a set of dimensionless variables that were substituted into the mass balance equations (Equations [4] and [5]). For cases when the characteristic time scale for the lateral solute exchange is faster than the convective transport time scale (as used to define the conditions for LEA), the ratio between the two time scales forms a small dimensionless parameter that multiplies the time derivative in the dimensionless version of Equation (5). The method of matched asymptotic expansions can then be used to form a uniform solution for this singular perturbations problem. The leading-order approximation, obtained when this parameter approaches zero, yields the solution obtained by the LEA and is valid only for the longer times ("outer" solution). A transition period exists near $t = 0$ within which the initial concentration equilibrates. The LEA validity, as well as the length of the transition period during which the initial mobile and immobile concentrations reach local equilibrium depends on how small the ratio between the two time scales is.

A closer look at the process of chemical transfer from the moving solute within the crack into the stagnant matrix brings forth the conclusion that the distribution of solute concentration within the moving solute varies only slightly across the crack, perpendicular to the flow, except at the interface with the matrix. The solute transferred across the interface is doing so by diffusion (Tang et al., 1981; Neretnieks et al., 1982; Neretnieks, 1983; Moreno and Rasmuson, 1986; Fujikawa and Fukui, 1990; Maloszewski and Zuber, 1990; and others). A mass balance equation is then written for each medium, and the mass continuity is preserved via the boundary condition at the interface between these two media. The mass balance equation for the preferential path (Tang et al., 1981) is

$$\frac{\partial c_m}{\partial t} - D_m \frac{\partial^2 c_m}{\partial z^2} + v \frac{\partial c_m}{\partial z} = -q \tag{6}$$

where D_m is the effective dispersion coefficient for the crack, and q (M L^{-3} T^{-1}) is the solute flux (concentration per unit time) from the crack to the matrix. The lateral transport in the matrix is described by

$$\frac{\partial c_{im}}{\partial t} = D_{im} \frac{\partial^2 c_{im}}{\partial x^2} \tag{7}$$

where x is the length in the matrix perpendicular to the flow direction in the crack, z, and D_{im} is the effective diffusion coefficient for the matrix. The boundary condition at the interface between the crack and the matrix ($x = 0$) is that local equilibrium is valid

$$c_m(z,t) = c_{im}(x = 0, z, t) \tag{8}$$

Solutions of this model reveal that the matrix diffusion can be viewed as a beneficial safety mechanism in soil contamination problems involving a concentrated source in fractured media. The dissolved chemicals are not spreading instantaneously throughout the matrix as assumed in the mobile–immobile models. The degree of spreading depends strongly on the difference in the time scales characterizing the dispersive-convective transport in the fracture and the transport by diffusion in the matrix. For example, in a system with large matrix porosity and low fracture fluid velocity the contamination of the matrix soil will be much less than in a system with low matrix porosity and high fracture fluid velocity.

Piquemal (1993) pointed out that it is erroneous, in general, to assume that the concentrations in the crack and the matrix are in local equilibrium. Developing the equations by the volume-averaging technique, he concluded that it applies only when the mean concentration of the flowing fluid is equal to the concentration in the immobile fluid at the boundary separating them. In many models where local equilibrium between the mobile and immobile regions is assumed, the solute flow is described by a one-dimensional convective-dispersive model which *a priori* ignores the lateral concentration gradient in the mobile region. Therefore, the assumption is self-consistent in these models.

In the mobile–immobile model, the preferential paths are not physically determined but have much larger volume than the immobile volume and are assumed to be uniformly distributed throughout the soil profile. On the contrary, in the case of preferential path models the preferential paths are defined and the immobile volume is much larger than the volume of the preferential paths. The second difference is the dissolved chemical distribution in the matrix. Assuming a uniform concentration distribution with the lateral direction within the matrix (Equation [5] for the MIM model) rather than a diffusion front that propagates laterally as a function of $\sqrt{t}$ t is equivalent to the assumption that concentration changes taking place at any point along the interface between the mobile and immobile domains equilibrate instantaneously with the whole stagnant solution in the horizontal direction. Since the chemical transfer between the mobile and immobile domains is controlled by the concentration difference (Equation [5]), this assumption has a major effect on the values of k and D estimated by fitting the MIM model to measured BTCs. The fitting-routine of Parker and van Genuchten (1984) is an often-used tool to estimate the MIM model parameters from measured BTCs. For field soils where dye tracers have shown that only few pores carry most of the solutes, the assumption of instantaneous equilibrium throughout the entire stagnant pore group is poor. The solute cannot move through the large lateral distances between the preferential paths by diffusion in the time frame of the transport in the preferential paths. Application of the MIM model to BTCs measured for soils with well-defined preferential paths provided poor predictions (Anderson and Bouma, 1977).

An important question for modeling flow and transport in preferential paths is whether the flow is laminar or turbulent. It has a great effect on the concentration distribution within the preferential path, especially in the direction perpendicular to the flow. Water flow in soil pores has traditionally been assumed laminar (following Poiseuille's law), but Chen and Wagenet (1992) argue that the flow is turbulent for macropores larger than 0.2 mm in diameter. Ela et al. (1991) measured rates of water flow into individual pores (size not given) ranging from 0 to 2550 ml/s, and Wang et al. (1991, 1994) have measured inflow rates of 1 to 17 ml/s into individual worm or ant holes (1.5 to 3.6 mm diameter). Reynolds numbers calculated for these open-channel pores and rates are out of range of laminar flow and into the transitional flow range (Chow, 1964). Turbulent flow theory was developed for large conduits, and laminar flow theory was developed for capillary-size pores. Large macropores are in between these sizes and may be only partially filled (Beven and Germann, 1981); thus, one may argue that there is no reason to assume one or the other type of flow.

Wallach and Steenhuis (1998) proposed a model for such field soils with well-defined preferential paths. The model follows the main assumptions made for the MIM model, i.e., the velocity of the mobile fluid in the well-defined preferential paths is constant and the dissolved chemical exchange between the saturated stagnant pore group and the mobile pore group is a rate-limiting process. For the sake of simplicity, the dissolved chemical transport in the crack is modeled by the kinematic-wave approach rather than by the convective-dispersive equation, and analytic solutions are obtained. The model takes advantage of the characteristic time scales for chemical transfer through the boundary layer along the

interface between the mobile and immobile domains being faster than the lateral transport of the chemicals within the stagnant porosity by diffusion. This assumption is expressed mathematically as multiplying the immobile concentration in the mass balance equations by a small dimensionless parameter, $\epsilon << 1$. The mass balance equation for the preferential paths becomes (Wallach and Steenhuis, 1998)

$$\frac{\partial c_m}{\partial t} + v \frac{\partial c_m}{\partial z} = -k_m \left[c_m - \varepsilon c_{im} \right] \tag{9}$$

where $k_m = k/\theta_m$. The boundary condition at the inlet to the preferential path is

$$c_m(0,t) = c_i(t) \tag{10}$$

Postulating that the concentration within the matrix is not laterally uniform and that it varies significantly only within a short lateral distance from the interface during the travel time of the mobile fluid to the outlet, the solute concentration is assumed to be uniformly distributed only within a thin matrix layer, θ_{im} (defined as the pore volume in this layer per sample volume), located at the interface between the two pore groups. The mass balance equation for this matrix layer is

$$\frac{\partial c_{im}}{\partial t} = k_{im} \left[c_m - \varepsilon c_{im} \right] \tag{11}$$

where $k_{im} = k/\theta_{im}$.

Equations (9) and (11) form a regular perturbations problem whose solution can be expressed by the series

$$c_m(z,t) = \sum_{i=0}^{n} \varepsilon^i A_i = A_0 + \varepsilon A_1 + \varepsilon^2 A_2 + \ldots \tag{12}$$

$$c_{im}(z,t) = \sum_{i=0}^{n} \varepsilon^i B_i = B_0 + \varepsilon B_1 + \varepsilon^2 B_2 + \ldots \tag{13}$$

Substituting Equations (12) and (13) into Equations (9 through 11) and equating terms that are multiplied by equal powers of ϵ, sequences of equations for the different orders of approximations are obtained.

The Zeroth-Order Approximation: This approximation, known also as the leading-order approximation, is obtained by equating terms multiplied by ϵ^0. The mass balance equations for the preferential path and matrix solutes are

$$\frac{\partial A_0}{\partial t} + v \frac{\partial A_0}{\partial z} = -\frac{k}{\theta_m} A_0 \tag{14}$$

$$\frac{\partial B_0}{\partial t} = \frac{k}{\theta_{im}} A_0 \tag{15}$$

The right-hand side of Equation (14) expresses the chemical transfer from the preferential paths to the matrix when the back transfer due to the accumulation of chemicals in the matrix is ignored. As a result, Equations (14) and (15) include one dependent variable each and can be solved independently, as opposed to Equations (4) and (7), which contain two dependent variables and, therefore, should be solved

simultaneously. The higher-order approximations in the following will account for the dissolved chemical flux from the matrix.

This problem is solved by applying the Laplace transform with respect to time. The transformed mobile concentration is represented as

$$a_i = \int_0^\infty A_i e^{-st}\, dt; \qquad b_i = \int_0^\infty B_i e^{-st}\, dt \tag{16}$$

where s is the transformation variable. Application of Equation (16) to Equations (14 through 15) leads to a pair of ordinary differential equations with a boundary condition in the Laplace domain. The equation for the preferential path with zero initial concentration is

$$sa_0 + v\frac{da_0}{dz} = -\frac{k}{\theta_m} a_0 \tag{17}$$

and its boundary condition at the soil surface in the Laplace domain is

$$\bar{c}_0(s) = \frac{c_i}{s} \tag{18}$$

The solution of Equation (17) with Equation (18) in the Laplace domain is

$$a_0(z,s) = \frac{c_i}{s}\exp\left[-\frac{kz}{\theta_m v}\cdot\frac{sz}{v}\right] \tag{19}$$

The solution for A_0 in the time domain is obtained through the application of the inverse Laplace transform to Equation (19)

$$A_0(z,t) = c_i \exp\left[-\frac{kz}{\theta_m v}\right]\cdot u\left(t-\frac{z}{v}\right) \tag{20}$$

where $u(t)$ is the unit step function. The leading solution for the mobile domain is of an advancing sharp front with a simultaneous transfer of dissolved chemicals to the active matrix layer. The chemical transfer depends only on the mobile concentration and its concentration at any point z as a function of the travel time from the inlet to this point.

Applying the Laplace transform Equation (16) to Equation (15) with initial zero immobile concentration and substituting $a_0(z,s)$ from Equation (19) gives

$$b_0(z,s) = \frac{k}{\theta_{im}}\frac{c_i}{s^2}\exp\left[-\frac{kz}{\theta_m v}-\frac{sz}{v}\right] \tag{21}$$

whose inverse Laplace transform gives the zero*th*-order approximation of the chemical distribution in the matrix. Since the local concentration in the matrix depends on the local chemical concentration in the preferential path, it also varies with time and depth

$$B_0(z,t) = \frac{c_i k}{\theta_{im}} \exp\left[-\frac{kz}{\theta_m v}\right]\left\{t - \frac{z}{v}\right\} \cdot u\left(t - \frac{z}{v}\right) \tag{22}$$

The First-Order Approximation: The first-order approximation is obtained by equating terms multiplied by ϵ when Equations (12) and (13) are substituted into Equations (9 through 11).

The mass balance equations for the preferential path and matrix solutions are

$$\frac{\partial A_1}{\partial t} = v\frac{\partial A_1}{\partial z} = -\frac{k}{\theta_m}\left(A_1 - B_0\right) \tag{23a}$$

$$\frac{\partial B_1}{\partial t} = \frac{k}{\theta_{im}}\left(A_1 - B_0\right) \tag{23b}$$

Applying the Laplace transform Equation (16) to Equation (23) gives

$$a_1(z,s) - \frac{k^2}{\theta_m \theta_{im}}\frac{c_i}{s^2}\exp\left[-\frac{kz}{\theta_m v} - \frac{sz}{v}\right] \tag{24}$$

whose inverse Laplace transform is

$$A_1(z,t) = \frac{c_i k^2}{\theta_m \theta_{im}}\exp\left[-\frac{kz}{\theta_m v}\right]\frac{z}{v}\left[t - \frac{z}{v}\right] \cdot u\left(t - \frac{z}{v}\right) \tag{25}$$

Higher order approximations can be similarly obtained and solved, but the first two terms in the asymptotic expansion provide a sufficiently good approximation. The breakthrough concentration from a soil column of length L is obtained by substituting Equations (20) and (25) into Equation (12) after replacing z by L, as

$$c_1(L,t) = c_i \exp\left[-\frac{kL}{\theta_m v}\right]\left\{1 + \varepsilon \frac{c_i k^2}{\theta_m \theta_{im}}\frac{L}{v}\left[t - \frac{L}{v}\right]\right\} \cdot u\left(t - \frac{L}{v}\right) \tag{26}$$

The shape of the BTC predicted by Equation (26) is similar to BTCs measured for initially drained undisturbed soil materials with well-defined preferential paths (Anderson and Bouma, 1977; White et al., 1984, 1986; Kluitenberg and Horton, 1990). These BTCs are characterized by a breakthrough at very early times followed by a sharp concentration increase afterward when the chloride concentration increases rapidly from 0 to a higher value after which the concentration increases at a lower rate until the initial concentration is reached. Note that the sharp increase in concentration immediately after the initial breakthrough and sharp break later on is due to the elimination of the diffusion-dispersion term from the mass balance equation for the preferential paths.

Equation (26) was fitted to BTCs that were measured by Anderson and Bouma (1977) at the bottom of saturated columns filled with structured soils that were initially drained. The measured BTCs are taken from Figure 3 in Anderson and Bouma (1977) and are shown together with the predicted BTCs in Figure 7.1. After determining the velocity v by dividing the column length by the breakthrough time, k is estimated by the concentration where the BTC changes its shape from the vertical increase to a moderate increase, and ϵ is determined from the slope at which the BTC reaches the initial concentration. The estimated k is close to the values calculated independently for this BTC by Skopp et al. (1981). Equation (26) fits well the measured BTCs (Figure 7.1).
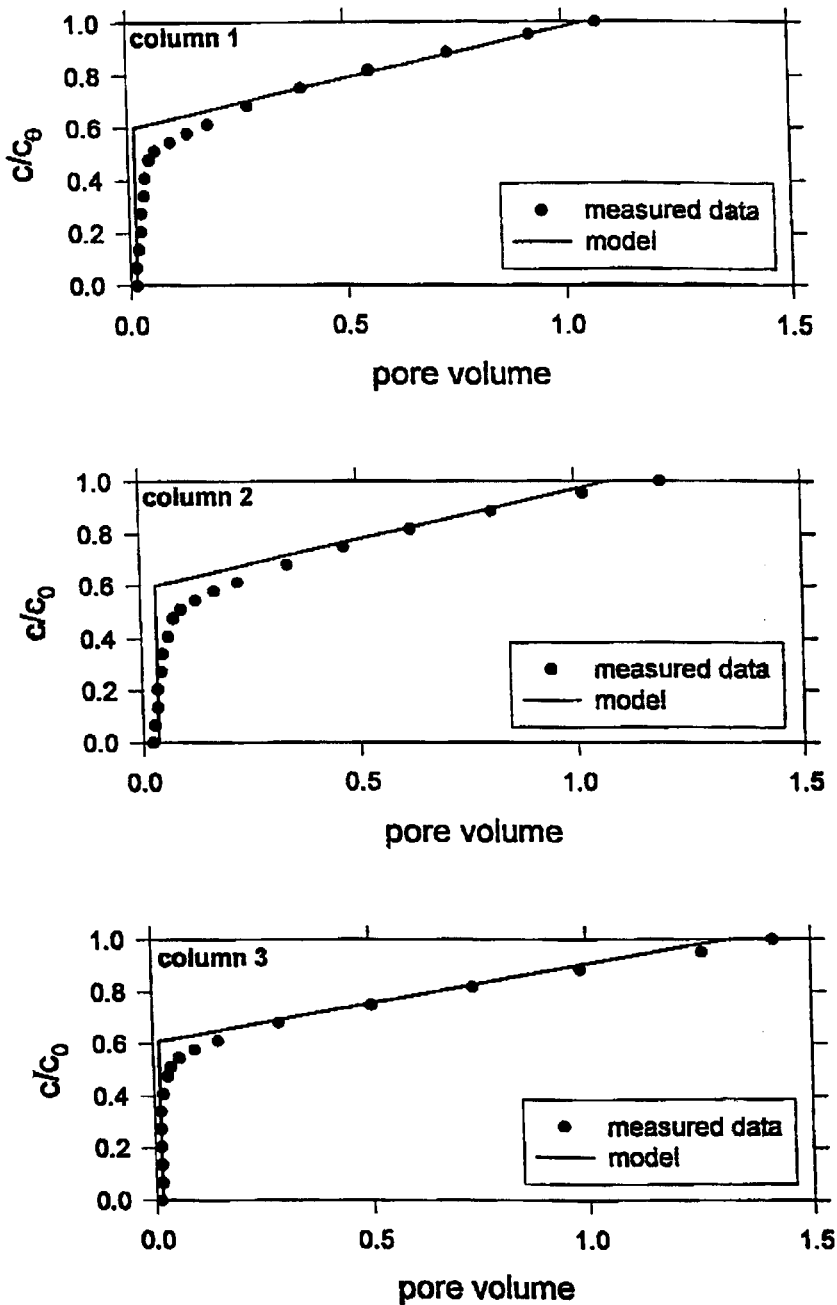
FIGURE 7.1 Measured (Data from Anderson, J.L. and Bouma, J. 1977. Water movement through pedal soil. I. Saturated flow. *Soil Sci. Soc. Am. J.* 41:413-418.) and predicted BTCs at the outlet of three initially drained columns.

7.3 Dual-Porosity Models with a Distribution Layer at the Soil Surface

The models presented in the previous section assume that the soil surface layer has a structure similar to that of the subsoil and preferential flow takes place along the whole soil pedon. Recent observations

(Steenhuis et al., 1994; Parlange et al., 1988; Ritsema and Dekker, 1995; and others) have shown that a thin (2 to 30 cm) soil surface layer (distribution layer) exists in many cases at the soil surface where flow takes place throughout this layer. The precipitation in the distribution layer is funneled into the preferential paths in the subsoil (Figure 7.2). The distribution layer typically has a higher conductivity (especially in the horizontal direction) than the soil profile below it. Its thickness depends on various factors such as tillage, vegetation, and bioturbation. One of the approaches to model the distribution layer is to assume that it is a linear reservoir (Steenhuis et al., 1994). A linear reservoir is equivalent to a well-mixed reservoir where the momentary solute concentration is uniformly distributed throughout its depth. The lumped mass balance equation for the distribution layer is

$$\frac{dc_0}{dt'} = -\frac{J_w}{\theta l} c_0 \qquad (27)$$
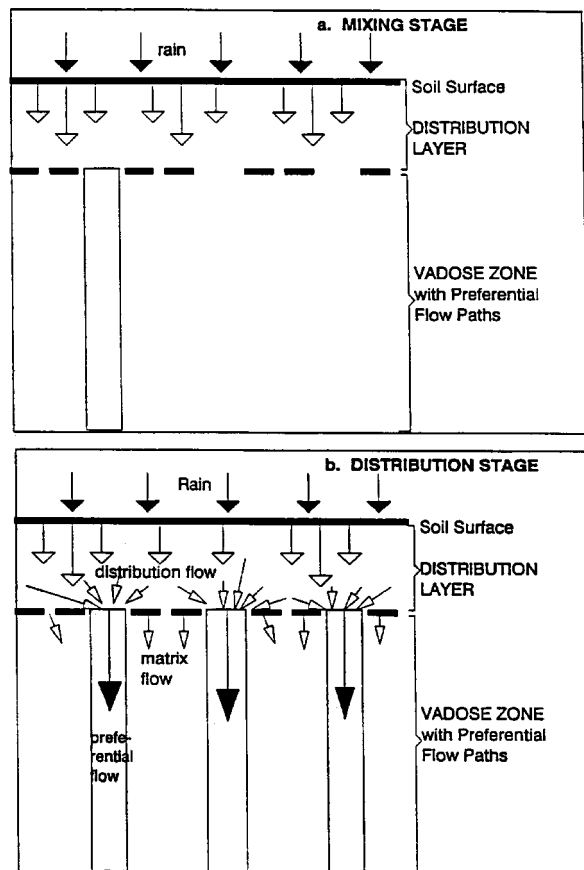
FIGURE 7.2 Conceptual framework of preferential flow model. (a) The mixing stage in the distribution layer and (b) the distribution into the preferential flow paths.

where θ is its saturated water content, J_w is the flux leaving the distribution zone, and l is the depth of the distribution layer. $c_0(t)$ is the dissolved chemical concentration in the distribution layer. Steenhuis et al. (1994) assumed that there was no interaction in preferential flow paths between the preferentially moving water and the matrix. The concentration of solutes leaving the distribution layer (and in preferential flow paths) after the solute is completely mixed in the distribution zone is

$$c_0 = \frac{M}{l\theta} \exp\left(-\frac{\int J_w \, dt}{l\theta}\right) \tag{28}$$

where M is the amount of solute applied per unit area. Figure 7.3 shows the results of an experiment carried out in the Cornell Orchard on a Rhinebeck loamy soil with a 30-cm distribution zone on top of a glacial till with macropores in the structural cracks. A pulse of chloride was added on day zero followed by daily irrigations of 3 cm. The concentration in the preferential flow paths was measured in wick and gravity pan samplers at 60 cm depth. Although the fit is quite good, it is obvious that it cannot fit the BTCs such as were obtained by Anderson and Bouma (1977). A more complicated model is needed in which the distribution layer as well as the interaction in the subsoil between macropores and matrix (introduced above) is taken into account.
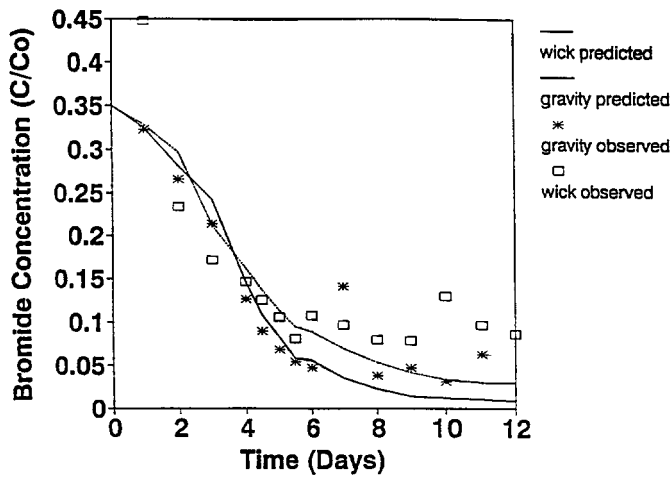
FIGURE 7.3 Observed and predicted outflow concentrations of the wick and gravity pan samplers for the Cornell Orchard site.

In order to do so, assume a time interval, τ, when the solutes that enter the soil surface enter the preferential paths. Thus, the concentration that enters the preferential flow paths (not assuming instantaneous mixing) is

$$c_0(\tau) = c_i\left(1 - e^{-\alpha\tau}\right) \tag{29}$$

to be used in Equation (10). Following the small perturbation solution above, Wallach and Steenhuis (1998) obtained the zero*th* and first approximations concentration in the preferential paths at point z

$$A_0(z,t) = c_i \exp\left[-\frac{kz}{\theta_m v}\right]\left\{1 - \exp\left[-\alpha\left(t + \tau - \frac{z}{v}\right)\right]\right\} \cdot u\left(t - \frac{z}{v}\right) \tag{30}$$

$$A_1(z,t) = \frac{c_i k^2}{\theta_m \theta_{im}} \exp\left[-\frac{kz}{\theta_m v}\right]\frac{z}{v}\left\{t - \frac{z}{v} - \frac{\exp[-\alpha\tau]}{\alpha}\left[1 - \exp\left[-\alpha\left(t + \tau - \frac{z}{v}\right)\right]\right]\right\} \cdot u\left(t - \frac{z}{v}\right) \tag{31}$$

This model was also validated by using BTCs that were measured by Anderson and Bouma (1977) for initially saturated columns containing structured soils. Both, measured (Figure 1 in Anderson and Bouma, 1977) and predicted BTCs (Equation [31]) are shown in Figure 7.4. The initially-saturated experimental runs have much longer initial breakthrough times and BTCs tailing compared with the initially drained runs where the now air-filled large pores were drained quickly. Emptying these paths by draining the columns prior to the runs enables the infiltrating solute in the initially drained columns to flow directly through them and bypass most of the columns matrix. A very good fit is obtained between the model output and the measured BTCs. While τ expresses the travel time in the vertical direction within the distribution layer prior to entering the preferential path, the average residence time in the distribution layer, α^{-1}, expresses both the vertical flow and the horizontal travel time to the preferential paths beneath it. In all three columns (Figure 7.4) $\alpha^{-1} > \tau$, which indicates that the average residence time of the converging flow in the distribution layer is larger than the time for downward vertical flow. Note that the dispersion in the BTCs in Figure 7.4 is only due to the interaction between the mobile and stagnant solutions and is not the result of the dispersion during the transport in the preferential paths, since this transport was modeled by the kinematic-wave model (Equation [4]) using $D = 0$ in the convective-dispersive equation.

7.4   Multi-Pore Group Models

The main purpose of the models described above was to find the relationship between the flowing fluid in the preferential paths and the stagnant fluid within the matrix and its effect on the BTC. As will be described now, the matrix fluid, which was assumed to be immobile in the models above, may be mobile as well and appear later at the outlet with lower concentration during a longer time period. However, even if the matrix fluid flows with much lower velocities than the preferential flow, its main effect on the BTC at earlier times will be as if it were stagnant. Dye infiltration experiments by Omoti and Wild (1979) revealed rapid transport through pores too small to be considered macropores. A secondary micropore system (smaller than the macropores), also contributes to preferential flow under variably saturated–unsaturated conditions and was also observed by Germann and Beven (1981). Roth et al. (1991) found that rainfall-driven movement of a single surface-applied tracer pulse separated into a slowly moving pulse through the soil matrix and a series of fast preferential flow pulses. The sporadic preferential flow pulses resulted in rapid solute transport to 220 cm and accounted for 58% of the total mass. However, flow through the secondary pore system within the soil matrix accounted for the remaining 42% and moved the solutes to 84 cm. Jardin et al. (1990) monitored rainfall-driven vertical movement of bromide through a large soil pedon and found a rather continuous movement through a small-pore region that lasted for several days following rainfall, while rapid movement occurred as discrete pulses through the large-pore region during events. Jury and Flühler (1992) noted that substantial water flow may be occurring in the surrounding matrix as well as in the preferential flow region. They also noted that solutes partition into a rapid or preferential flow region and a slower but still mobile matrix flow region, each of which may embody a smaller but significant degree of water flow, and concluded that the dual-porosity model is unsatisfactory for representing media wherein transport occurs both in the preferential flow region and in the bulk matrix.

To model preferential flow through a soil profile with multi-pore groups is to assume that the soil is a composite of flow paths with different velocities where mixing may or may not take place among the different flow paths. Dependent on the scale of observation, such flow regions may be pores, fissures, wetted fingers, fine porous matrix, or any region with a relatively well-defined velocity. If the flow regions are non-interacting, meaning that the boundaries of each region are impermeable for solute transfer, an individual solute parcel remains for its entire travel time within the region it entered. If the particle can move laterally through the boundary of its flow region, it experiences the velocities of other flow regions. The velocity, assumed constant within a flow region, can be a deterministic or a random variable distributed over the entire flow cross-section. The fate of a particle that enters a bundle of hypothetically parallel flow regions at the inlet end is now considered.
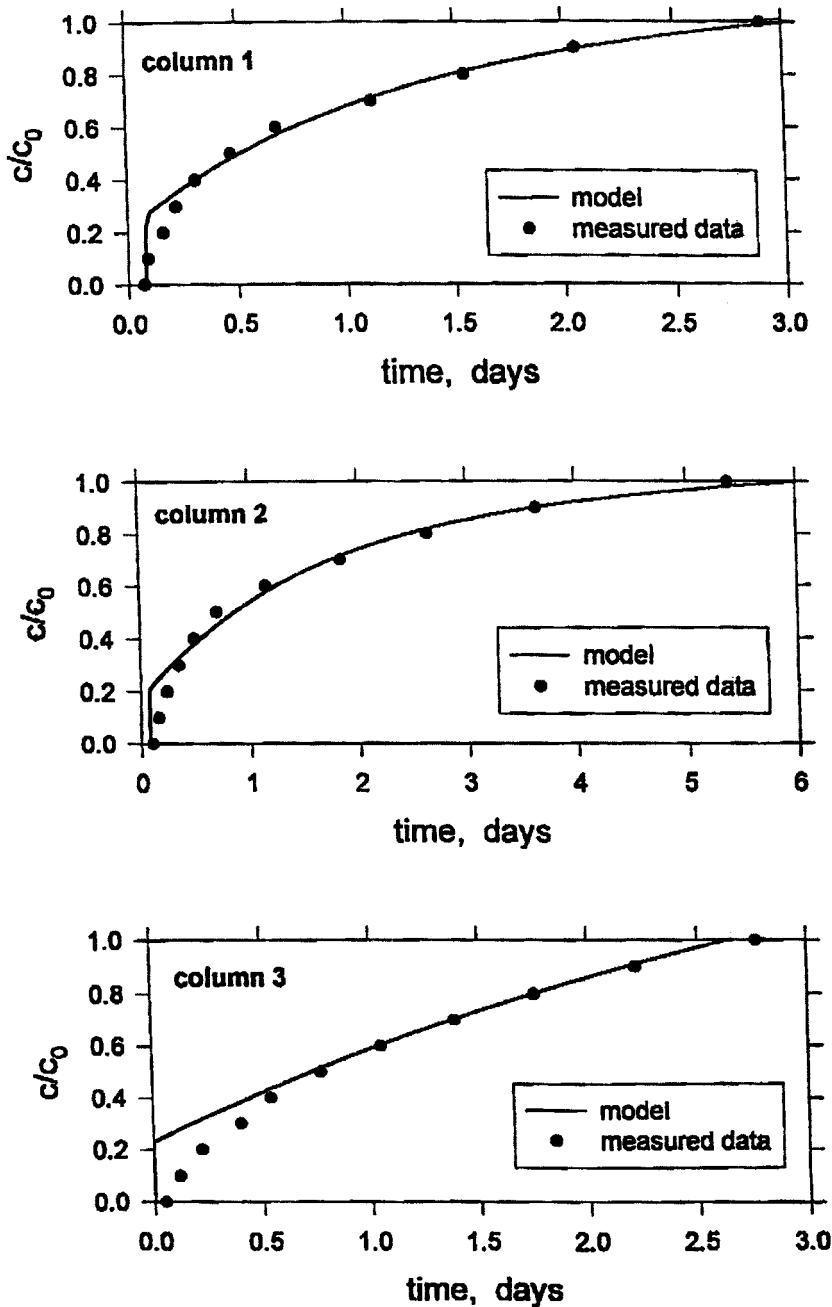
FIGURE 7.4 Measured (Data from Anderson, J.L. and Bouma, J. 1977. Water movement through pedal soil. I. Saturated flow. *Soil Sci. Soc. Am. J.* 41:413-418.) and predicted BTCs at the outlet of three initially saturated columns.

The straightforward extension of the mobile–immobile model is a two-region convective-dispersive model in which mass is transported from one region to the other in response to differences in the residence fluid concentration. Approximate solutions to this problem have been obtained by Skopp et al. (1981), Van Duin and van der Zee (1986), and Gerke and van Genuchten (1993). Skopp et al. (1981) developed a model where the liquid-filled pores are partitioned into two distinct pore size classes. One region represents macro- or inter-aggregate porosity, and the other represents a mobile matrix porosity. The

regions have different hydraulic and transport characteristics and an interaction coefficient characterizes the linear transfer between them. A small interaction coefficient was assumed, and a regular perturbation method was used to solve the model equations. The leading-order solution ignores interaction between the two regions, and flow and transport take place in each one of them independently. The interaction between the two regions appears in the higher-order approximations. If the interaction coefficient is sufficiently large, the convective-dispersive equation is obtained. Comparing the model's output to the measured data of Anderson and Bouma (1977) showed that the interaction term between the two regions is higher than anticipated and the concentration distribution just after the breakthrough time was not well described. Gerke and van Genuchten (1993) also used the two-mobile porosity concept to separate non-steady water flow and transport equations for the matrix and macropore. The chemical exchange between the two regions is described by transfer coefficients which are often functions of the pore size.

Lindstrom and Boersma (1971) separated the pore space of a saturated soil into a finite number of subdomains with characteristic pore sizes. They estimated the flow velocity in each pore domain by Poiseuille's law and calculated total transport by superimposing analytical solutions of the convection-dispersion equation. Their model was not successful in describing experimental breakthrough curves (Rao et al., 1976) since it did not include any mixing of solutes between the pore classes. Steenhuis et al. (1990) followed a similar conceptual framework and proposed a mathematical model that considers many flow and transport domains. Extensions and improvements to the model were later proposed by Stagnitti et al. (1991, 1995) and Durner and Flühler (1996). The preferential flow paths were characterized by taking piecewise linear approximations of the hydraulic conductivity, resulting in a number of pore groups with a mobile transport velocity in each.

Exchange processes between pore groups require a calibration to a particular field or laboratory experiment. Unlike the usual modeling assumptions applied in the CDE, the concentration of solutes in the percolating water is dependent on the varying rate of applied water and the time period between rainfall and chemical application. This is achieved by relating the solute flux to the water flux. Thus, transient field conditions can be simulated. The model can be applied to both large-scale field experiments and small-scale laboratory experiments. As the model has been fully described elsewhere (Steenhuis et al., 1990; Parlange et al., 1996; Stagnitti et al., 1995), we will present only a brief description here.

The total amount of moisture, $\theta(x,t)$ in the soil at time, t, and point, x, is the sum of all individual moisture contents for each capillary bundle, p

$$\theta(x,t) = \sum_{p=0}^{N} \theta_p(x,t) \tag{32}$$

where θ_p is the individual moisture content for the pth pore group. The maximum amount of moisture that each group can hold and transmit is $\Delta M_p = M_p - M_{p-1}$, where M_p and M_{p-1} are various moisture contents representing upper and lower limiting values for the pth group and are a function of the size of pores in each group. When $\theta_p = \Delta M_p$, then all the pathways for the pth group are completely saturated with soil moisture. The group's moisture content, θ_p is a function of the vertical percolation rate, q_p, effects of precipitation and evapotranspiration, and loss or gain of moisture from interaction and exchanges with other groups. The mass balance equations for the water and solute in the different pore groups and the algorithm to get the moisture and BTCs for different scenarios are presented in detail in Stagnitti et al. (1995). Since the networks of flow paths in the soil are interconnected, the exchange of moisture and solutes between the various capillary bundles ranges from a complete exchange of moisture and solute to no exchange or mixing. The basic idea behind the different mixing degrees is the existence of a virtual "common" pool where moisture can be extracted from any group and placed into it, while any fraction of moisture may move back from the common pool into the pth group after mixing. The values of the mixing coefficients depend on the soil type and would normally have to be found by careful experimentation. The preferential flow model has been successfully applied to a number of field and

laboratory experiments (e.g., Stagnitti et al., 1991, 1995; Steenhuis et al., 1990; Nijssen et al., 1991; Parlange et al., 1996).

Recently, Skopp and Gardner (1992), who generalized the model of Skopp et al. (1981), and Durner (1992) presented an analytical model for a continuous distribution of local flow velocities perpendicular on an axis orthogonal to the main transport direction. Skopp and Gardner (1992) used the method of moments to relate the dispersion coefficient to flow velocity, which can then be substituted into the convective-dispersive equation for large interaction between the flow domain perpendicular to the main flow direction.

Durner and Flühler (1996) expanded the previous model by presenting a numerical solution to a conceptual multi-domain model for the transport of solutes in saturated–unsaturated soils. The pore space in this model is represented by a continuous distribution of pore size on a virtual structural coordinate, orthogonal to the spatial coordinate. The chemical transport is simulated by the convective-dispersive equation. Following the method of Steenhuis et al. (1990) a two-step procedure for convective transport and mixing was applied. In the first step, the transport in each pore domain along the spatial axis is solved. In the second step, solute exchange between the pore domains is modeled by a finite difference approximation of the diffusion equation. The model presents the gradual change of convection-dominated transport to convective-dispersive transport with time and depth. The length scale where this change takes place depends on the pore size distribution of the porous medium, on the intensity of lateral mixing, and on the degree of saturation.

For Further Information

Preferential flow is a fast-developing science. Summaries of the work can be found in conference proceedings and special issues of journals. A recent special issue of *Geoderma* (Steenhuis et al., 1996) covers the range from field experiments, where preferential flow was found, to theoretical development. In an older publication, Gish and Shirmohammadi (1991) summarize the preferential flow research up to 1991. A very good overview article about the occurrence of preferential flow was written by Flury (1996). The other references on the different types of preferential flow are mentioned in the text.

Gish, T.J. and A. Shirmohammadi, Eds. 1991. *Preferential Flow*. Proceedings of American Society of Agricultural Engineers National Symposium, December 16-17, 1991, Chicago, IL.

Flury, M. 1996. Experimental evidence of transport of pesticides through field soils. A review. *J. Env. Qual.* 25:25-45.

Steenhuis, T.S., C.J. Ritsema, and L.W. Dekker, Eds. 1996. *Fingered Flow in Unsaturated Soil: From Nature to Model*. Special Issue of *Geoderma*, 70.

References

Anderson, J. L. and Bouma, J. 1977. Water movement through pedal soil. I. Saturated flow. *Soil Sci. Soc. Am. J.* 41:413-418.

Bahr, J. M. and Rubin, J. 1987. Direct comparison of kinetic and local equilibrium formulations for solute transport affected by surface reactions. *Water Resources Res.* 23:438-452.

Baker, R. S. and Hillel, D. 1990. Laboratory tests of a theory of fingering during infiltration into layered soils. *Soil Sci. Soc. Am. J.* 54:20-29.

Barenblatt, G. I., Zheltov, I. P., and Kochina, I. N. 1960. Basic concepts in the theory of seepage of homogeneous liquids in fissured rocks (strata). *Prikl. Mat. Mekh.* 24:852-864; *J. Appl. Math. Mech.* 24:1286-1303.

Bear, J. 1972. *Dynamics of Fluids in Porous Media*. Elsevier, New York.

Beven, K. and Germann, P. 1981. Water flow in soil macropores. II. A combined flow model. *J. Soil Sci.* 32:15-29.

Biggar, J. W. and Nielsen, D. R. 1962. Miscible displacement. II. Behavior of tracers. *Soil Sci. Soc. Am. Proc.* 26:125-128.

Biggar, J. W. and Nielsen, D. R. 1967. Miscible displacement and leaching phenomenon. *Agron. Monog.* 11:254-274.

Boll, J., Selker, J. S., Shalit, G., and Steenhuis, T. S. 1997. Frequency distribution of water and solute transport properties derived from pan sampler data. *Water Resources Res.* 33:2655-2664.

Brusseau, M. L. and Rao, P. S. C. 1989. Modeling solute transport in structured soil: A review. *Geoderma.* 46:169-192.

Brusseau, M. L., Gerstl, Z., Augustijn, D., and Rao, P. S. C. 1994. Simulating solute transport in an aggregated soil with the dual-porosity model: Measured and optimized parameter values. *J. Hydrology.* 163:187-193.

Chen, C. and Wagenet, R. J. 1992. Simulation of water and chemicals in macropore soils. Part 2. Application of linear filter theory. *J. Hydrology.* 130:127-149.

Chow, V. T. 1964. *Handbook of Applied Hydrology.* McGraw-Hill, Inc., New York.

Dekker, L. W. and Ritsema, C. J. 1994. Fingered flow: The creator of sand columns in dune and beach sands. *Earth Surface Processes and Landforms.* 19:153-164.

Durner, W. 1992. New concepts for the functional description of soil hydraulic properties and for the modeling of pollutant transport during stormwater infiltration into multi porosity soils, in *Hydrological and Pollutational Aspects of Stormwater Infiltration.* Eds. M. Grottger and W. Schilling. Proc. European Workshop, Schriftenreibe der EAWAG, 3, CH-8600 Dübendorf, 125-143.

Durner, W. and Flühler, H. 1996. Multi-domain model for pore-size dependent transport of solutes in soils, in *Fingered Flow in Unsaturated Soil: From Nature to Model.* Eds. T.S. Steenhuis, C.J. Ritsema, and L.W. Dekker. Special Issue of *Geoderma.* 70:281-298.

Ela, S. D., Gupta, S. C., Rawls, W. J., and Moncrief, J. F. 1991. Role of earthworm macropores formed by *Aporrectodea tuberculata* on preferential flow of water through a Typic Hapludoll, in *Preferential Flow.* Eds. T.J. Gish and A. Shirmohammadi. Proc. American Society of Agricultural Engineers National Symposium, December 16-17, 1991, Chicago, Illinois. 68-76.

Flühler, H., Durner, W., and Flury, M. 1996. Lateral solute mixing processes — A key for understanding field-scale transport of water and solutes, in *Fingered Flow in Unsaturated Soil: From Nature to Model.* Eds. T.S. Steenhuis, C.J. Ritsema, and L.W. Dekker. Special Issue of *Geoderma.* 70:165-183.

Fujikawa, Y. and Fukui, M. 1990. Adsorptive solute transport in fractured rock: Analytical solutions for delta-type source conditions. *J. Contam. Hydrol.* 6:85-102.

Gerke, H. H. and van Genuchten, M. T. 1993. A dual-porosity model for simulating preferential movement of water and solutes in structured porous media. *Water Resources Res.* 29:305-319.

Germann, P. and Beven, K. 1981. Water flow in soil macropores. *J. Soil Sci.* 32:1-13.

Glass, R. J., Oosting, G. H., and Steenhuis, T. S. 1989a. Preferential solute transport in layered homogeneous sands as a consequence of wetting front instability. *J. Hydrology.* 110:87-105.

Glass, R. J., Parlange, J.-Y., and Steenhuis, T. S. 1989b. Wetting front instability. 1. Theoretical discussion and dimensional analysis. *Water Resources Res.* 25:1187-1194.

Glass, R. J., Steenhuis, T. S., and Parlange, J.-Y. 1989c. Wetting front instability. 2. Experimental determination of relationships between system parameters and two-dimensional unstable flow field behavior in initially dry porous media. *Water Resources Res.* 25:1195-1207.

Glass, R. J., Parlange, J.-Y., and Steenhuis, T. S. 1991. Immiscible displacement in porous media: Stability analysis of three-dimensional, axisymmetric disturbances with application to gravity-driven wetting front instability. *Water Resources Res.* 27:1947-1956.

Glass, R. J. and Nichol, M. J. 1996. Physics of gravity fingering of immiscible fluids within porous media: An overview of current understanding and selected complicating factors, in *Fingered Flow in Unsaturated Soil: From Nature to Model.* Eds. T.S. Steenhuis, C.J. Ritsema, and L.W. Dekker. Special Issue of Geoderma. 70:117-132.

Heijs, A. W. J., Ritsema, C. J., and Dekker, L. W. 1996. Three-dimensional visualization of preferential flow patterns in two soils, in *Fingered Flow in Unsaturated Soil: From Nature to Model.* Eds. T.S. Steenhuis, C.J. Ritsema, and L.W. Dekker. Special Issue of Geoderma. 70:87-101.

Hendrickx, J. M. H. and Dekker, L. W. 1991. Experimental evidence of unstable wetting fronts in homogeneous non-layered soils, in *Preferential Flow.* Eds. T.J. Gish and A. Shirmohammadi. Proc. American Society of Agricultural Engineers National Symposium, December 16-17, 1991, Chicago, Illinois, 22-31.

Hill, D. E. and Parlange, J.-Y. 1972. Wetting front instability in layered soils. *Soil Sci. Soc. Am. J.* 36:697-702.

Hillel, D. and Baker, R. S. 1988. A descriptive theory of fingering during infiltration into layered soils. *Soil Sci.* 146:51-56.

Jardin, P. M., Wilson, G. V., and Luxmoore, R. J. 1990. Unsaturated solute transport through a forest soil during rain storm events. *Geoderma.* 46:103-118.

Jennings, A. A. and Kirkner, D. J. 1984. Instantaneous equilibrium approximation analysis. *J. Hydraul. Eng.* 110:1700-1717.

Jury, W. A., Gardner, W. R., and Gardner, W. H. 1991. *Soil Physics.* John Wiley & Sons, New York.

Jury, W. A. and Flühler, H. 1992. Transport of chemicals through soil: Mechanisms, models and field applications. *Adv. Agron.* 47:141-201.

Kluitenberg, G. J. and Horton, R. 1990. Effect of solute application method on preferential transport of solutes in soil. *Geoderma.* 46:283-297.

Krupp, H. K. and Elrick, D. E. 1969. Density effects in miscible displacement experiments. *Soil Sci.* 107:372-380.

Kung, K.-J. S. 1990a. Preferential flow in a sandy vadose soil. 1. Field observations. *Geoderma.* 46:51-58.

Kung, K.-J. S. 1990b. Preferential flow in a sandy vadose soil. 2. Mechanism and implications. *Geoderma.* 46:59-71.

Lindstrom, F. T. and Boersma, L. 1971. A theory on the mass transport of previously distributed chemicals in water saturated sorbing porous medium. *Soil Sci.* 111:192-199.

Liu, Y., Bierck, B. R., Selker, J. S., Steenhuis, T. S., and Parlange, J.-Y. 1993. High intensity x-ray and tensiometer measurements in rapidly changing preferential flow fields. *Soil Sci. Soc. Am. J.* 57:1188-1192.

Liu, Y., Steenhuis, T. S., and Parlange, J.-Y. 1994a. Closed-form solution for finger width in sandy soils at different water contents. *Water Resources Res.* 30:949-952.

Liu, Y., Steenhuis, T. S., and Parlange, J. -Y. 1994b. Formation and persistence of fingered flow fields in coarse grained soils under different moisture contents. *J. Hydrology.* 159:187-195.

Maloszewski, P. and Zuber, A. 1990. Mathematical modeling of tracer behavior in short-term experiments in fissured rocks. *Water Resources Res.* 26:1517-1528.

Moreno, L. and Rasmuson, A. 1986. Contaminant transport through a fractured porous rock: Impact of the inlet boundary condition on the concentration profile in the rock matrix. *Water Resources Res.* 22:1728-1730.

Neretnieks, I., Eriksen, T., and Tahtinen, P. 1982. Tracer movement in a single fissure in granitic rock: Some experimental results and their interpretation. *Water Resources Res.* 18:849-858.

Neretnieks, I. 1983. A note on fracture flow dispersion mechanisms in the ground. *Water Resources Res.* 19:364-370.

Nijssen, B. M., Steenhuis, T. S., Kluitenberg, G. J., Stagnitti, F., and Parlange, J.-Y. 1991. Moving water and solutes through the soil: Testing of a preferential flow model, in *Preferential Flow.* Eds. T.J. Gish and A. Shirmohammadi. Proc. American Society of Agricultural Engineers National Symposium, December 16-17, 1991, Chicago, Illinois, 223-232.

Nkedi-Kizza, P., Biggar, J. W., van Genuchten, M. Th., Wierenga, P. J., Selim, H. M., Davidson, J. M., and Nielsen, D. R. 1983. Modeling tritium and chloride transport through an aggregated Oxisol. *Water Resources Res.* 19:691-700.

Omoti, U. and Wild, A. 1979. Use of fluorescent dyes to mark the pathways of solute movement through soils under leaching conditions. 2. Field experiments. *Soil Sci.* 128:98-104.

Parker, J. C. and van Genuchten, M. T. 1984. Determining transport parameters from laboratory and field tracer experiments. *Bulletin 84-3*. Virginia Agricultural Experimental Station, Blacksburg, Virginia.

Parker, J. C. and Valocchi, A. J. 1986. Constraints on the validity of equilibrium and first order kinetic transport models in structured soils. *Water Resources Res.* 22:399-407.

Parlange, J.-Y. and Hill, D. E. 1976. Theoretical analysis of wetting front instability in soils. *Soil Sci.* 122:236-239.

Parlange, J.-Y., Steenhuis, T. S., Glass, R. J., Pickering, N. B., Waltman, W. J., Bailey, N. O., Andreini, M. S., and Throop, J. A. 1988. The flow of pesticides through preferential paths in soils. *New York's Food and Life Science Quarterly.* 18:20-23.

Parlange, J.-Y., Steenhuis, T. S., Stagnitti, F., Simmelink, E., and Nijssen, B. 1996. Recent advances in modelling vadose zone transport, in *Subsurface Water Hydrology.* Eds. V.P. Singh and B. Kumar. Kluwer Academic Publishers, Netherlands, 127-151.

Parlange, J.-Y., Steenhuis, T. S., Haverkamp, R., Barry, D. A., Culligan, P. J., Hogarth, W. L., Parlange, M. B., and Ross, P. 1998. Soil properties and water movement, in *Vadose Zone Hydrology — Cutting Across Disciplines.* Eds. M.B. Parlange and J.W. Hopmans. Oxford University Press, New York. In press.

Passioura, J. B. 1971. Hydrodynamic dispersion in aggregate media. 1. Theory. *Soil Sci.* 111:339-334.

Philip, J. R. 1975. Stability analysis of infiltration. *Soil Sci. Soc. Am. Proc.* 39:1042-1049.

Piquemal, J. 1993. On the modelling conditions of mass transfer in porous media presenting capacitance effects by a dispersion-convection equation for the mobile fluid and a diffusion equation for the stagnant fluid. *Transport in Porous Media.* 10:271-283.

Raats, P. A. C. 1973. Unstable wetting fronts in uniform and nonuniform soils. *Soil Sci. Soc. Am. Proc.* 37:681-685.

Rao, P. S. C., Jessup, R. E., Ahuja, L. R., and Davidson, J. 1976. Evaluation of a capillary bundle model for describing solute dispersion in aggregate soils. *Soil Sci. Soc. Am. J.* 40:815-820.

Rao, P. S. C., Rolston, D. E., Jessup, R. E., and Davidson, J. M. 1980. Solute transport in aggregated porous media: Theoretical and experimental evaluation. *Soil Sci. Soc. Am. J.* 44:1139-1146.

Rice, R. C., Jaynes, D. B., and Bowman, R. S. 1991. Preferential flow of solutes and herbicide under irrigated fields. *Trans. Am. Soc. Agr. Eng.* 34:914-918.

Ritsema, C. J. and Dekker, L. W. 1993a. Preferential flow mechanism in a water repellent sandy soil. *Water Resources Res.* 29:2183-2193.

Ritsema, C. J. and Dekker, L. W. 1993b. Soil moisture and dry bulk density patterns in bare dune sands. *J. Hydrology.* 154:107-131.

Ritsema, C. J. and Dekker, L. W. 1995. Distribution flow: A general process in water repellent soils. *Water Resources Res.* 31:1187-1200.

Ritsema, C. J., Steenhuis, T. S., Parlange, J.-Y., and Dekker, L. W. 1996. Predicted and observed finger diameters in field soils, in *Fingered Flow in Unsaturated Soil: From Nature to Model.* Eds. T.S. Steenhuis, C.J. Ritsema, and L.W. Dekker. Special Issue of *Geoderma.* 70:185-196.

Roth, K., Jury, W. A., Flugler, H., and Attinger, W. 1991. Transport of chloride through an unsaturated field soil. *Water Resources Res.* 27:2533-2541.

Rubin, J. 1983. Transport of reacting solutes in porous media: Relation between mathematical nature of problems formulation and chemical nature of reaction. *Water Resources Res.* 19:1231-1252.

Selker, J. S., Leclerq, P., Parlange, J.-Y., and Steenhuis, T. S. 1992a. Fingered flow in two dimensions. Part 1. Measurement of matric potential. *Water Resources Res.* 28:2513-2521.

Selker, J. S., Parlange, J.-Y., and Steenhuis, T. S. 1992b. Fingered flow in two dimensions. Part 2. Predicting finger moisture profile. *Water Resources Res.* 28:2523-2528.

Selker, J. S., Steenhuis, T. S., and Parlange, J.-Y. 1993. Wetting front instability in homogeneous sandy soils under continuous infiltration. *Soil Sci. Soc. Am. J.* 56:1346-1350.

Seyfried, M. S. and Rao, P. S. C. 1987. Solute transport in undisturbed columns of an aggregated tropical soil: Preferential flow effects. *Soil Sci. Soc. Am. J.* 51:1434-1443.

Skopp, J., Gardner, W. R., and Taylor, E. J. 1981. Solute movement in structured soils: Two-regions model with small interaction. *Soil Sci. Soc. Am. J.* 45:837-842.

Skopp, J. and Gardner, W. R. 1992. Miscible displacement: An interacting flow region model. *Soil Sci. Soc. Am. J.* 56:1680-1686.

Smettem, K. R. J. and Collis-George, N. 1985. Statistical characterization of soil biopores using a soil peel method. *Geoderma.* 36:27-36.

Sposito, G., Jury, W. A., and Gupta, V. K. 1986. Fundamental problems in the stochastic convective-dispersion model of solute transport in aquifers and field soils. *Water Resources Res.* 22:77-88.

Stagnitti, F., Steenhuis, T. S., Parlange, J.-Y., Nijssen, B. M., and Parlange, M. B. 1991. Preferential solute and moisture transport in hillslopes. Institution of Engineers, Australia. 22:919-924.

Stagnitti, F., Parlange, J.-Y., Steenhuis, T. S., Boll, J., Pivetz, B., and Barry, D. A. 1995. Transport of moisture and solutes in the unsaturated zone by preferential flow, in *Environmental Hydrology.* Ed. V.P. Singh. Kluwer Academic Pub., Netherlands, 193-224.

Starr, J. L. and Parlange, J.-Y. 1977. Plate-induced tailing in miscible displacement experiments. *Soil Sci.* 124:56-60.

Starr, J. L., DeRoo, H. C., Frink, C. R., and Parlange, J.-Y. 1978. Leaching characteristics of a layered field soil. *Soil Sci. Soc. Am. J.* 42:376-391.

Steenhuis, T. S., Parlange, J.-Y., and Andreini, M. S. 1990. A numerical model for preferential solute movement in structured soils. *Geoderma.* 46:193-208.

Steenhuis, T. S., Boll, J., Shalit, G., Selker, J. S., and Merwin, I. A. 1994. A simple equation for predicting preferential flow solute concentrations. *J. Env. Quality.* 23:1058-1064.

Steenhuis, T. S., Vandenheuvel, K., Weiler, K. W., Boll, J., Daliparthy, J., Herbert, S., and Kung, K.-J. S. 1998. Mapping and interpreting soil textural layers to assess agri-chemical movement at several scales along the eastern seaboard (USA). *Nutrient Cycling in Agroecosystems.* 50:91-98.

Tang, D. H., Frind, E. O., and Sudicky, E. A. 1981. Contaminant transport in fractured porous media: Analytical solution for a single fracture. *Water Resources Res.* 17:555-564.

Valocchi, A. J. 1985. Validity of the local equilibrium assumption for modeling sorbing solute transport through homogeneous soils. *Water Resources Res.* 21:808-820.

Van Duin, C. J. and van der Zee, S. E. A. T. M. 1986. Solute transport parallel to an interface separating two different porous materials. *Water Resources Res.* 22:1779-1789.

van Genuchten, M. Th. and Wierenga, P. J. 1977. Mass transfer studies in sorbing porous media: Experimental evaluation with tritium. *Soil Sci. Soc. Am. J.* 41:278-285.

van Genuchten, M. Th. and Alves, W. J. 1982. Analytical solutions of the one-dimensional convection-dispersion solute transport equation. *Tech. Bull.* 1661. U.S. Dept. Agr., 1-151.

van Genuchten, M. Th. and Dalton, F. N. 1986. Models for simulating salty movement in aggregated field soils. *Geoderma.* 38:165-183.

Wallach, R. and Steenhuis, T. S. 1998. A model for non-reactive solute transport in structured soils with continuous preferential flow paths. *Soil Sci. Soc. Am. J.* In press.

Wallach, R. 1998. A small perturbations solution for nonequilibrium chemical transport through soils with relatively high desorption rate. *Water Resources. Res.* 34:149-154.

Wang, D., McSweeney, K., Norman, J. M., and Lowery, B. 1991. Preferential flow in soils with ant burrows, in *Preferential Flow.* Eds. T.J. Gish and A. Shirmohammadi. Proc. American Society of Agricultural Engineers National Symposium, December 16-17, 1991, Chicago, Illinois. 183-191.

Wang, D., Norman, J. M., Lowery, B., and McSweeney, K. 1994. Nondestructive determination of hydro-geometrical characteristics of soil macropores. *Soil Sci. Soc. Am. J.* 58:294-303.

Warren, J. E. and Root, P. J. 1963. The behaviour of naturally fractured reservoirs. *Soc. Petrol. Eng. J.* 3:245-255.

Weiler, K. W., Steenhuis, T. S., Vandenheuvel, K., Daliparthy, J., and Herbert, S. 1997. Capillary barrier effects on moisture patterns during infiltration. *Geoderma.* Submitted.

White, R. E., Thomas, G. W., and Smith, M. S. 1984. Modelling water flow through undisturbed soil cores using a transfer function model derived from ^{3}HOH and Cl transport. *J. Soil Sci.* 35:159-168.

White, R. E., Dyson, J. S., Gerstl, Z., and Yaron, B. 1986. Leaching of herbicides in the undisturbed cores of a structured clay soil. *Soil Sci. Soc. Am. J.* 50:277-283.

SYMBOLS

α^{-1}: average residence time in distribution layer
a, b: Laplace transform of A, B
A, B: terms in series for mobile and immobile concentration
c: solute concentration
d: finger diameter
D: dispersion
ϵ: small number
J: mass flux
κ: π or 4.8 in 2 and 3 dimensions, respectively
k: mobile–immobile transfer coefficients
K: soil water conductivity
l: depth of distribution layer
L: column length
M: applied mass per unit area
N: number of pore groups
θ: water content
q: solute flux
Q: imposed surface flux
s: Laplace variable
τ: mixing time
t: time
u: step-function
v: pore water velocity
x: direction perpendicular to flow
z: direction of flow

SUBSCRIPTS

e: effective property
F: finger tip
I: initial
m: mobile
im: immobile
n: order of expansion
p: pore group
w: water

8

Well Hydraulics and Aquifer Tests

J. Boonstra
International Institute for Land Reclamation and Improvement The Netherlands

8.1   Introduction

Estimating the physical properties of water-bearing layers is an essential part of groundwater studies. One of the most effective ways of determining these properties is to conduct and analyze aquifer tests. The principle of an aquifer test is simple: water is pumped from a well tapping the *aquifer*, and the discharge of the well and the changes in water levels in the well and in *piezometers* at known distances from the well are measured. The change in water level induced by the pumping is known as the *drawdown*. In the literature, aquifer tests based on the analysis of drawdowns during pumping are commonly referred to as *pumping tests*.

The physical properties can also be found from a recovery test. In such a test, a well that has been discharging for some time is shut down, and thereafter the recovery of the aquifer's hydraulic head is measured in the wells; the change in water level during this recovery period is known as the *residual drawdown*.

Owing to the high costs of aquifer tests, the number that can be performed in most groundwater studies has to be restricted. Nevertheless, one can perform an aquifer test without using piezometers, thereby cutting costs, although one must then accept a certain, sometimes appreciable, error. To distin-

guish such tests from normal aquifer tests, they are called single-well tests. In these tests, measurements are only taken inside the pumped well.

Analyzing and evaluating aquifer test data is as much an art as a science. It is a science because it is based on theoretical models that the geologist or engineer must understand and on thorough investigations that he must conduct into the geological formations in the area of the test. It is an art because different types of aquifers can exhibit similar drawdown behaviors, which demand interpretational skills on the part of the geologist or engineer.

The well-flow equations presented in Section 8.2 are from well hydraulics. Any lengthy derivations of the equations have been omitted because these can be found in the original publications listed in the References. The equations are presented in their final form, emphasizing the assumptions and conditions that underlie them. The essentials of how to conduct an aquifer test are summarized in Section 8.3. The remaining sections deal with the analysis and evaluation of aquifer test data from a variety of aquifer types and analysis methods, outlining the procedures that are to be followed for their successful application. All these procedures are illustrated with field data.

8.1.1   Aquifer Types

In the analysis and evaluation of aquifer test data, three main types of aquifer are distinguished: the confined, unconfined, and leaky aquifer (Figure 8.1).

A confined aquifer is a completely saturated aquifer whose upper and lower boundaries are *aquicludes*. In confined aquifers the pressure of the water is usually higher than that of the atmosphere, and the water level in wells tapping such aquifers stands above the top of the aquifer (see Figure 8.1A). The piezometric surface is the imaginary surface to which the water will rise in wells penetrating the aquifer. When the water level in wells tapping such aquifers stands above the ground surface, they are called free-flowing wells or artesian wells.

An unconfined aquifer is a partly saturated aquifer bounded below by an aquiclude and above by the free water table or phreatic surface (see Figure 8.1B). At the free water table, the pressure of the groundwater equals that of the atmosphere. The water level in a well penetrating an unconfined aquifer does not, in general, rise above the water table, except when there is vertical flow.

A leaky aquifer, also known as a semiconfined aquifer, is a completely saturated aquifer that is bounded below by an aquiclude and above by an *aquitard*. The overlying aquitard may be partly saturated when it extends to the land surface (see Figure 8.1.C) or is fully saturated when it is overlain by an unconfined aquifer that is bounded above by the water table (see Figure 8.1.D). The piezometric level in a well tapping a leaky aquifer may coincide with the water table if there is a hydrologic equilibrium between the aquifer's recharge and discharge; it may rise or fall below the water table in areas with upward or downward flow, in other words, in discharge or recharge areas.

A multilayered aquifer is a special case of the leaky aquifer. The aquicludes in Figures 8.1C and 8.1D are then aquitards, with other aquifers occurring at greater depths. In deep sedimentary basins such an interbedded system of permeable and less permeable layers is very common.

8.1.2   Physical Properties

This section summarizes the physical properties and derived parameters of aquifers and aquitards which appear in the various well-flow equations.

The hydraulic conductivity is the constant of proportionality in Darcy's law and is defined as the volume of water that will move through a porous medium in unit time under a unit hydraulic gradient through a unit area measured at right angles to the direction of flow. The hydraulic conductivity of an aquifer is denoted by the symbol K. Common orders of magnitude of the hydraulic conductivities for different aquifer materials range from 10^{-8} m/day for clay deposits to 10^3 m/day for coarse sands.

For confined aquifers, the saturated thickness is equal to the physical thickness of the aquifer between the aquicludes above and below it (see Figure 8.1A). The same is true for the confined parts of a leaky
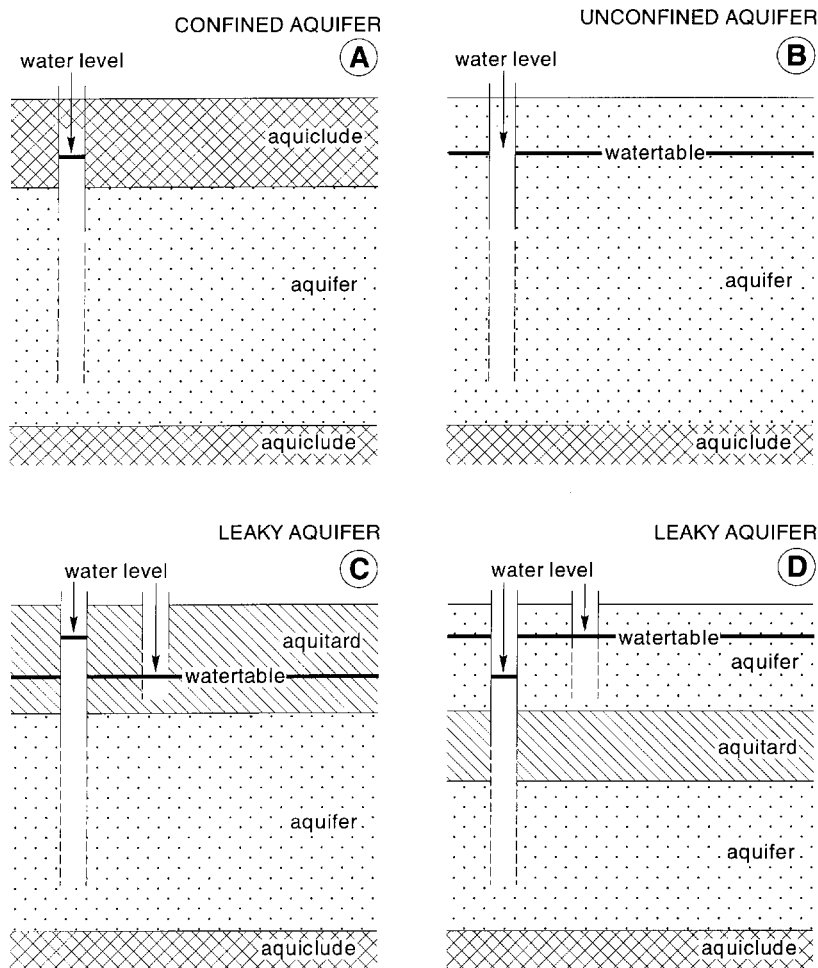
FIGURE 8.1 Different types of aquifers.

aquifer bounded by an aquitard and an aquiclude (see Figures 8.1C and 8.1D). In both cases, the saturated thickness is a constant.

For unconfined aquifers, the saturated thickness is equal to the difference between the free water table and the aquiclude (see Figure 8.1B). Because the water table changes its position with time, the saturated thickness of an unconfined aquifer is not constant, but variable. Whether constant or variable, the saturated thickness of an aquifer is denoted by the symbol D. Its order of magnitude can range from several meters to hundreds or even thousands of meters.

For aquitards in leaky aquifers, the saturated thickness can be variable or constant. In Figure 8.1C, the aquitard is partly saturated and has a free water table. Its saturated thickness depends upon the position of the water table. In Figure 8.1D, the aquitard is bounded by two aquifers and is fully saturated. Its saturated thickness is physically determined and thus constant. The saturated thickness of an aquitard is denoted by the symbol D'. It may range from a few meters to hundreds of meters.

The transmissivity is the product of the average hydraulic conductivity and the saturated thickness of the aquifer. Consequently, the transmissivity is the rate of flow under a hydraulic gradient equal to unity through a cross section of unit width over the whole saturated thickness of the water-bearing layer. The transmissivity of an aquifer is denoted by the symbol T. Its order of magnitude can be derived from those of K and D, and may range from hundreds to tens of thousands of m²/day.

The specific storage of a saturated confined aquifer is the volume of water that a unit volume of aquifer releases from storage under a unit decline in head. This release of water under conditions of decreasing hydraulic head stems from the compaction of the aquifer due to increasing effective stress and from the expansion of the water due to decreasing water pressure. The specific storage depends on the elasticity of both the aquifer material and the water. For a certain location it can be regarded as a constant. It is denoted by the symbol S_s; its order of magnitude is 10^{-6} to 10^{-4} day^{-1}.

The storativity of a saturated confined aquifer of thickness D is defined as the volume of water released from storage per unit surface area of the aquifer per unit decline in the component of hydraulic head normal to that surface. In a vertical column of unit area extending through the confined aquifer, the storativity equals the volume of water released from the aquifer when the piezometric surface drops over a unit distance. Storativity is thus defined as $S = S_s D$. The storativity of a saturated aquifer is a function of its thickness. Storativity is a dimensionless quantity because it involves a volume of water per volume of aquifer. Its values in confined aquifers range from 5×10^{-5} to 5×10^{-3}.

The specific yield is the volume of water that an unconfined aquifer releases from storage per unit surface area of aquifer per unit decline of the water table. In unconfined aquifers, the effects of the elasticity of the aquifer material and the water are negligible, except for a short time after the beginning of pumping. Values of the specific yield of unconfined aquifers are much higher than the storativities of confined aquifers. The orders of magnitude of the specific yield denoted by the symbol S_y range from 0.01 for clay deposits to 0.47 for coarse sands. Specific yield is also referred to as effective porosity, unconfined storativity, or drainable pore space. Small interstices do not contribute to the effective porosity, because the retention forces in them are greater than the weight of water. Hence, no groundwater will be released from these small interstices by gravity drainage.

The hydraulic resistance characterizes the resistance of an aquitard to vertical flow, either upward or downward. It is the ratio of the saturated thickness of the aquitard D' and its hydraulic conductivity for vertical flow K' and is thus defined as $c = D'/K'$. Its order of magnitude may range from a few tens to thousands of days. Aquitards having c-values of 2500 days or more are regarded to act as aquicludes, although theoretically an aquiclude has an infinitely high c-value.

The leakage factor describes the spatial distribution of leakage through one or two aquitards into a leaky aquifer or vice versa. It is the square root of the product of the aquifer's transmissivity and the aquitard's hydraulic resistance and is thus defined as $L = \sqrt{(Tc)}$; it is expressed in meters. Large values of L originate from a high transmissivity of the aquifer and/or a high hydraulic resistance of the aquitard. In both cases the influence of leakage will be small and the area over which leakage takes place, large.

8.2   Well-Flow Equations

The well-flow equations presented in this section were developed under the following common assumptions and conditions: (1) the aquifer has a seemingly infinite areal extent; (2) the aquifer is *homogeneous, isotropic*, and of uniform thickness over the area influenced by the test; (3) prior to pumping, the hydraulic head is horizontal (or nearly so) over the area that will be influenced by the test; (4) the pumped well penetrates the entire thickness of the aquifer and thus receives water by horizontal flow; (5) the aquifer is pumped at a constant-discharge rate; (6) the water removed from storage is discharged instantaneously with decline of head; and (7) the diameter of the pumped well is small, i.e., the storage inside the well can be neglected.

These simplifying assumptions and limiting conditions cannot always be fulfilled in the field. In these cases, special well-flow equations, for instance for bounded, sloping, or anisotropic aquifers or for large-diameter wells, can be used; these can be found in Kruseman and de Ridder (1990).

8.2.1   Confined Aquifers

Theis (1935) was the first to develop an equation for unsteady-state flow which introduced the time factor and the storativity. He noted that when a fully penetrating well pumps an extensive confined

aquifer at a constant rate, the influence of the discharge extends outward with time. The rate of decline of head, multiplied by the storativity and summed over the area of influence, equals the discharge. The Theis equation, which was derived from the analogy between the flow of groundwater and the conduction of heat, is written as

$$s(r,t) = \frac{Q}{4\pi T} \int_u^\infty \frac{e^{-y}}{y} dy = \frac{Q}{4\pi T} W(u)$$ (1)

with

$$u = \frac{r^2 S}{4Tt}$$ (2)

where $s(r,t)$ is the drawdown in m measured in a well, Q is the constant well discharge in m³/day, T is the transmissivity of the aquifer in m²/day, $W(u)$ is the dimensionless Theis well function, r is the distance in m from the pumped well, S is the dimensionless storativity of the aquifer, and t is the time in days since pumping started. In Figure 8.2, the Theis well function $W(u)$ is plotted versus $1/u$ on semi-log paper. This figure shows that, for large values of $1/u$, the Theis well function exhibits a straight-line segment. The Jacob method is based on this phenomenon. Cooper and Jacob (1946) showed that, for the straight-line segment, Equation (1) can be approximated by

$$s(r,t) = \frac{2.3Q}{4\pi T} \log \frac{2.25Tt}{r^2 S}$$ (3)

with

an error less than	1%	2%	5%	10%
for 1/u larger than	30	20	10	7

In most handbooks on this subject, the condition to use Equation (3) is taken as $1/u > 100$. From personal experience, this limiting condition can often be relaxed to $1/u > 10$.
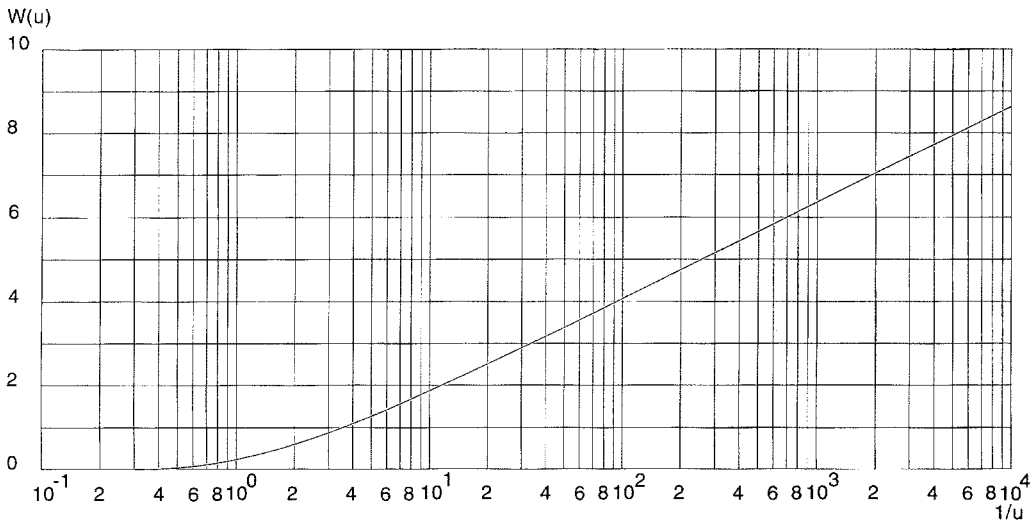
FIGURE 8.2 Theis well function W(u) versus 1/u for fully penetrating wells in confined aquifers.

When a confined aquifer is pumped, the cone of depression will continuously deepen and expand. Even at late pumping times, the water levels in the piezometers will never stabilize to a real steady state. Although the water levels continue to drop, the cone of depression will eventually deepen uniformly over the area influenced by the pumping. At that stage, the hydraulic gradient has become constant; this phenomenon is called pseudo-steady state. For this situation, Thiem (1906), using two or more piezometers, developed an equation, the so-called Thiem–Dupuit equation, which can be written as

$$s(r_1, t) - s(r_2, t) = \frac{2.3Q}{2\pi T} \log \frac{r_2}{r_1} \qquad (4)$$

Equation (4) can also be derived by applying Equation (3) to two piezometers at distances r_1 and r_2 at large times.

8.2.2 Unconfined Aquifers

Figure 8.3 shows a pumped unconfined aquifer underlain by an aquiclude. There are the following basic differences between unconfined and confined aquifers when they are pumped: (1) an unconfined aquifer is partly dewatered during pumping, resulting in a decreasing saturated aquifer thickness, whereas a confined aquifer remains fully saturated, (2) the water produced by a well in an unconfined aquifer comes from the physical dewatering of the aquifer, whereas in a confined aquifer it comes from the expansion of the water in the aquifer due to a reduction of the water pressure, and from the compaction of the aquifer due to increased effective stresses, and (3) the flow toward a well in an unconfined aquifer has clear vertical components near the pumped well, whereas there are no such vertical flow components in a confined aquifer, provided, of course, that the well is fully penetrating.
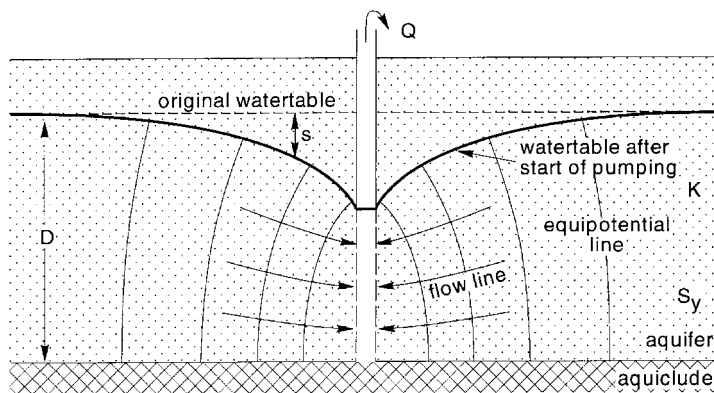
FIGURE 8.3 Schematic cross section of a pumped unconfined aquifer.

In an unconfined aquifer, the water levels in piezometers near the well often tend to decline at a slower rate than that described by the Theis equation. Time-drawdown curves on semi-log paper therefore usually show a typical S shape: a relatively steep early-time segment, a flat intermediate segment, and a relatively steep segment again at later times, as is depicted in Figure 8.4. Nowadays, the widely used explanation of this S-shaped time-drawdown curve is based on the concept of delayed yield. It is caused by a time lag between the early elastic response of the aquifer and the subsequent downward movement of the water table due to gravity drainage.

During the early stage of an aquifer test — a stage that may last for only a few minutes — the discharge of the pumped well is derived uniquely from the elastic storage within the aquifer. Hence, the reaction
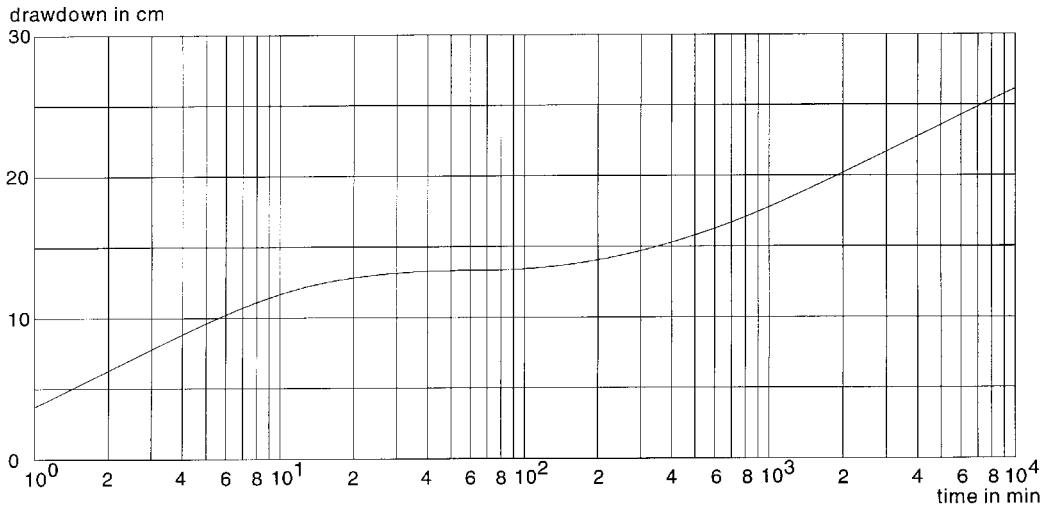
FIGURE 8.4 Time-drawdown plot in an unconfined aquifer showing delayed yield.

of the unconfined aquifer immediately after the start of pumping is similar to the reaction of a confined aquifer as described by the flow equation of Theis.

Only after some time does the water table start to fall and the effect of the delayed yield becomes apparent. The influence of the delayed yield is comparable to that of leakage: the average drawdown slows down with time and no longer conforms to the Theis curve. After a few minutes to a few hours of pumping, the time-drawdown curve approaches a horizontal position.

The late-time segment of the time-drawdown curve may start from several minutes to several days after the start of pumping. The declining water table can now keep pace with the increase in the average drawdown. The flow in the aquifer is essentially horizontal again and as in the early pumping time the time-drawdown curve approaches the Theis curve.

Jacob (1950) showed that if the drawdowns in an unconfined aquifer are small compared to the initial saturated thickness of the aquifer, the condition of horizontal flow toward the well is approximately satisfied, so that Equations (1) to (3) can also be applied to determine the physical properties. The only changes required are that the storativity S be replaced by the specific yield S_y of the unconfined aquifer, and that the transmissivity T be defined as the transmissivity of the initial saturated thickness of the aquifer.

When the drawdowns in an unconfined aquifer are large compared with the aquifer's original saturated thickness, the observed drawdowns need to be corrected before this equation can be used. Jacob (1944) proposed the following correction

$$s_c(r,t) = s(r,t) - \frac{s^2(r,t)}{2D} \tag{5}$$

where $s_c(r,t)$ is the corrected drawdown in m, $s(r,t)$ is the observed drawdown in m, and D is the saturated aquifer thickness in m prior to pumping. This correction is only needed when the maximum drawdown at the end of the test is larger than 5% of the original saturated aquifer thickness.

As with confined aquifers, the cone of depression will continuously deepen and expand in unconfined aquifers. For pseudo-steady-state conditions, Equation (4) can also be used to describe the drawdown behavior in two piezometers in an unconfined aquifer, provided that the observed drawdowns are corrected according to Equation (5).

8.2.3 Leaky Aquifers

When a leaky aquifer is pumped (Figure 8.5), the piezometric level of the aquifer in the well is lowered. This lowering spreads radially outward as pumping continues, creating a difference in hydraulic head between the aquifer and the aquitard. Consequently, the groundwater in the aquitard will start moving vertically downward to join the water in the aquifer. The aquifer is thus partially recharged by downward percolation from the aquitard. As pumping continues, the percentage of the total discharge derived from this percolation increases. After a certain period of pumping, equilibrium will be established between the discharge rate of the pump and the recharge rate by vertical flow through the aquitard. This steady state will be maintained as long as the water table in the aquitard is kept constant.
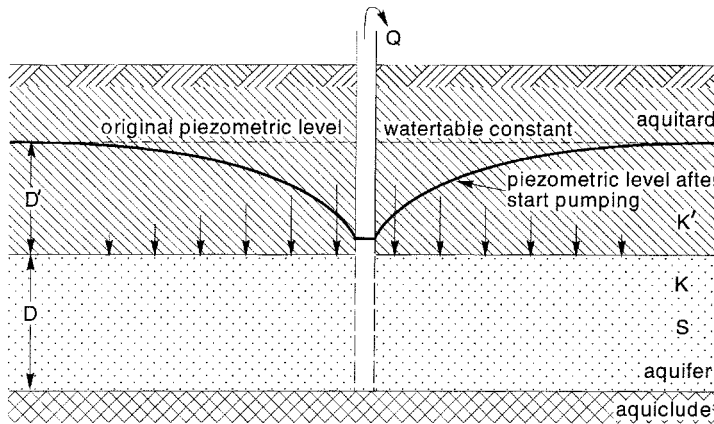
FIGURE 8.5 Schematic cross section of a pumped leaky aquifer.

According to Hantush and Jacob (1955), the drawdown due to pumping a leaky aquifer can be described by the following equation

$$s(r,t) = \frac{Q}{4\pi T}\int_u^\infty \frac{1}{y}\exp\left(-y - \frac{r^2}{4L^2 y}\right)dy = \frac{Q}{4\pi T}W(u, r/L) \tag{6}$$

where $W(u,r/L)$ is the dimensionless Hantush well function, $L\,(=\sqrt{(Tc)})$ is the leakage factor or characteristic length in m, $c\,(= D'/K')$ is the hydraulic resistance of the aquitard in days, D' is the saturated thickness of the aquitard in m, K' is the vertical hydraulic conductivity of the aquitard in m/day, and the other symbols as defined earlier.

Equation (6) has the same form as the Theis equation (Equation [1]), but there are two parameters in the integral: u and r/L. Equation (6) approaches the Theis equation for large values of L, when the exponential term $r^2/(4L^2 y)$ approaches zero. In Figure 8.6, the Hantush well function $W(u,r/L)$ is plotted versus $1/u$ on semi-log paper for an arbitrary value of r/L. This figure shows that the Hantush well function exhibits an S shape and, for large values of $1/u$, a horizontal straight-line segment indicating steady-state flow.

For the steady-state drawdown in a leaky aquifer, De Glee (1930, 1951) derived the following equation

$$s_m(r) = \frac{Q}{2\pi T}K_o\left(\frac{r}{L}\right) \tag{7}$$
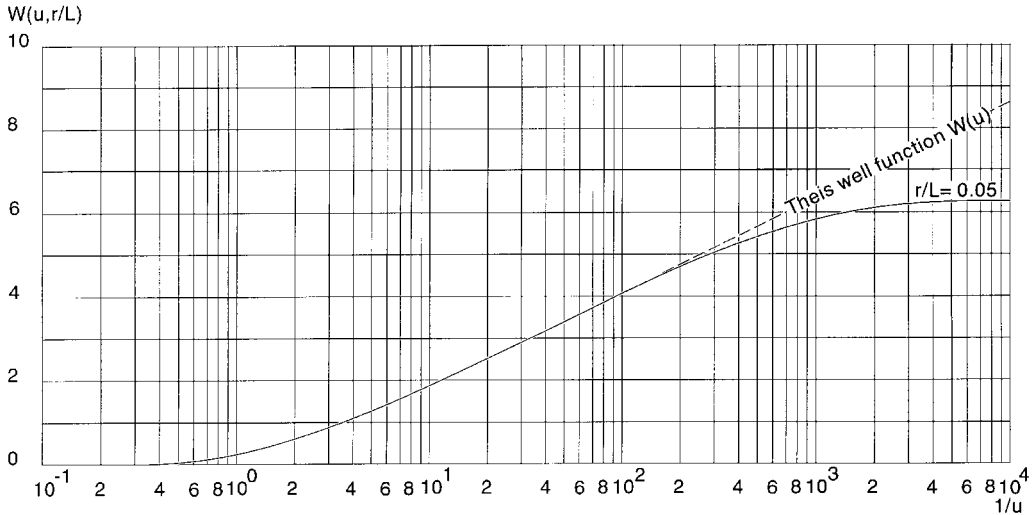
FIGURE 8.6 Hantush well function W(u,r/L) versus 1/u for fully penetrating wells in leaky aquifers.

where $s_m(r)$ is the steady-state, stabilized drawdown in m and $K_o(r/L)$ is the dimensionless modified Bessel function of the second kind and of zero order (Hankel function). Hantush (1956, 1964) noted that if r/L is small ($r/L < 0.05$) and $L > 3D$, Equation (7) can, for all practical purposes, be approximated by

$$s_m(r) = \frac{2.3Q}{2\pi T}\log\frac{1.12L}{r} \tag{8}$$

It is important to note that the flow system in a pumped leaky aquifer consists of a vertical component in the overlying aquitard and a horizontal component in the aquifer. In reality, the flow lines in the aquifer are not horizontal but curved (i.e., there are both vertical and horizontal flow components in the aquifer). The above equations can therefore only be used when the vertical-flow component in the aquifer is so small compared to the horizontal-flow component that it can be neglected. In practice, this condition is fulfilled when $L > 3D$.

8.2.4 Partial Penetration

When an aquifer is pumped by a partially penetrating well, the assumption that the well receives water from horizontal flow is no longer valid. Hence, the previous equations cannot be used to describe the flow to pumped wells. Due to a contraction of flow lines, partial penetration causes the flow velocity in the immediate vicinity of the well to be higher than it would be otherwise, leading to an extra loss of head. This effect is strongest at the well face, and decreases with increasing distance from the well. According to Hantush (1962) the drawdown due to pumping in a confined aquifer can be described by the following equation

$$s(r,t) = \frac{Q}{4\pi T}\left[W(u) + f_s\right] \tag{9}$$

and

$$f_s = \frac{2D}{\pi(b-d)} \sum_{n=1}^{\infty} \left(\frac{1}{n}\right) W\left(u, \frac{n\pi r}{D}\right) \times \left[\sin\left(\frac{n\pi b}{D}\right) - \sin\left(\frac{n\pi d}{D}\right)\right]\left(\cos\frac{n\pi z}{D}\right) \tag{10}$$

where $W(u, n\pi r/D)$ is the dimensionless Hantush well function, b and b' are the penetration depths in m of the pumped well and of the piezometer, d and d' are the non-screened parts in m of the pumped well and of the piezometer, and $z = (b' + d')/2$ (see Figure 8.7).
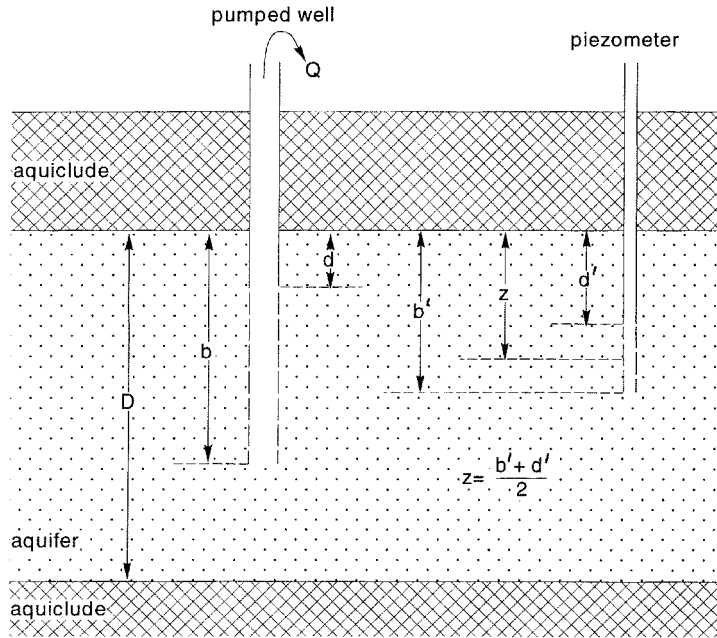
FIGURE 8.7 Schematic illustration of the parameters used for the analysis of partially penetrating wells.

In Figure 8.8, the expression $W(u) + f_s$ is plotted versus $1/u$ on semi-log paper. This figure shows that, for large values of $1/u$, this expression exhibits a straight-line segment. For this segment, Equation (10) reduces to

$$f_s = \frac{4D}{\pi(b-d)} \sum_{n=1}^{\infty} \left(\frac{1}{n}\right) K_o\left(\frac{n\pi r}{D}\right)\left[\sin\left(\frac{n\pi b}{D}\right) - \sin\left(\frac{n\pi d}{D}\right)\right]\left(\cos\frac{n\pi z}{D}\right) \tag{11}$$

For a particular well/piezometer configuration, Equation (11) yields constant values for f_s. The Hantush's modification of the Jacob equation was based on this phenomenon. Hantush (1962) showed that for the straight-line segment, Equation (9) can be approximated by

$$s(r,t) = \frac{2.3Q}{4\pi T}\log\frac{2.25Tt}{r^2 S} e^{f_s} \tag{12}$$

Equation (12) can be used to describe the drawdown behavior in a confined aquifer with partially penetrating well, provided that $t > D^2 S/2T$. Equation (12) can also be applied to unconfined aquifers. It is then assumed that due to partial penetration the drawdowns are small compared with the initial
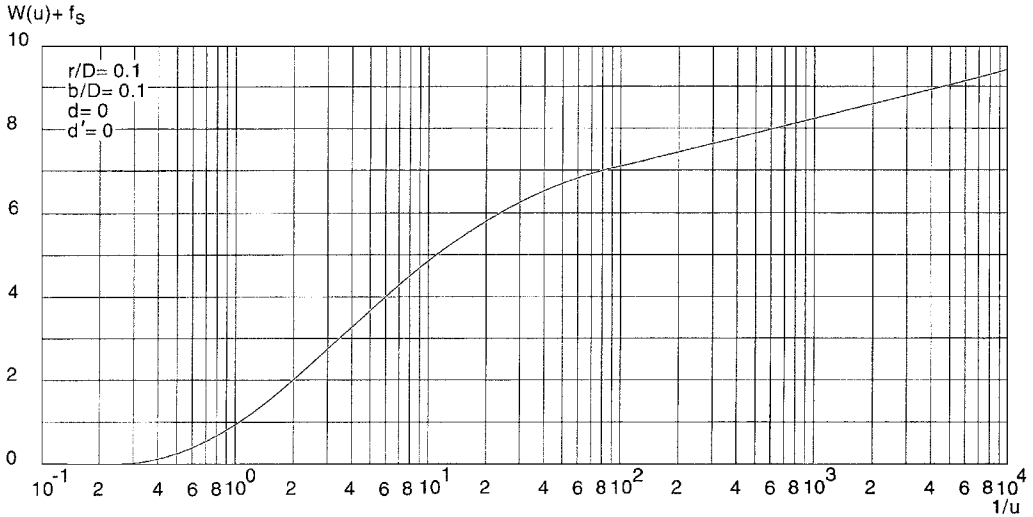
FIGURE 8.8 Hantush well function $W(u) + f_s$ versus $1/u$ for partially penetrating wells in confined aquifers.

saturated thickness of the aquifer. The only changes required are that the storativity S be replaced by the specific yield S_y of the unconfined aquifer, and that the transmissivity T be defined as the transmissivity of the initial saturated thickness of the aquifer.

For leaky aquifers, the drawdown behavior can be described by (Weeks, 1969)

$$s(r,t) = \frac{Q}{4\pi T}\Big[W\big(u, r/L\big) + f_s\Big] \tag{13}$$

The pumping time should also be long to apply Equation (13) ($t > D^2 S / 2T$).

8.2.5 Recovery Well-Flow Equations

The well-flow equations describing the drawdown behavior during the recovery period are based on the principle of superposition. Applying this principle, it is assumed that, after the pump has been shut down, the well continues to be pumped at the same discharge as before, and that an imaginary recharge, equal to the discharge, is injected into the well. The recharge and discharge thus cancel each other, resulting in an idle well as is required for the recovery period. For any of the well-flow equations presented in the previous sections, a corresponding recovery equation can be formulated.

On the basis of Equation (1), Theis (1935) developed his recovery method for confined aquifers. After a constant-rate pumping test, the residual drawdown during the recovery period is given by

$$s'(r, t/t') = \frac{Q}{4\pi T}\big\{W(u) - W(u')\big\} \tag{14}$$

where $s'(r,t)$ is the residual drawdown in m, $u' = r^2 S'/(4Tt')$, S' is the dimensionless storativity of the aquifer during recovery, and t' is the time in days since pumping stopped. In Figure 8.9, the Theis recovery well function $W(u) - W(u')$ is plotted versus u'/u on semi-log paper. This figure shows that, for small values of u'/u, the expression exhibits a straight-line segment. For this segment, Equation (14) can be approximated by
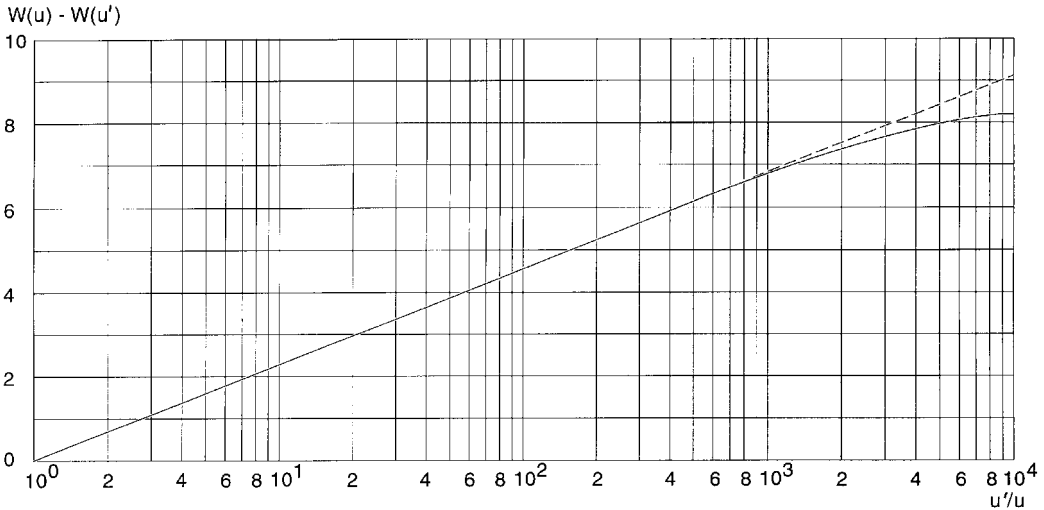
FIGURE 8.9 Theis recovery well function $W(u) - W(u')$ versus u'/u for fully penetrating wells in confined aquifers.

$$s'(r, t/t') = \frac{2.3Q}{4\pi T} \log \frac{S't}{St'} \tag{15}$$

Neuman (1975) showed that Equation (15) can also be used to describe the residual drawdown behavior in unconfined aquifers, because its delayed water table response to pumping is fully reversible (no hysteresis effects). The only changes required are that the storativity S be replaced by the specific yield S_y of the unconfined aquifer, and that the transmissivity T be defined as the transmissivity of the initial saturated thickness of the aquifer.

For leaky aquifers, the principle of superposition can also be applied. This results in the following equation

$$s(r, t) = \frac{Q}{4\pi T}\left[W\left(u, \frac{r}{L}\right) - W\left(u', \frac{r}{L}\right) \right] \tag{16}$$

Equation (16) can only be solved by numerical simulation. Hantush (1964) showed that, if pumping and recovery times are short, i.e., if $t + t' < (L^2 S)/20T$ or $t + t' < cS/20$, Equation (14) can also be used.

8.3 Performance of an Aquifer Test

The changes of water levels observed in both aquifer tests and single-well tests can be analyzed by using the time-drawdown and time-residual-drawdown relationships of the well-flow equations of Section 8.2. Figure 8.10A gives an example of the time-drawdown relationship for the pumped well or a piezometer during a pumping test followed by a recovery test. With aquifer tests, the above relationships can be used to make time-drawdown analyses (Section 8.4) for each piezometer separately. In addition, the distance-drawdown relationship (see Figure 8.10B) can be used to make a distance-drawdown analysis (Section 8.5), provided that drawdown data in at least two piezometers are available. The advantage of aquifer tests is thus that the various estimates of the aquifer's physical properties obtained from the time-drawdown analyses of each piezometer can be compared with an independent estimate of the physical properties obtained from a distance-drawdown analysis.
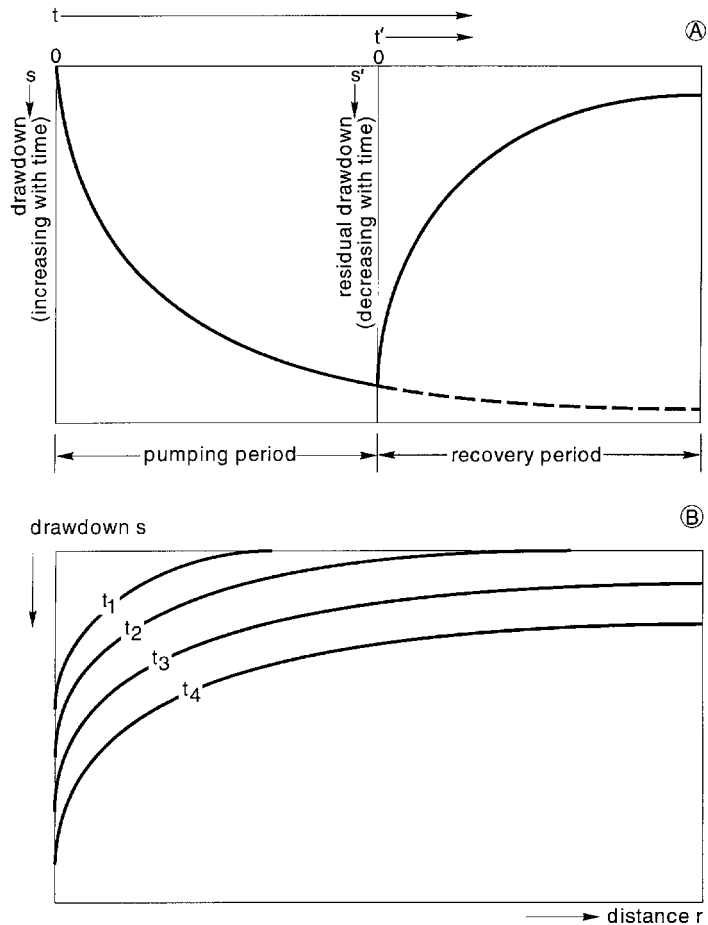
FIGURE 8.10 A) Time-drawdown relationship during a pumping test, followed by a recovery test; B) distance-drawdown relationship during a pumping test.

With single-well tests, only a single time-drawdown analysis based on the drawdowns of the pumped well can be made, i.e., no comparison of results from different time-drawdown analyses is possible, and no distance-drawdown analysis can be made. Consequently, the results of aquifer tests will be more accurate than the results of single-well tests. Moreover, aquifer-test results are representative of a larger volume of the aquifer than single-well test results.

The question of how many piezometers should be employed depends not only on the amount of information desired and the required degree of accuracy, but also on the funds available for the test. It is always best to have as many piezometers as conditions permit and to place them in various directions and distances from the pumped well.

Although no fixed rule can be given, placing piezometers at distances of between 10 and 100 m from the well will usually give reliable data. For thick or stratified confined aquifers, these distances must be greater, say between 100 to 250 m or more. An optimum result in determining the physical properties of the aquifer can be obtained by placing the screen of the piezometer at half the depth of the screen of the pumped well.

The design and construction of the pumped well and the piezometers will be discussed in Chapter 9; for the site selection, reference is made to Kruseman and de Ridder (1990).

8.3.1 Measurements

Ideally, an aquifer test should be performed under the natural conditions of a stable water table. This is, however, not always possible. Water tables rise and fall due to natural recharge and discharge of the groundwater reservoir (precipitation and evaporation), man-made recharge and discharge of the ground-water reservoir (irrigation losses and pumping from wells), changes in barometric pressure, and/or tidal movements in coastal aquifers. Such short-term variations of the water table have an adverse effect on the drawdown and recovery of the water table during testing. Hence, for some days prior to the actual test, the water levels in the well and the piezometers should be measured, say, twice a day. For each observation point, a time-versus-water-level curve, or hydrograph, should be drawn. From these, the trend and rate of water-level changes can be read. At the end of the test, i.e., after complete recovery, water-level readings should be continued at the observation points for one or two days. With these data, the hydrographs can be completed and the rate of water-level change during the test can be determined. This information can then be used to correct the drawdowns observed during the test itself.

The water-level measurements must be taken many times during the course of a test, and with as much accuracy as possible. Because water levels drop fast during the first hour or two of the test, readings should first be taken at brief intervals. These intervals should be gradually increased as pumping continues. Since in all the analysis procedures the time is plotted on a logarithmic scale, it is recommended to have the same number of readings in each log cycle of time. Table 8.1 shows an example of the sequence in time for taking water-level measurements, based on ten readings in each log cycle and resulting in approximately equidistant plotting positions. For piezometers far from the well and for those in aquitards above or below the aquifer, the brief time intervals in the first minutes of the pumping test can be disregarded.

TABLE 8.1 Typical Time Sequence for Taking Water-Level Measurements in an Aquifer Test

Time (s)	Time (min)	Time (min)	Time (h)	Time (h)
10	2.5	20	2.5	22
20	3.0	25	3.0	27
30	4.0	30	4.0	33
40	5.0	40	5.0	42
50	6.5	50	7.0	53
60	8.0	65	8.5	67
80	10.0	80	11.0	83
100	13.0	90	13.0	108
120	16.0	120	17.0	133

Note: Aim is to get equal number of observations in each log cycle of time

After the pump has been shut down, the water levels in the well and the piezometers will start to rise. In the first hour, they will rise rapidly, but as time goes on the rate of rise decreases. These rises can be measured in a recovery test. If the yield of the well was not constant throughout the pumping test, recovery-test data are more consistent than the drawdown data collected during pumping. Recovery-test data can thus be used as a check on the calculations that are based on the drawdown data. The schedule for recovery measurements is the same as that for the pumping period.

Water-level measurements can be taken in various ways, i.e., wetted-tape method, mechanical sounder, electric water-level indicator, floating-level indicator or recorder, pressure gauge, or pressure logger. For detailed information on these devices, reference is made to Kruseman and de Ridder (1990) and Driscoll (1986). Fairly accurate measurements of water levels can be made manually, but then the instant of each reading should be recorded with a chronometer. Experience has shown that it is possible to measure the depth to water within two millimeters. For piezometers close to the well, the wetted-tape method and

the mechanical sounder cannot be used because of the rapid water-level changes and the noise of the pump, respectively.

Although the pressure-gauge method is less accurate than the other methods (within 6 cm), it is the most practical method for measuring water levels in a pumped well. It should, however, not be used for measuring water levels in piezometers, because the drop in water levels during the initial stage of pumping will be small in these wells.

Among the arrangements to be made for a pumping test is the control of the discharge rate. To avoid complicated calculations later, the discharge rate should be kept constant throughout the test. The discharge should be kept constant by manipulating a valve in the discharge pipe. This gives more accurate control than changing the speed of the pump. During pumping tests, the discharge rate should be measured at least once every hour, and adjustments necessary to keep it constant should be made. The discharge rate can be measured with various devices, i.e., commercial water meter, flume, container, orifice weir, orifice bucket, or jet-stream method. For detailed information on these devices, reference is also made to Kruseman and de Ridder (1990) and Bos (1989). The water delivered by the well should be prevented from returning to the aquifer. This can be done by conveying the water through a large-diameter pipe over a convenient distance, at least 300 m, depending on the location of the piezometers, and then discharging it into a canal or natural channel. Preferably, the water should be discharged away from the line of piezometers. The pumped water can also be conveyed through a shallow ditch, but precautionary measures should be taken to seal the bottom of the ditch with clay or plastic sheets to prevent leakage.

8.3.2   Duration of a Pumping Test

The question of how long a pumping test should last is difficult to answer because the period of pumping depends on the type of aquifer and the degree of accuracy desired in establishing its physical properties. Economizing on the pumping period is not recommended because the costs of running the pump a few extra hours are low compared with the total costs of the test. Moreover, better and more reliable data are obtained if pumping continues until the cone of depression has stabilized and does not seem to be expanding further as pumping continues. At the beginning of the test, the cone develops quickly because the pumped water is initially derived from the aquifer storage immediately around the well. But, as pumping continues, the cone expands and deepens more slowly because, with each additional meter of horizontal expansion, a larger volume of stored water becomes available. This may often lead inexperienced observers to conclude that the cone has stabilized, or in other words that steady state has been reached. Inaccurate measurements of the drawdowns in the piezometers — drawdowns that are becoming additionally smaller and smaller as pumping continues — can also lead to this wrong conclusion. In reality, the depression cone will continue to expand until the recharge of the aquifer, if any, equals the discharge.

The unsteady-state flow, also known as non-equilibrium flow, is time dependent, i.e., the water level as observed in piezometers changes with time. During a pumping test, the unsteady-state flow condition occurs from the moment pumping starts until the steady state is reached. Theoretically, in an infinite, horizontal, completely confined aquifer of constant thickness which is pumped at a constant rate, there will always be an unsteady state, because the aquifer is not recharged by an outside source. In practice, well flow is considered to be in unsteady state as long as the changes in the water level in the piezometers are measurable, or as long as the hydraulic gradient changes in a measurable way.

The steady-state flow, also known as equilibrium flow, is independent of time, i.e., the water level, as observed in piezometers, does not change with time. It occurs, for instance, when there is equilibrium between the discharge of a pumped well and the recharge of the pumped aquifer by an outside source. Such outside sources may be recharge from surface water of nearby rivers, canals, or lakes, or recharge from groundwater of an unconfined aquifer with constant water table overlying an aquitard which covers a pumped leaky aquifer. Because real steady-state conditions seldom occur, it is said in practice that a
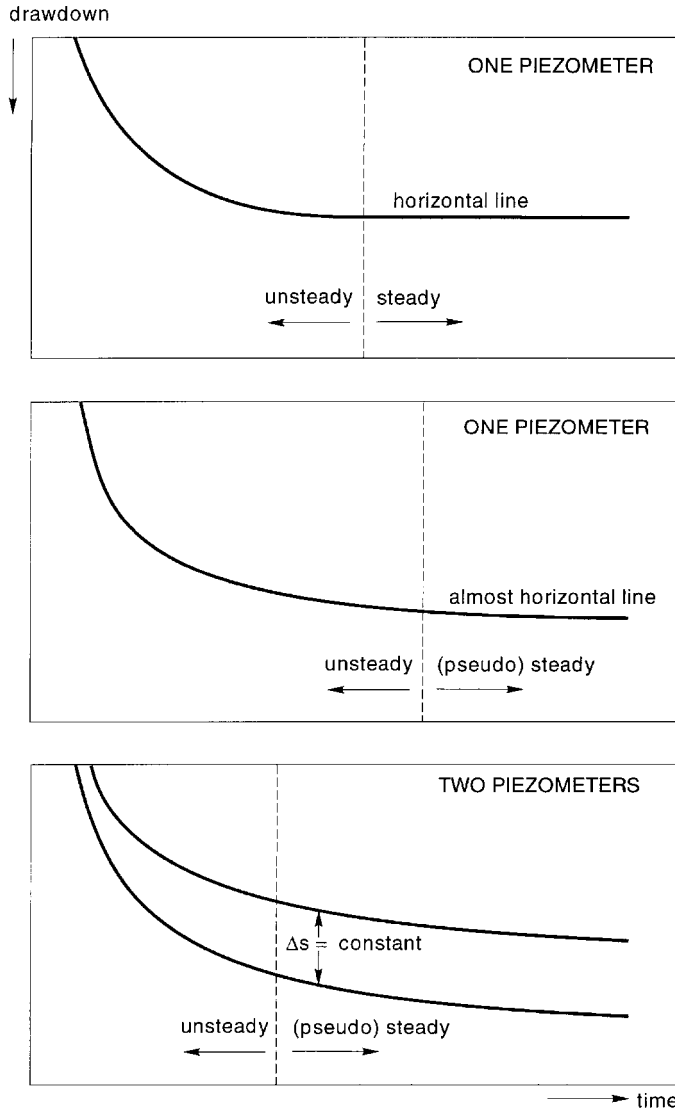
FIGURE 8.11 Time-drawdown plots showing the changes in drawdown during an aquifer test and their interpretations.

steady-state condition is reached when the changes in the water level as observed in piezometers are negligibly small, or when the hydraulic gradient has become constant.

To establish whether unsteady or steady-state conditions prevail, the changes in head during the pumping test should be plotted. Figure 8.11 shows possible shapes of time-drawdown plots and their interpretations.

In some wells, a steady state occurs a few hours after the start of pumping; in others, it does not occur until after a few days or weeks, whereas in yet other wells it never occurs, even though pumping continues for years. Kruseman and de Ridder (1990) suggest that, under average conditions, steady-state flow is generally reached in leaky aquifers after 15 to 20 hours of pumping, and in a confined aquifer, after 24 hours. In an unconfined aquifer, the cone of depression expands more slowly, so a longer period of pumping is required, say, three days.

Preliminary plotting of drawdown data during the test will often show what is happening and may indicate whether or not a test should be continued.

8.3.3   Processing the Data

All measurements of the water level, time, and discharge of the pump should preferably be noted on preprinted forms. After three hours of pumping, sufficient time will become available in the field to draw the time-drawdown curves of each observation point. These graphs will be helpful in checking whether the test is running well and in deciding on the time that the pump can be shut down when steady or pseudo-steady-state flow has been reached.

Because the well-flow equations require drawdown data instead of depth-to-water table data, the latter should therefore be converted to drawdown data. In other words, the initial depth to the water table must be subtracted from the depth-to-water table data during the test. It may be necessary to correct the observed drawdowns for external influences, i.e., influences independent of the test (trend correction). If, during the post-recovery period, the same constant water level is observed as during the pre-testing period, it can be assumed that no external forces influenced the hydraulic head in the aquifer during the test. If, however, the water level is subject to unidirectional or rhythmic changes due to tidal influences for example, the observed drawdowns will have to be corrected before being used in the analysis. Such phenomena are likely to occur during pumping tests of long duration.

When the evaluation of the test data has been completed, a report on the results should be written. A copy of the report should be kept on file for further reference and later studies. Samples of the different layers penetrated by the borings should be stored too, because they may be needed for other studies in a later phase of investigations. The basic field measurements of the test should be put on file as well. The conclusions drawn from the test may become obsolete in the light of new insights, but the hard facts, carefully collected in the field, remain facts and can always be re-evaluated.

8.3.4   Interpreting the Data

Calculating physical properties would be relatively easy if the aquifer system (i.e., aquifer plus well) were precisely known. This is generally not the case, so interpreting aquifer tests is primarily a matter of identifying an unknown system. System identification relies on models, the characteristics of which are assumed to represent the characteristics of the real aquifer system. In Section 8.2, well-flow equations were presented based on idealized theoretical models: homogeneous and isotropic aquifers of seemingly infinite extent, no well bore storage inside the pumped well, etc.

To identify an aquifer system, one must compare its drawdown behavior with that of the various theoretical models. The model that compares best with the real system is then selected for the calculation of the physical properties. System identification includes the construction of diagnostics plots. In the literature, both log-log plots and semi-log plots of the drawdown versus the time are used for this purpose. In the following sections, only the semi-log plots are used for identification and analysis purposes, because (1) they have more diagnostic value than log-log plots due to the occurrence of straight-line segments, and (2) they can also be used to analyze the field data of single-well tests. The characteristic shapes of the curves can help in selecting the appropriate model.

The choice of theoretical model is a crucial step in the interpretation of aquifer tests. If the wrong model is chosen, the physical properties for the real aquifer will not be correct. A troublesome fact is that theoretical solutions to well-flow problems are usually not unique. Some models, developed for different aquifer systems, yield similar responses to a given stress exerted on them. This makes system identification and model selection a difficult affair.

8.4   Time-Drawdown Analyses

All the analysis methods have in common that they are based on graphical solution procedures. For the pumping period, time-drawdown plots when plotted on semi-log paper will show characteristic shapes according to the governing equations as was illustrated in Figures 8.2, 8.4, 8.6, and 8.8. For the recovery period, time-ratio-residual-drawdown plots when plotted on semi-log paper will give characteristic shapes as was illustrated in Figure 8.9.

For the three main types of aquifers, the analysis procedures are now discussed together with sample calculations. These analysis procedures can be applied both to aquifer test data and to single-well test data. With the latter, the value of the specific yield or storativity can, in principle, not be determined.

8.4.1   Aquifer Test in a Confined Aquifer

The physical properties of a confined aquifer can be found by developing the time-drawdown relationship based on Equation (3). If the pumping time is long enough, a plot of the drawdowns observed at a particular distance r from the pumped well versus the logarithm of time t, will appear as a straight line. If the slope of the straight-line segment is expressed as the drawdown difference ($\Delta s = s_1 - s_2$) per log cycle of time (log $t_2/t_1 = 1$), rearranging Equation (3) gives

$$T = \frac{2.3Q}{4\pi\,\Delta s} \tag{17}$$

If this line is extended until it intercepts the time-axis where $s = 0$, the interception point has the coordinates $s = 0$ and $t = t_o$. Substituting these values into Equation (3), after rearrangement, gives

$$S = \frac{2.25Tt_o}{r^2} \tag{18}$$

This procedure will now be illustrated with field data. An aquifer test was made in a confined aquifer. The well was pumped at a constant rate of 528 m³/day for 12 hours. Table 8.2 shows the drawdowns of a piezometer at 90 m distance from the pumped well at various times. Figure 8.12 shows the time-drawdown plot on semi-log paper. From visual inspection, the data plot exhibits a straight line in the time range from 20 to 360 minutes. Through these points a best-fitting straight line was drawn; Figure 8.12 shows that its slope Δs is 21 cm = 0.21 m per log cycle of time. Substitution of the appropriate values into Equation (17) gives

$$T = \frac{2.3\times528}{4\times3.14\times0.21} = 460 \text{ m}^2/\text{day}$$

TABLE 8.2 Time-Drawdown Values Observed in a Piezometer at 90 m Distance from the Pumped Well of an Aquifer Test

Time (min)	Drawdown (cm)	Time (min)	Drawdown (cm)	Time (min)	Drawdown (cm)
1	2.5	30	20.0	150	35.0
2	3.9	40	22.6	360	42.6
4	6.1	50	24.7	550	44.0
6	8.0	60	26.4	720	44.5
9	10.6	90	30.4		
20	16.8	120	33.0		
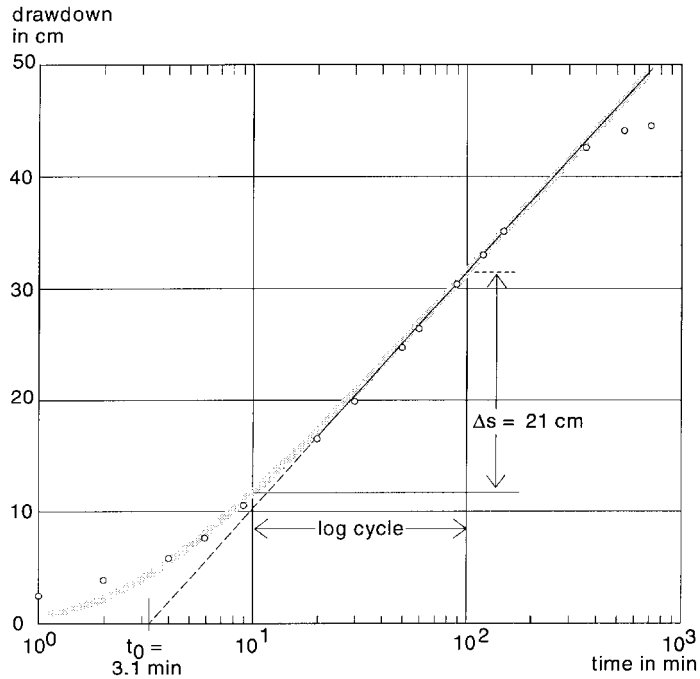
FIGURE 8.12 Time-drawdown plot of field data of an aquifer test in a confined aquifer.

From Figure 8.12 it can be seen that the extrapolated straight line intersects the x-axis where $s = 0$ at t_o = 3.1 min = 2.2×10^{-3} day. Substitution of the appropriate values into Equation (18) yields

$$S = \frac{2.25 \times 460 \times 2.2 \times 10^{-3}}{90^2} = 2.8 \times 10^{-4}$$

Finally, the selected data points to draw the best-fitting straight line in Figure 8.12 should have corresponding $1/u$ values larger than 10, which is a practical condition for the applicability of the Jacob method. Rearranging Equation (2) and substituting $1/u > 10$, this condition gives

$$t > \frac{10r^2 S}{4T} = \frac{10 \times 90^2 \times 2.8 \times 10^{-4}}{4 \times 460} = 1.2 \times 10^{-2} \text{ day} = 18 \text{ min}$$

So, theoretically, all the observed drawdown values with t values larger than 18 minutes have $1/u$ values larger than 10 and can thus be expected to lie on a straight line. Since the straight line was initially drawn through the plotted points in the time range from 20 to 360 minutes, the above condition was already fulfilled. For an explanation for the fact that the last two drawdown values do not fall on the straight line, the reader is referred to Section 8.6.

The well was not observed during the recovery period, so no second estimate for the physical properties could be made based on the time-residual-drawdown relationship. Also no other piezometers were drilled, so the analysis of this aquifer test yielded only one value of the transmissivity and one value of the storativity and thus the consistency of these parameter values could not be checked.

To check the consistency of the analysis results in another way, the following procedure was followed. For all the data points, their u and corresponding $W(u)$ values were calculated and substituted into Equation (1) together with the appropriate values for Q and T. These theoretical drawdown values based

on the Theis equation were plotted as a shaded line in Figure 8.12. It clearly shows that there is a good match between these theoretical drawdown values and those observed in the field. The above physical properties can thus be regarded as representative for the aquifer under consideration. With the software package SATEM (Boonstra, 1989), these theoretical drawdown values are always plotted in the data plot to check the consistency of the analysis results.

Finally it can be noted that the more strict criterion, i.e., that the data points should have $1/u$ values larger than 100, would have resulted in no analysis at all; in that case, only the last three data points in Figure 8.12 ($t > 180$ min) would have been eligible to draw a best-fitting straight line. This test clearly demonstrates that the limiting condition of large $1/u$ values in order to be able to apply the Jacob method, should be treated judiciously.

8.4.2   Aquifer Test in an Unconfined Aquifer

Figure 8.4 shows that even with a delayed yield effect, the time-drawdown curve will eventually exhibit a straight-line segment under a slope for late-time conditions. So, if the pumping time is sufficiently long, the physical properties of an unconfined aquifer can thus also be found by using Equations (17) and (18), provided that the drawdowns are corrected according to Equation (5). This procedure will now be illustrated with field data. An aquifer test was made in a 470-m thick, unconfined aquifer. The well only penetrating the aquifer to a depth of 60 m with a screen length of 40 m, was pumped at a constant rate of 6350 m³/day for 100 hours. Table 8.3 shows the drawdowns observed in a piezometer at a distance of 91.5 m from the pumped well as a function of time. The depth of this piezometer was 46 m with the last two meters screened.

TABLE 8.3 Time-Drawdown Values Observed in a Piezometer at 91.5 m Distance from the Pumped Well of an Aquifer Test

Time (min)	Drawdown (cm)	Time (min)	Drawdown (cm)	Time (min)	Drawdown (cm)
1	1.0	40	18.0	600	25.9
2	1.0	50	18.6	750	26.5
3	2.1	60	19.2	1000	27.4
4	3.4	75	20.1	1250	28.7
6	6.1	100	20.7	1500	29.3
8	8.8	125	21.3	1750	29.9
10	11.0	150	21.6	2000	30.5
12	12.2	175	21.9	2500	32.6
15	13.4	200	22.3	3000	33.5
18	14.3	250	22.9	4000	35.7
21	15.2	300	23.8	5000	36.9
25	15.8	400	24.4	6000	36.9
30	16.8	500	25.3		

The first step is to determine whether the observed drawdown values are small compared to the aquifer's thickness. The last observed drawdown value in Table 8.3 is less than 5% of the aquifer thickness. So, the uncorrected drawdown values can thus be used in the analysis. Figure 8.13 shows the time-drawdown plot on semi-log paper. From visual inspection, the data plot exhibits a straight line in the time range from 30 to 1000 minutes. Through these points, a best-fitting straight line was drawn; Figure 8.13 shows that its slope Δs is 7 cm = 0.07 m per log cycle of time. Substitution of the appropriate values into Equation (17) gives

$$T = \frac{2.3 \times 6350}{4 \times 3.14 \times 0.07} = 16603 \text{ m}^2/\text{day}$$
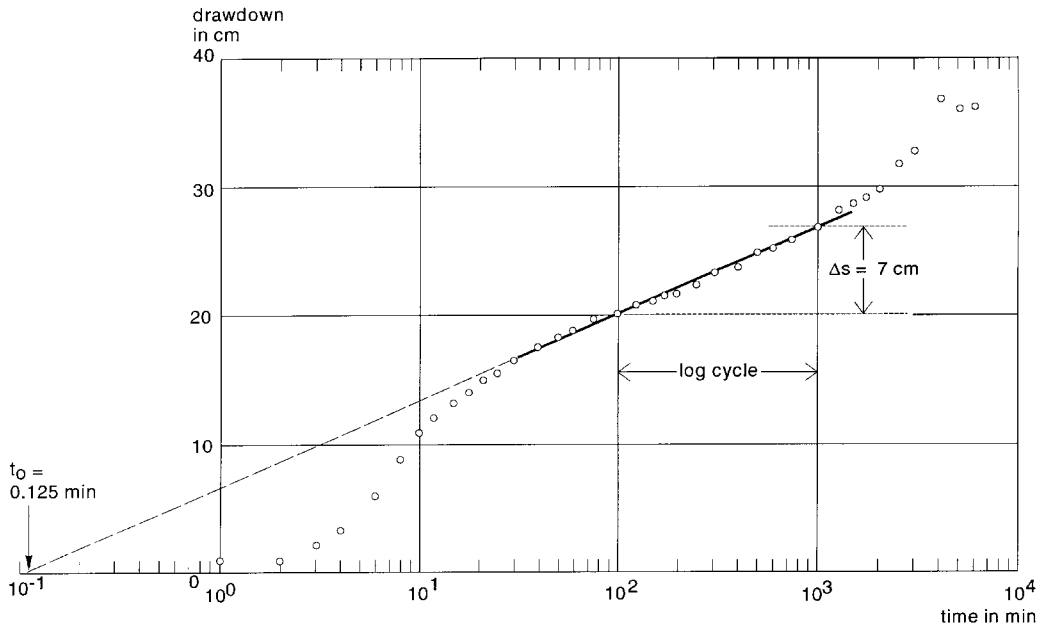
FIGURE 8.13 Time-drawdown plot of field data of an aquifer test in an unconfined aquifer with partially penetrating pumped well.

Equation (18) cannot be used to estimate the specific yield due to the fact that there is a partially penetrating pumped well. Substituting the values $s = 0$ and $t = t_o$ into Equation (12), the specific yield value can then be obtained by

$$S_y = \frac{2.25 T t_o}{r^2} e^{f_s} \tag{19}$$

From Figure 8.13 it can be seen that the extrapolated straight line intersects the x-axis where $s = 0$ at t_o $= 0.125$ min $= 8.7 \times 10^{-5}$ day. Before the value of S_y can be calculated, the value of f_s needs to be determined. Table 8.4 shows the different components of the first five values in the summation series of Equation (11) with $b = 60$ m, $d = 20$ m, $b' = 46$ m, $d' = 44$ m, and $D = 470$ m. Substituting the appropriate values into Equation (11) gives

$$f_s = \frac{4 \times 470}{3.14(60 - 20)} \times 0.2641 = 14.96 \times 0.2641 = 3.95$$

TABLE 8.4 Values of Components in Summation Series of Equation (11)

n	K_o	sin	sin	cos	component	summation
1	0.7629	0.3904	0.1333	0.9551	0.1874	0.1874
2	0.3095	0.7188	0.2642	0.8244	0.0580	0.2454
3	0.1401	0.9332	0.3904	0.6197	0.0157	0.2611
4	0.0665	0.9994	0.5096	0.3594	0.0029	0.2640
5	0.0326	0.9071	0.6197	0.0668	0.0001	0.2641

Substitution of the appropriate values into Equation (19) then yields

$$S_y = \frac{2.25 \times 16603 \times 8.7 \times 10^{-5}}{91.5^2} e^{3.95} = 3.9 \times 10^{-4} \times 51.9 = 0.02$$

Finally, the limiting condition to use Equation (12) for evaluating the physical properties of the tested unconfined aquifer should be checked:

$$t > \frac{D^2 S_y}{2T} = \frac{470^2 \times 0.02}{2 \times 16603} = 0.133 \text{ day} = 192 \text{ min}$$

So, theoretically, all the observed drawdown values with t values larger than 192 minutes, can be expected to lie on a straight line. Since the straight line was initially drawn through the plotted points in the time range from 30 to 1000 minutes, the above condition was thus violated. When another straight line is drawn through the points in the time range from 200 to 1000 minutes, its slope will be somewhat steeper. A repetition of the above calculations then yields a transmissivity value of 15,828 m²/day and a specific yield value of 0.016. These values do not differ significantly from the first ones.

For this aquifer test, two other piezometers were drilled at distances of 15.2 and 30.5 m from the pumped well (Boonstra, 1992). The time-drawdown analyses from these piezometers gave two other estimates of the physical properties: transmissivities of 16,691 and 14,104 m²/day and specific yield values of 0.011 and 0.013, respectively. This nicely illustrates the fact that with aquifer tests with more than one piezometer, a comparison between the various estimates of the physical properties can be made. When the various estimates are consistent in their values as with this test, more reliability can be given to the analyses results. It should be noted that the range in transmissivity values is usually smaller than the range in specific yield values due to the nature of the analysis procedure. This phenomenon also pertains to aquifer test results in confined and leaky aquifers.

8.4.3 Aquifer Test in a Leaky Aquifer

The physical properties of a leaky aquifer can be found by developing the time-drawdown relationship based on Equation (6). Figure 8.6 shows that the Hantush well function exhibits an inflection point and, for large values of $1/u$, a horizontal straight-line segment indicating steady-state flow. The inflection-point method was based on these phenomena. Hantush (1956) showed that for the inflection point (s_p, t_p) the following relationships hold

a) The drawdown value s_p is given by

$$s_p = 0.5 s_m \tag{20}$$

b) The u_p value is given by

$$u_p = \frac{r^2 S}{4 T t_p} = \frac{r}{2L} \tag{21}$$

c) The slope of the curve at the inflection point Δs_p per log cycle of time is given by

$$\Delta s_p = \frac{2.3 Q}{4 \pi T} e^{-r/L} \tag{22}$$

d) At the inflection point, the relation between the drawdown and the slope of the curve is given by

$$2.3 \frac{s_p}{\Delta s_p} = e^{r/L} K_o \left(r/L \right) \qquad (23)$$

This procedure will now be illustrated with field data. An aquifer test was made in a leaky aquifer. The well was pumped at a constant rate of 545 m³/day for 48 hours. One of the piezometers was located 20 m away from the pumped well. Table 8.5 shows the observed drawdowns as a function of time. Figure 8.14 shows that steady-state conditions did occur at the end of the pumping test (s_m = 2.44 m). According to Equation (20), the s_p value is then 1.22 m. Locating the inflection point in Figure 8.14 using this s_p value gives a t_p value of 22 min = 1.5×10^{-2} day.

TABLE 8.5 Time-Drawdown Values Observed in a Piezometer at 20 m Distance from the Pumped Well of an Aquifer Test

Time (min)	Drawdown (m)	Time (min)	Drawdown (m)	Time (min)	Drawdown (m)
1	0.265	60	1.615	720	2.320
2	0.347	70	1.675	840	2.337
3	0.490	80	1.725	960	2.355
4	0.584	90	1.767	1080	2.377
5	0.635	120	1.895	1200	2.373
6	0.707	150	1.977	1320	2.385
7	0.793	180	2.020	1440	2.400
8	0.839	210	2.075	1620	2.410
9	0.882	240	2.113	1800	2.420
10	0.930	300	2.170	1980	2.440
15	1.079	360	2.210	2160	2.440
20	1.187	420	2.245	2340	2.450
25	1.275	480	2.270	2520	2.450
30	1.345	540	2.281	2700	2.440
40	1.457	600	2.298	2880	2.440
50	1.554	660	2.310		
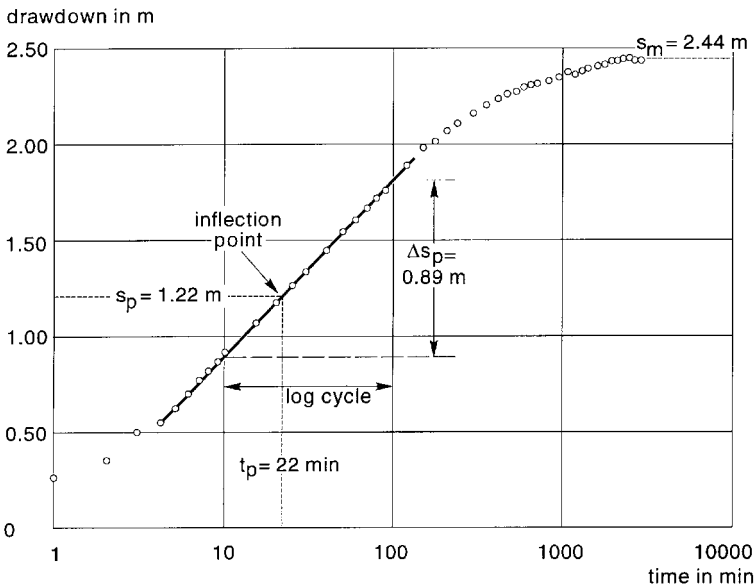
FIGURE 8.14 Time-drawdown plot of field data of an aquifer test in a leaky aquifer.

The slope of the curve at this inflection point is 0.89 m per log cycle of time. Substitution of the appropriate values into Equation (23) yields

$$e^{r/L} K_o\left(r/L\right) = 2.30 \frac{s_p}{\Delta s_p} = 2.3 \frac{1.22}{0.89} = 3.15$$

A table with r/L, $K_o(r/L)$, and $e^{r/L} K_o(r/L)$ values is provided in Appendix 8.1. The r/L value is then 0.0575. Substitution of the r value yields $L = 20/0.0575 = 348$ m. Rearranging Equation (22) and substituting the appropriate values yields

$$T = \frac{2.3Q}{4\pi \Delta s_p} e^{-r/L} = \frac{2.3 \times 545}{4 \times 3.14 \times 0.89} e^{-0.0575} = 106 \text{ m}^2/\text{day}$$

The c value is then $L^2/T = 348^2/106 = 1142$ days. Rearranging Equation (21) and substituting the appropriate values yields

$$S = \frac{4Tt_p}{r^2} \frac{r}{2L} = \frac{4 \times 106 \times 1.5 \times 10^{-2}}{20^2} \times \frac{20}{2 \times 348} = 4.6 \times 10^{-4}$$

It should be noted that the accuracy of the calculated physical properties depends to a large extent on the accuracy of the value of the (extrapolated) steady-state drawdown s_m. The calculations should therefore be checked by substituting the different values into Equation (6); calculations of s should be made for the observed values of t. If the values of t are not too small, the values of s should fall on the observed data curve. If the calculated data deviate from the observed data, the (extrapolated) value of s_m should be adjusted. Sometimes, the observed data curve can be drawn somewhat steeper or flatter through the plotted points, and so Δs_p can be adjusted too. With the new values of s and/or Δs_p, the calculation is then repeated.

With the software package SATEM (Boonstra, 1989), this time-consuming work can be avoided. This program follows the same procedure as described above. In addition, it displays the drawdowns calculated with Equation (6) on the monitor, together with the data observed in the field. This makes it easy to check whether the correct (extrapolated) steady-state drawdown s_m has been selected.

8.4.4 Single-Well Tests

With single-well tests, basically the same procedures can be applied as with aquifer tests. The r value now represents the effective radius of the single well. This is difficult to determine under field conditions; as a "best" estimate, the outer radius of the well screen is often used.

A complicating factor is the phenomenon that, due to nonlinear well losses, the water levels inside the well can be considerably lower than those directly outside the well screen. This implies that drawdown data from the pumped well can, in general, only be used for the analysis when corrected for these nonlinear well losses using the results of so-called step-drawdown tests (see Chapter 9).

The procedure with single-well tests will now be illustrated with field data from an unconfined aquifer. The well was pumped at a constant rate of 3853 m³/day for 10 hours. The outer radius of the well screen was 0.20 m. Table 8.6 shows the observed drawdowns as a function of time. At the same site, a step-drawdown test was also made. From the analysis of this test, the nonlinear well loss was calculated as 1.93 m (see Chapter 9 under the heading *Jacob's Method*).

The first step is to determine whether the observed drawdown values are small compared with the aquifer's thickness. The depth of the pumped borehole was 271 m. Substituting this value and the last observed drawdown value corrected for nonlinear well loss into Equation (5) gives a maximum correction

TABLE 8.6 Time-Drawdown Values of a Single-Well Test

Time (min)	Drawdown (m)	Time (min)	Drawdown (m)	Time (min)	Drawdown (m)
15	4.161	50	4.486	240	4.808
20	4.283	56	4.471	300	4.776
25	4.257	60	4.474	360	4.885
30	4.357	80	4.534	420	4.960
36	4.358	105	4.618	480	4.906
40	4.399	120	4.672	540	4.972
46	4.456	180	4.748	600	5.016

value of $(5.016 - 1.93)^2/(2 \times 271) = 0.02$ m. So, no corrections on the observed drawdowns are needed with respect to Equation (5).

Figure 8.15 shows the time-drawdown plot of the observed drawdowns in the pumped well corrected for nonlinear well loss on semi-log paper. From visual inspection, the data plot exhibits a straight line over the whole time range. Through all the points a best-fitting straight line was drawn; Figure 8.15 shows that its slope Δs is 0.50 m per log cycle of time. Substitution of the appropriate values into Equation (17) gives

$$T = \frac{2.3 \times 3853}{4 \times 3.14 \times 0.50} = 1410 \text{ m}^2/\text{day}$$
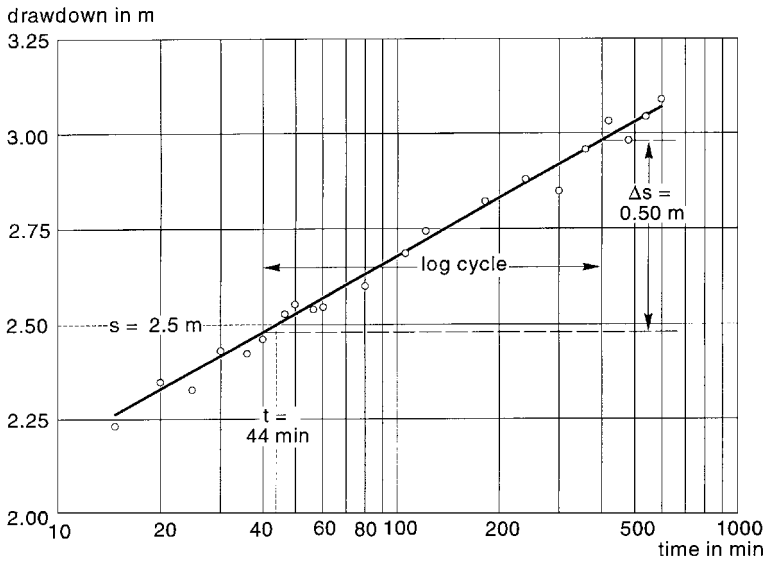
FIGURE 8.15 Time-drawdown plot of field data of a single-well test in an unconfined aquifer.

From Figure 8.15 it can be seen that the intersection of the straight line with the x-axis where $s = 0$ cannot be determined directly. When such a situation occurs, the following procedure can be used. The Δs value was calculated as 0.50 m. Figure 8.15 shows that for a drawdown of 2.50 m — being 5 times the Δs value — the corresponding time is 44 minutes. The t_0 value is then equal to 44×10^{-5} min $= 3.1 \times 10^{-7}$ day. Substitution of the appropriate values into Equation (18) yields

$$S_y = \frac{2.25 \times 1410 \times 3.1 \times 10^{-7}}{0.20^2} 0.025$$

Finally, the selected data points to draw the best-fitting straight line in Figure 8.15 should have corresponding $1/u$ values larger than 10, which is a practical condition for the applicability of the Jacob method. Rearranging Equation (2) and substituting $1/u > 10$, this condition gives

$$t > \frac{10r^2 S_y}{4T} = \frac{10 \times 0.2^2 \times 0.025}{4 \times 1410} = 1.8 \times 10^{-6} \text{ day} = 2.6 \times 10^{-3} \text{ min}$$

So, theoretically, all the observed drawdown values with t values larger than 2.6×10^{-3} minutes, can be expected to lie on a straight line. In other words, all the observed drawdown data as plotted in Figure 8.15 can be used to determine the slope of the straight line. It should be noted that the condition $1/u > 10$ is usually fulfilled in single-well tests because of the small r value.

With the above procedure, we can also use the uncorrected drawdown data and still get representative transmissivity values, because the slope of the straight-line segment in the time-drawdown plot on semi-log paper is not affected by the constant value of the nonlinear well loss. Specific-yield values based on the corrected data, should, however, be treated with caution, because they are highly sensitive to the value of the effective radius of the pumped well.

For this single-well test, field measurements were also made during the recovery period. Table 8.7 shows the observed residual drawdowns as a function of time. Figure 8.16 shows the plot of residual drawdown versus time-ratio on semi-log paper. From visual inspection, the data plot exhibits a straight line in the time-ratio range from 2 to 40. Through these points a best-fitting straight line was drawn; Figure 8.16 shows that its slope $\Delta s'$ is 0.56 m per log cycle of time ratio. Rearranging Equation (15) and assuming that $S_y' = S_y$ gives

$$T = \frac{2.3Q}{4\pi \, \Delta s'} = \frac{2.3 \times 3853}{4 \times 3.14 \times 0.56} = 1260 \text{ m}^2/\text{day}$$

TABLE 8.7 Time-Residual Drawdown Values of a Single-Well Test

Time (min)	Residual Drawdown (m)	Time (min)	Residual Drawdown (m)	Time (min)	Residual Drawdown (m)
5	0.888	46	0.588	240	0.226
10	0.847	50	0.563	330	0.158
15	0.817	56	0.517	450	0.193
20	0.760	60	0.514	570	0.060
25	0.683	75	0.460	630	0.075
30	0.648	105	0.393	750	0.015
34	0.636	120	0.321	870	0.015
40	0.588	180	0.310	990	0.001

From Figure 8.16 it can be seen that the straight line intersects the axis where $s'(r,t/t') = 0$, at $(t/t')_0 = 1.39$. Substituting the appropriate values into Equation (15) yields

$$S_y' = \frac{S_y}{1.39} = 0.7 S_y$$

The above implies that the specific-yield value during recovery is less than during pumping. This phenomenon is often encountered because of air entrapment when the pores are again filled with water.

The above transmissivity value corresponds well with that found from the time-drawdown analysis. The application of time-drawdown and time-recovery analyses thus enables us to check the calculated transmissivity value. When the two values are close to each other, it implies that the data are consistent, i.e., that the results of the test are reliable.
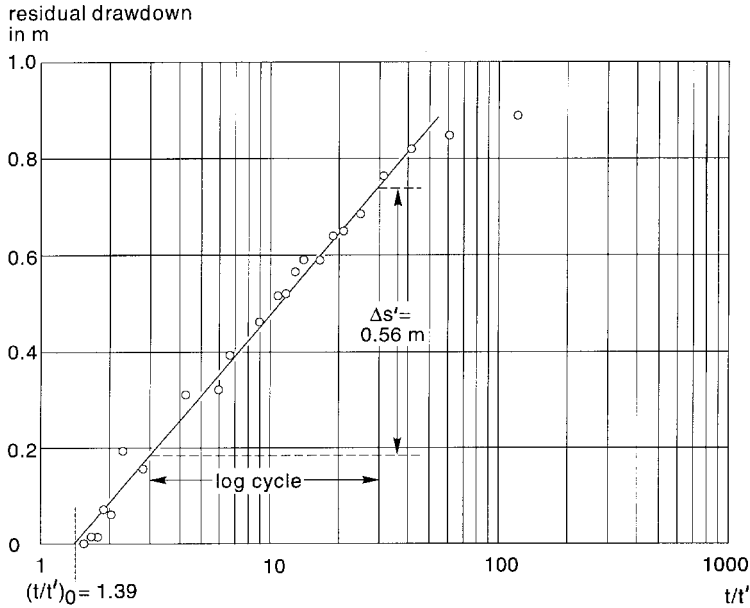
FIGURE 8.16 Time-ratio-residual-drawdown plot of field data of a single-well test in an unconfined aquifer.

8.5   Distance-Drawdown Analyses

These analysis methods also have in common that they are based on graphical solution procedures. For all three main types of aquifers, plots of steady-state drawdowns of piezometers versus the logarithms of the distances of these piezometers from the pumped well will show characteristic shapes; we will refer to these diagnostic plots as distance-drawdown plots. Steady-state drawdowns of at least two piezometers are required for this type of analysis. The presented distance-drawdown analyses can only be applied to aquifer tests with fully penetrating pumped wells. The advantage of distance-drawdown analyses is that an independent estimate of the aquifer's transmissivity can be made, in addition to the estimates obtained from the time-drawdown analyses of each piezometer separately. For leaky aquifers, an independent estimate of the aquitard's hydraulic resistance can be made as well.

8.5.1   Confined and Unconfined Aquifers

Theoretically, in confined and unconfined aquifers which are pumped at a constant rate, there will always be an unsteady state, because these aquifers are not recharged by an outside source. Instead of steady state, the concept of pseudo-steady-state is used for these types of aquifers. It is said that the aquifer is in pseudo-steady-state when the hydraulic gradient has become constant, i.e., when the time-drawdown curves are running parallel (see Figure 8.11).

   The physical properties of a confined or unconfined aquifer can be found by developing the distance-drawdown relationship based on Equation (4). From this equation it can be seen that the corresponding distance-drawdown plot will show a straight line. If the slope of this straight line is expressed as the drawdown difference ($\Delta s = s_1 - s_2$) per log cycle of distance ($\log r_2/r_1 = 1$), rearranging Equation (4) gives

$$T = \frac{2.3Q}{4\pi \Delta s} \tag{24}$$

This procedure will now be illustrated with field data. An aquifer test was made in a shallow unconfined aquifer; prior to pumping, the aquifer's saturated thickness was only 6.5 m. The well was pumped at a constant rate of 167 m³/day for 520 minutes. The water table was observed in seven piezometers. Figure 8.17 shows the time-drawdown graphs of these seven wells; the plotted drawdowns were already corrected using Equation (5). As can be seen from this figure, the curves run almost parallel in the last few hours of the test. Table 8.8 shows the drawdown values in the seven piezometers at the end of the test; these were assumed to represent the pseudo-steady-state drawdowns. Figure 8.18 shows the distance-drawdown plot on semi-log paper. From visual inspection, the data plot exhibits a straight line in the distance range from 4 to 20 m. Through these points a best-fitting straight line was drawn; Figure 8.18 shows that its slope Δs is 0.22 m per log cycle of distance. Substitution of the appropriate values into Equation (24) gives

$$T = \frac{2.3 \times 167}{2 \times 3.14 \times 0.22} = 278 \text{ m}^2/\text{day}$$
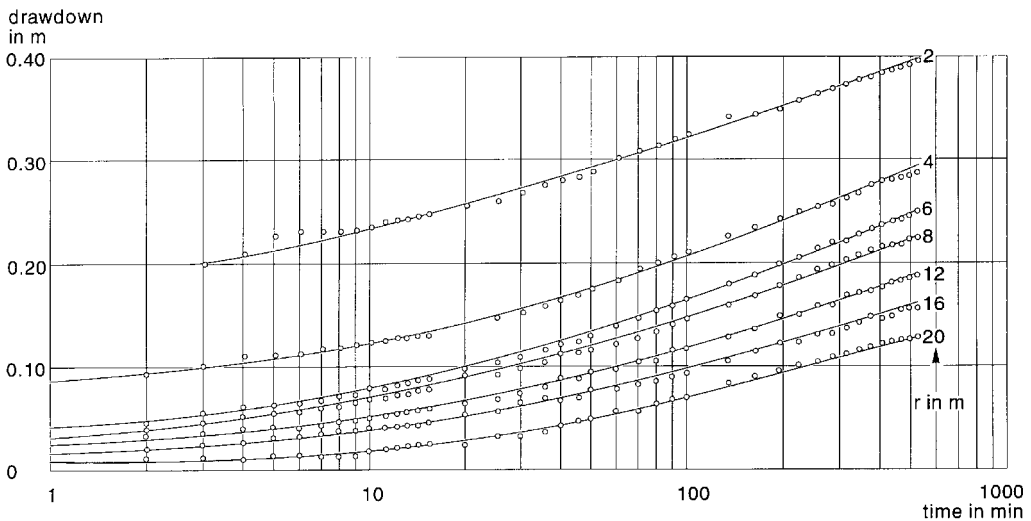
FIGURE 8.17 Time-drawdown curves of seven piezometers showing field data of an aquifer test in an unconfined aquifer.

TABLE 8.8 Pseudo-Steady-State Drawdown Values of an Aquifer Test

Distance to Pumped Well (m)	Pseudo-Steady-State Drawdown	
	Uncorrected (m)	Corrected (m)
2	0.407	0.394
4	0.294	0.287
6	0.252	0.247
8	0.228	0.224
12	0.193	0.190
16	0.161	0.159
20	0.131	0.130

The fact that the drawdown of the piezometer at 2 m distance is not lying on the straight line can be explained by an additional head loss which usually occurs near the well because of the relatively strong curvature of the water table. It should be noted that, because the aquifer is very thin, the piezometers were drilled at relatively short distances from the pumped well. The above transmissivity value corre-
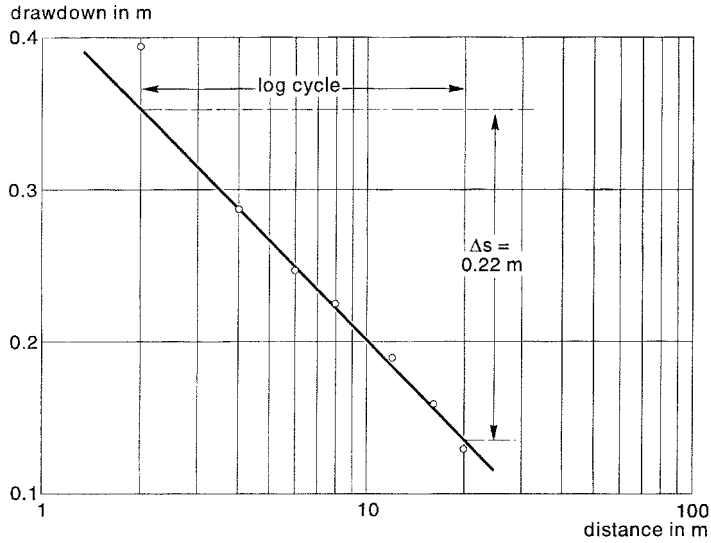
FIGURE 8.18 Distance-drawdown plot of field data of an aquifer test in an unconfined aquifer.

sponds well with those found from the time-drawdown analyses based on the individual piezometers; the latter ranged from 210 to 260 m²/day.

8.5.2 Leaky Aquifers

The physical properties of a leaky aquifer can be found by developing the distance-drawdown relationship based on Equation (8). The transmissivity of a leaky aquifer can also be found from Equation (24) (Boonstra and de Ridder, 1994). If the straight line is extended until it intercepts the distance-axis where $s_m = 0$, the interception point has the coordinates $s_m = 0$ and $r = r_0$. Substituting these values into Equation (8) gives

$$\log \frac{1.12L}{r_0} = 0 \quad \text{or} \quad \frac{1.12L}{r_0} = 1 \quad \text{or} \quad L = \frac{r_0}{1.12} \tag{25}$$

This procedure will now be illustrated with field data. An aquifer test was made in a leaky aquifer with a thickness of 10 m. The well was pumped at a constant rate of 761 m³/day for 480 minutes. Table 8.9 shows the extrapolated steady-state drawdowns in the six piezometers. Figure 8.19 shows the distance-drawdown plot on semi-log paper. From visual inspection, the data plot exhibits a straight line over the whole distance range. Through all the points a best-fitting straight line was drawn; Figure 8.19 shows that its slope Δs_m is 0.14 m per log cycle of distance. Substitution of the appropriate values into Equation (24) gives

$$T = \frac{2.3 \times 761}{2 \times 3.14 \times 0.14} = 1990 \text{ m}^2/\text{day}$$

From Figure 8.19 it can be seen that the extrapolated straight line intersects the x-axis where $s_m = 0$ at $r_o = 1000$ m. Substitution of this value into Equation (25) gives

$$L = \frac{r_0}{1.12} = \frac{1000}{1.12} = 890 \text{ m}$$

TABLE 8.9 Extrapolated Steady-State Drawdown Values of an Aquifer Test

Distance from Pumped Well (m)	Steady-State Drawdown (m)
10	0.282
30	0.235
60	0.170
90	0.147
120	0.132
400	0.059
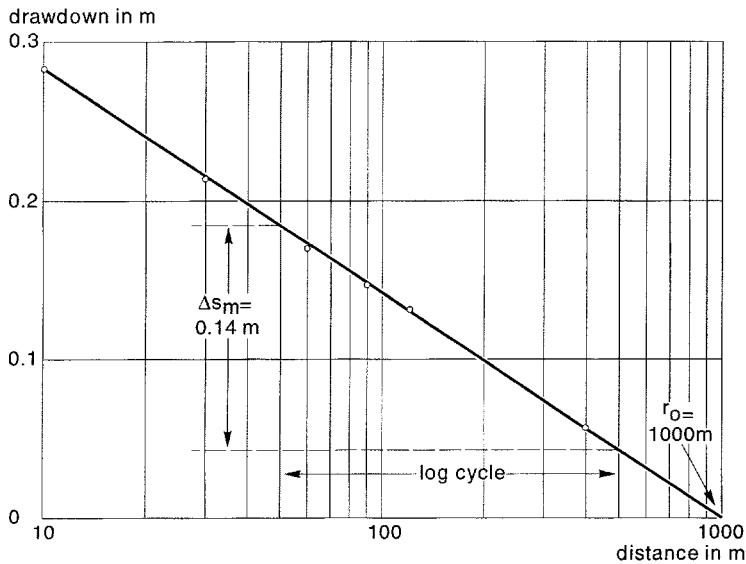
FIGURE 8.19 Distance-drawdown plot of field data of an aquifer test in a leaky aquifer.

The c value can now be calculated as $c = L^2/T = 890^2/1990 = 398$ days. Finally, the limiting conditions should be checked. Substituting the above L value into the condition $L > 3D$ gives $D < 296$ m, so this condition is fulfilled. Substituting the appropriate values into the condition $r/L < 0.2$ gives $r < 178$ m. According to this condition, the drawdown value of the piezometer at a distance of 400 m should be eliminated from the analysis. Figure 8.19, however, shows that this point, too, lies on the straight line, so in this case this condition is not a limiting factor. The above transmissivity and hydraulic resistance values correspond well with those found from the time-drawdown analyses based on the individual piezometers; the latter ranged from 1730 to 1980 m²/day for the aquifer's transmissivity and from 350 to 510 days for the aquitard's hydraulic resistance.

8.6   Concluding Remarks

The diagnostic plots of time-drawdown data based on the well-flow equations are theoretical curves. The time-drawdown curves based on field data will often deviate from these theoretical shapes. These deviations can stem from the fact that one or more of the general assumptions and conditions listed in Section 8.2 are not met in the field, or that the method selected is not the correct one for the test site. Some of the common departures from the theoretical curves will now be discussed.

All real aquifers are limited by geological or hydrological boundaries. If, however, at the end of the pumping period, no such boundaries have been met within the cone of depression, it is said that the aquifer has a seemingly infinite areal extent. When the cone of depression intersects an impervious boundary (e.g., a fault or an impermeable valley wall), it can expand no farther in that direction. The

cone must expand and deepen more rapidly at the fault or valley wall to maintain the yield of the well. This will result in time-drawdown plots in which the last part of the late-time drawdown data will deviate from a straight line under a slope. This part of the plot should be disregarded when the slope of the straight-line segment is being determined. The same phenomenon can occur when, in one of the directions, the sediments become finer and the hydraulic conductivity decreases. Well interference where two wells are close to each other will also result in a similar phenomenon (see Chapter 9).

An opposite effect is encountered when the cone of depression intersects an open water body. If the open water body is hydraulically connected with the aquifer, the aquifer is recharged at an increasing rate as the cone of depression spreads with time. This results in a flattening of the slope of the time-drawdown curve at later times, which resembles the recharge that occurs in a leaky aquifer. The same phenomenon occurs when, in one of the directions, the hydraulic conductivity and/or the aquifer thickness increases. The above cases will result in time-drawdown plots in which the last part of the late-time drawdown data will deviate from a straight line under a slope. This part of the plot should be disregarded when the slope of the straight-line segment is being determined.

It will be clear that there are various reasons why time-drawdown data depart from the theoretical curves. It will also be clear that different phenomena can cause approximately the same anomalies. So, if one is to make a correct analysis, one must have a proper knowledge of the geology of the test site. Because, unfortunately, this knowledge is often fragmentary, determining physical properties is more an art than a science. This is one of the main reasons it is strongly recommended to continue to monitor the water table behavior during the recovery period. This allows a second estimate of the aquifer's transmissivity, which can then be compared with the one found during the pumping period. Even with single well tests, this second estimate is possible. With computer software, even a third estimate of the transmissivity and a second estimate for the specific yield or storativity is then possible (Boonstra, 1989).

Finally, a few remarks on the difference between aquifer tests and single-well tests. The results of aquifer tests are more reliable and more accurate than those of single-well tests. Another advantage is that aquifer tests allow estimates to be made of both the aquifer's transmissivity and its specific yield or storativity, which is not possible with single-well tests. Further, if an aquifer test uses more than one piezometer, separate estimates of the physical properties can be made for each well, allowing the various values to be compared. Moreover, one can make yet another estimate of the physical properties by using not only the time-drawdown relationship, but also the distance-drawdown relationships.

For Further Information

Kruseman and de Ridder (1990) provide an excellent, well-illustrated handbook to well-flow equations that cover a wider range of conditions than presented here, i.e., bounded, sloping, and consolidated fractured aquifers and their application to field data.

Boonstra (1989) presents a software package that facilitates the analysis of aquifer test data of (un)confined and leaky aquifers, including partially penetrating wells, based on time-drawdown, distance-drawdown, and time-ratio-residual-drawdown relationships.

See References for full citations of these works.

References

Boonstra, J. 1989. *SATEM: Selected aquifer test evaluation methods : a microcomputer program.* ILRI Publication 48, Wageningen, 80 p.

Boonstra, J. 1992. Aquifer tests with partially penetrating wells: theory and practice. *J. Hydrol.* 137: 165-179.

Boonstra, J. and de Ridder, N. A. 1994. Single well and aquifer tests, in *Drainage Principles and Applications.* Ed. H.R. Ritsema. ILRI Publication 16, Wageningen, 341-381.

Bos, M. G. Ed. 1989. *Discharge Measurement Structures.* 3rd rev. ed. ILRI Publication 20, Wageningen, 401 p.

Cooper, H. H. and Jacob, C. E. 1946. A generalized graphical method for evaluating formation constants and summarizing well field history. *Am. Geophys. Union Trans.* 27, 526-534.

De Glee, G. J. 1930. *Over grondwaterstromingen bij wateronttrekking door middel van putten.* Waltman, Delft, 175 p.

De Glee, G. J. 1951. Berekeningsmethoden voor de winning van grondwater, in *Winning van grondwater : derde vacantiecursus in drinkwatervoorziening.* Delft University of Technology. 38-80.

Driscoll, F. G. 1986. *Groundwater and Wells.* 2nd Ed. Johnson Division, St. Paul, 1089 p.

Hantush, M. S. 1956. Analysis of data from pumping tests in leaky aquifers. *Am. Geophys. Union Trans.*, 37, 702-714.

Hantush, M. S. 1962. Aquifer tests on partially-penetrating wells. *Am. Soc. Civ. Eng. Trans.*, 127, Part I, 284-308.

Hantush, M. S. 1964. Hydraulics of wells, in V.T. Chow Ed. *Advanced Hydroscience 1*, Academic Press, New York. 281-432.

Hantush, M. S. and Jacob, C. E. 1955. Non-steady radial flow in an infinite leaky aquifer. *Am. Geophys. Union Trans.* 36, 95-100.

Huisman, L. 1972. *Groundwater Recovery.* MacMillan. 366 p.

Jacob, C. E. 1944. Notes on determining permeability by pumping tests under water table conditions. *U.S. Geol. Surv.* Open File Rept.

Jacob, C. E. 1950. Flow of groundwater, in H. Rouse Ed. *Engineering Hydraulics.* Wiley, New York. 321-386.

Kruseman, G. P. and de Ridder, N. A. 1990. *Analysis and Evaluation of Pumping Test Data.* ILRI Publication 47, Wageningen, 377 p.

Neuman, S. P. 1972. Theory of flow in unconfined aquifers considering delayed response of the water table. *Water Resources Res.* 8, 1031-1045.

Neuman, S. P. 1975. Analysis of pumping test data from anisotropic unconfined aquifers considering delayed gravity response. *Water Resources Res.* 11, 329-342.

Theis, C. V. 1935. The relation between the lowering of the piezometric surface and the rate and duration of discharge of a well using groundwater storage. *Am. Geophys. Union Trans.* 16, 519-524.

Thiem, G. 1906. *Hydrologische Methoden.* Gebhardt, Leipzig, 56 p.

Glossary

Aquiclude Water-bearing layer in which both the horizontal and vertical flow components are so small that they can be neglected. In an aquiclude, the groundwater flow is assumed to be zero.

Aquifer Water-bearing layer in which the vertical flow component with respect to the horizontal flow component is so small that it can be neglected. In an aquifer, the groundwater flow is assumed to be predominantly horizontal.

Aquitard Water-bearing layer in which the horizontal flow component with respect to the vertical flow component is so small that it can be neglected. In an aquitard, the groundwater flow is assumed to be predominantly vertical.

Drawdown Change in observed water level as a function of pumping time with respect to original water level prior to pumping.

Homogeneous An aquifer or aquitard is called homogeneous when its hydraulic conductivity is the same throughout such a water-bearing layer; its value is thus independent of the location.

Isotropic An aquifer or aquitard is called isotropic when its hydraulic conductivity is the same in all directions at a particular location in such a water-bearing layer; its value is thus independent of the orientation.

Piezometer Small-diameter pipe placed in the subsoil so that there is no leakage around the pipe and all water enters the pipe through its open bottom. A piezometer indicates only the hydrostatic pressure of the groundwater at the specific point in the subsoil at its open lower end. To prevent clogging, piezometers are sometimes equipped with a short screen of a few decimeters length.

Residual drawdown Change in observed water level as a function of recovery time with respect to original water level prior to pumping.

APPENDIX 8.1 Values of $K_o(r/L)$ and $e^{r/L} K_o(r/L)$ as Function of r/L

r/L	$K_o(r/L)$	$e^{r/L} K_o(r/L)$	r/L	$K_o(r/L)$	$e^{r/L} K_o(r/L)$	r/L	$K_o(r/L)$	$e^{r/l} K_o(r/L)$	r/L	$K_o(r/L)$	$e^{r/L} K_o(r/L)$
1.0(−2)	4.72	4.77	3.8(−2)	3.39	3.52	6.6(−2)	2.84	3.03	9.4(−2)	2.49	2.73
1.1(−2)	4.63	4.68	3.9(−2)	3.36	3.50	6.7(−2)	2.82	3.02	9.5(−2)	2.48	2.72
1.2(−2)	4.54	4.59	4.0(−2)	3.34	3.47	6.8(−2)	2.81	3.01	9.6(−2)	2.47	2.72
1.3(−2)	4.46	4.52	4.1(−2)	3.31	3.45	6.9(−2)	2.79	2.99	9.7(−2)	2.46	2.71
1.4(−2)	4.38	4.45	4.2(−2)	3.29	3.43	7.0(−2)	2.78	2.98	9.8(−2)	2.45	2.70
1.5(−2)	4.32	4.38	4.3(−2)	3.26	3.41	7.1(−2)	2.77	2.97	9.9(−2)	2.44	2.69
1.6(−2)	4.25	4.32	4.4(−2)	3.24	3.39	7.2(−2)	2.75	2.96	1.0(−1)	2.43	2.68
1.7(−2)	4.19	4.26	4.5(−2)	3.22	3.37	7.3(−2)	2.74	2.95	1.1(−1)	2.33	2.60
1.8(−2)	4.13	4.21	4.6(−2)	3.20	3.35	7.4(−2)	2.72	2.93	1.2(−1)	2.25	2.53
1.9(−2)	4.08	4.16	4.7(−2)	3.18	3.33	7.5(−2)	2.71	2.92	1.3(−1)	2.17	2.47
2.0(−2)	4.03	4.11	4.8(−2)	3.15	3.31	7.6(−2)	2.70	2.91	1.4(−1)	2.10	2.41
2.1(−2)	3.98	4.06	4.9(−2)	3.13	3.29	7.7(−2)	2.69	2.90	1.5(−1)	2.03	2.36
2.2(−2)	3.93	4.02	5.0(−2)	3.11	3.27	7.8(−2)	2.67	2.89	1.6(−1)	1.97	2.31
2.3(−2)	3.89	3.98	5.1(−2)	3.09	3.26	7.9(−2)	2.66	2.88	1.7(−1)	1.91	2.26
2.4(−2)	3.85	3.94	5.2(−2)	3.08	3.24	8.0(−2)	2.65	2.87	1.8(−1)	1.85	2.22
2.5(−2)	3.81	3.90	5.3(−2)	3.06	3.22	8.1(−2)	2.64	2.86	1.9(−1)	1.80	2.18
2.6(−2)	3.77	3.87	5.4(−2)	3.04	3.21	8.2(−2)	2.62	2.85	2.0(−1)	1.75	2.14
2.7(−2)	3.73	3.83	5.5(−2)	3.02	3.19	8.3(−2)	2.61	2.84	2.1(−1)	1.71	2.10
2.8(−2)	3.69	3.80	5.6(−2)	3.00	3.17	8.4(−2)	2.60	2.83	2.2(−1)	1.66	2.07
2.9(−2)	3.66	3.76	5.7(−2)	2.98	3.16	8.5(−2)	2.59	2.82	2.3(−1)	1.62	2.04
3.0(−2)	3.62	3.73	5.8(−2)	2.97	3.14	8.6(−2)	2.58	2.81	2.4(−1)	1.58	2.01
3.1(−2)	3.59	3.70	5.9(−2)	2.95	3.13	8.7(−2)	2.56	2.80	2.5(−1)	1.54	1.98
3.2(−2)	3.56	3.67	6.0(−2)	2.93	3.11	8.8(−2)	2.55	2.79	2.6(−1)	1.50	1.95
3.3(−2)	3.53	3.65	6.1(−2)	2.92	3.10	8.9(−2)	2.54	2.78	2.7(−1)	1.47	1.93
3.4(−2)	3.50	3.62	6.2(−2)	2.90	3.09	9.0(−2)	2.53	2.77	2.8(−1)	1.44	1.90
3.5(−2)	3.47	3.59	6.3(−2)	2.88	3.07	9.1(−2)	2.52	2.76	2.9(−1)	1.40	1.88
3.6(−2)	3.44	3.57	6.4(−2)	2.87	3.06	9.2(−2)	2.51	2.75	3.0(−1)	1.37	1.85
3.7(−2)	3.41	3.54	6.5(−2)	2.85	3.04	9.3(−2)	2.50	2.74	3.1(−1)	1.34	1.83

APPENDIX 8.1 (continued) Values of K_o (r/L) and $e^{r/L} K_o(r/L)$ as Function of r/L

r/L	K_o(r/L)	$e^{r/L} K_o$(r/L)	r/L	K_o(r/L)	$e^{r/L} K_o$(r/L)	r/L	K_o(r/L)	$e^{r/l} K_o$(r/L)	r/L	K_o(r/L)	$e^{r/L} K_o$(r/L)
3.2(−1)	1.31	1.81	6.0(−1)	7.78(−1)	1.42	8.8(−1)	5.01(−1)	1.21	2.6	5.54(−2)	7.46(−1)
3.3(−1)	1.29	1.79	6.1(−1)	7.65(−1)	1.41	8.9(−1)	4.94(−1)	1.20	2.7	4.93(−2)	7.33(−1)
3.4(−1)	1.26	1.77	6.2(−1)	7.52(−1)	1.40	9.0(−1)	4.87(−1)	1.20	2.8	4.38(−2)	7.21(−1)
3.5(−1)	1.23	1.75	6.3(−1)	7.40(−1)	1.39	9.1(−1)	4.80(−1)	1.19	2.9	3.90(−2)	7.09(−1)
3.6(−1)	1.21	1.73	6.4(−1)	7.28(−1)	1.38	9.2(−1)	4.73(−1)	1.19	3.0	3.47(−2)	6.98(−1)
3.7(−1)	1.18	1.71	6.5(−1)	7.16(−1)	1.37	9.3(−1)	4.66(−1)	1.18	3.1	3.10(−2)	6.87(−1)
3.8(−1)	1.16	1.70	6.6(−1)	7.04(−1)	1.36	9.4(−1)	4.59(−1)	1.18	3.2	2.76(−2)	6.77(−1)
3.9(−1)	1.14	1.68	6.7(−1)	6.93(−1)	1.35	9.5(−1)	4.52(−1)	1.17	3.3	2.46(−2)	6.67(−1)
4.0(−1)	1.11	1.66	6.8(−1)	6.82(−1)	1.35	9.6(−1)	4.46(−1)	1.16	3.4	2.20(−2)	6.58(−1)
4.1(−1)	1.09	1.65	6.9(−1)	6.71(−1)	1.34	9.7(−1)	4.40(−1)	1.16	3.5	1.96(−2)	6.49(−1)
4.2(−1)	1.07	1.63	7.0(−1)	6.61(−1)	1.33	9.8(−1)	4.33(−1)	1.15	3.6	1.75(−2)	6.40(−1)
4.3(−1)	1.05	1.62	7.1(−1)	6.50(−1)	1.32	9.9(−1)	4.27(−1)	1.15	3.7	1.56(−2)	6.32(−1)
4.4(−1)	1.03	1.60	7.2(−1)	6.40(−1)	1.31	1.0	4.21(−1)	1.14	3.8	1.40(−2)	6.24(−1)
4.5(−1)	1.01	1.59	7.3(−1)	6.30(−1)	1.31	1.1	3.66(−1)	1.10	3.9	1.25(−2)	6.17(−1)
4.6(−1)	9.94(−1)	1.57	7.4(−1)	6.20(−1)	1.30	1.2	3.19(−1)	1.06	4.0	1.12(−2)	6.09(−1)
4.7(−1)	9.76(−1)	1.56	7.5(−1)	6.11(−1)	1.29	1.3	2.78(−1)	1.02	4.1	9.98(−3)	6.02(−1)
4.8(−1)	9.58(−1)	1.55	7.6(−1)	6.01(−1)	1.29	1.4	2.44(−1)	9.88(−1)	4.2	8.93(−3)	5.95(−1)
4.9(−1)	9.41(−1)	1.54	7.7(−1)	5.92(−1)	1.28	1.5	2.14(−1)	9.58(−1)	4.3	7.99(−3)	5.89(−1)
5.0(−1)	9.24(−1)	1.52	7.8(−1)	5.83(−1)	1.27	1.6	1.88(−1)	9.31(−1)	4.4	7.15(−3)	5.82(−1)
5.1(−1)	9.08(−1)	1.51	7.9(−1)	5.74(−1)	1.26	1.7	1.65(−1)	9.06(−1)	4.5	6.40(−3)	5.76(−1)
5.2(−1)	8.92(−1)	1.50	8.0(−1)	5.65(−1)	1.26	1.8	1.46(−1)	8.83(−1)	4.6	5.73(−3)	5.70(−1)
5.3(−1)	8.77(−1)	1.49	8.1(−1)	5.57(−1)	1.25	1.9	1.29(−1)	8.61(−1)	4.7	5.13(−3)	5.64(−1)
5.4(−1)	8.61(−1)	1.48	8.2(−1)	5.48(−1)	1.25	2.0	1.14(−1)	8.42(−1)	4.8	4.60(−3)	5.59(−1)
5.5(−1)	8.47(−1)	1.47	8.3(−1)	5.40(−1)	1.24	2.1	1.01(−1)	8.25(−1)	4.9	4.12(−3)	5.53(−1)
5.6(−1)	8.32(−1)	1.46	8.4(−1)	5.32(−1)	1.23	2.2	8.93(−2)	8.06(−1)	5.0	3.69(−3)	5.48(−1)
5.7(−1)	8.18(−1)	1.45	8.5(−1)	5.24(−1)	1.23	2.3	7.91(−2)	7.89(−1)			
5.8(−1)	8.04(−1)	1.44	8.6(−1)	5.16(−1)	1.22	2.4	7.02(−2)	7.74(−1)			
5.9(−1)	7.91(−1)	1.43	8.7(−1)	5.09(−1)	1.21	2.5	6.23(−2)	7.60(−1)			

Note: 3.7(−2) means 3.7×10^{-2} or 0.037

Example: r/L = 5.0(−1) = 5×10^{-1} = 0.5; K_o(0.5) = 0.924; $e^{0.5} K_o$(0.5) = 1.52

9
Well Design and Construction

J. Boonstra

International Institute for Land Reclamation and Improvement The Netherlands

9.1   Introduction

Wells are drilled either for exploration or exploitation. With exploration wells, the objective is to collect information on the geology of the underlying aquifer. Apart from describing the various geological layers encountered (well logs), the physical properties are commonly assessed by conducting aquifer tests. These aquifer tests are commonly preceded by step-drawdown tests to determine the proper discharge rate. Proper in this respect means that good measurable drawdowns are produced in the piezometers, and that the drawdown in the pumped well does not reach the top of the well screen during the pumping period.

For exploitation or production wells, the objective can be summarized as pumping the required amount of water at the lowest cost, considering investment, operation, and maintenance. Exploitation wells are drilled for water supply for municipal, industrial, and irrigation purposes, and for water table control for drainage purposes In the latter two cases, the exploitation wells are usually placed in so-called well fields; well fields are also needed when the required pumping capacity for municipal and industrial purposes cannot be met by a single exploitation well.

In the design of exploitation wells, two main aspects can be distinguished: the maximum expected drawdown in the well and the production capacity of the underlying aquifer system. The maximum expected drawdown determines the length of the pump casing section and is mainly composed of aquifer losses and well losses. To these aquifer losses, interference effects and partially penetrating effects should be added, when applicable. The production capacity of the underlying aquifer system determines the

minimum length of the screen section which mainly depends on maximum screen velocities and screen type. To this end, well-field equations and step-drawdown tests are presented to relate aquifer losses to discharge rate and recharge rate of the aquifer, and to determine well losses and production capacity of exploration wells which will be used later as exploitation wells. Step-drawdown tests are also required to analyze well performance of exploitation wells with time for maintenance and rehabilitation purposes.

9.2 Well-Field Equations

Chapter 8 described the flow to single wells pumping in extensive aquifers. It is usually assumed that the aquifer is not recharged during the relatively short pumping period of aquifer tests. For exploitation wells, this assumption is not realistic because they are operated on a more or less continuous basis over a long time span. Most aquifer systems receive continuous or intermittent recharge from percolating rainfall and/or losses from surface water systems. When recharge is intermittent because of seasonal effects, the aquifer may perform without recharge for periods of several months or even longer. There is clear evidence that after such a period water tables will completely recover once the aquifer is recharged again by rainfall.

In this section, we assume that the aquifer is replenished at a constant rate, R, expressed as a volume per unit surface per unit of time (m^3/m^2 day). The well-flow equations that will be presented are based on a steady-state situation. The flow is said to be in a steady state as soon as the recharge and the discharge balance each other. In such a situation, beyond a certain distance from the well, there will be no drawdown induced by pumping. This distance is called the radius of influence of the well, r_e. When the wells in a well field are placed in a triangular pattern, their individual radii of influence hardly overlap (Figure 9.1). The simplifying assumption is then made that the discharge and the drawdown of each well will not be affected by those of neighboring wells. In other words, the theory of a single well can be used.
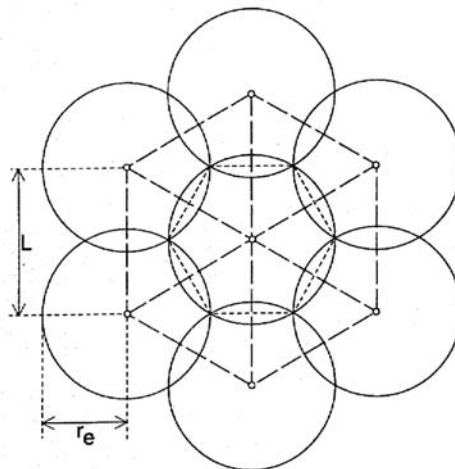
FIGURE 9.1 Layout of a well field where the wells are located in a pattern of equilateral triangles (well spacing $L = r_e \sqrt{3}$ o).

In a well field, there is a direct relationship between the discharge rate of the well, the recharge rate of the aquifer by percolation, and the area affected by pumping. The decline of the water level due to pumping is determined by the discharge rate of the well and the permeability and thickness of the aquifer. The discharge rate and the drawdown in the well are important factors in calculating the pumping costs of well fields. In an unconfined aquifer, the steady-state flow through an arbitrary qlinder at a distance r from the well is given by

$$Q_r = \pi\left(r_e^2 - r^2\right)R \tag{1}$$

where Q_r is the volumetric flow rate in m³/day at a distance r from the well, r_e is the radius of influence of the well in m, and R is the recharge rate of the aquifer per unit surface area in m/day. According to Darcy's law, Q_r equals the product of the cylindrical area of flow and the flow velocity. Hence, the discharge at distance r from the well can also be expressed by

$$Q_r = 2\pi r h K \frac{\partial h}{\partial r} \tag{2}$$

where h is the hydraulic head in m, K is the hydraulic conductivity of the aquifer in m/day, and $\partial h / \partial r$ is the is the hydraulic gradient in the aquifer at distance r. Since, in steady state, the discharge of the well, Q, equals the vertical recharge of the area within the radius of influence, the following relationship can be used

$$Q = \pi r_e^2 R \tag{3}$$

Combining Equations (1) and (3) yields

$$Q_r = Q - \pi r^2 R \tag{4}$$

or, combining Equations (2) and (4) and separating r and h

$$\left(\frac{Q}{r} - \pi r R\right)\partial r = 2\pi K h \partial h \tag{5}$$

Integration between the limits $r = r_w$, $h = h_w$, and $r = r_e$, $h = h_w$ yields

$$Q \ln\left(\frac{r_e}{r_w}\right) - \frac{1}{2}\pi R\left(r_e^2 - r_w^2\right) = \pi K\left(h_e^2 - h_w^2\right) \tag{6}$$

The quantity $1/2\pi R r_w^2$ is very small in comparison with $1/2\pi R r_e^2$ and can be neglected. If, moreover, the drawdown in the well is small in comparison with the original hydraulic head, the right-hand side of Equation (6) can be expressed as (Peterson et al., 1952)

$$\pi K\left(h_e + h_w\right)\left(h_e - h_w\right) \approx \pi K 2D\left(h_e - h_w\right) = 2\pi K D \Delta h_r \tag{7}$$

where D is the saturated thickness of the aquifer before pumping in m, and Δh_r is the drawdown due to radial flow toward the pumped well in m. Since, according to Equation (3),

$$r_e^2 = \frac{Q}{\pi R} \tag{8}$$

Equation (6) can be written as

$$\Delta h_r = \frac{Q}{2\pi K D}\left[2.3\log\left(\frac{r_e}{r_w}\right) - \frac{1}{2}\right] \tag{9}$$

If $r_e/r_w > 100$, and if we accept an error of 10%, the term -1/2 can be neglected and Equation (9) reduces to

$$\Delta h_r = \frac{2.3Q}{2\pi T} \log \frac{r_e}{r_w} \tag{10}$$

where $T = KD$ is the transmissivity of the aquifer in m²/day. Equation (10) can be used to calculate the drawdown in a single well or, when no interference between the wells is assumed, the drawdown in a well field. From Figure 9.1, it can be seen that the distance between the wells (L) should then be equal to $r_e \sqrt{3}$.

Figure 9.2 shows a leaky aquifer whose overlying layer, the aquitard, is replenished by percolating rain or excess irrigation water at a constant rate R. Depending on the recharge rate and the hydraulic resistance of the aquitard, a difference in head between the free water table in the aquitard and the piezometric level of the aquifer will develop. Under steady-state conditions, the same recharge rate will replenish the underlying aquifer. So, Equation (10) can also be used to calculate the head loss in a well or well field for leaky aquifers.
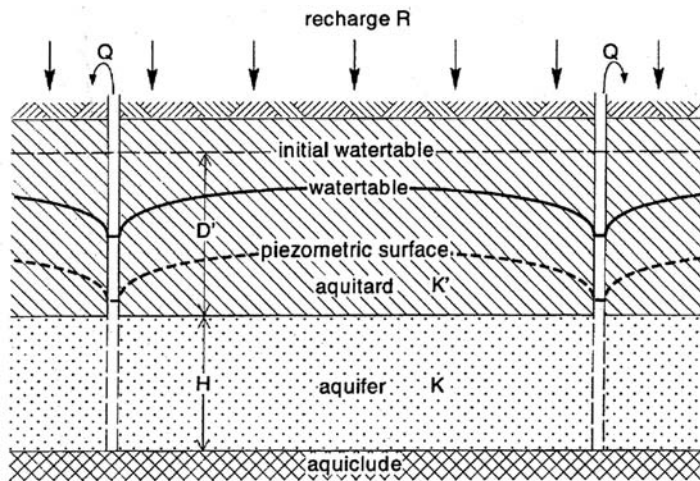
FIGURE 9.2 Schematic cross section of wells located in a leaky aquffer.

The application of Equation (10) will now be illustrated with an example calculation. In Chapter 8, data were presented for a single-well test. The analysis of this test yielded a transmissivity value of some 1400 m²/day. Suppose that from a water-balance analysis, the long-term average net recharge to the aquifer was estimated as 550 mm per year (= 1.5 mm/day). To calculate the drawdown in a well with a constant discharge of 200 m³/hr, the radius of influence needs to be calculated. Using Equation (3), the radius of influence will be

$$r_e = \sqrt{\frac{200 \times 24}{3.14 \times 0.0015}} = 1010 \text{ m}$$

Substituting this value into Equation (10) with an outer well screen radius (r_w) of 0.2 m gives

$$\Delta h_r = \frac{2.3 \times 200 \times 24}{2 \times 3.14 \times 1400} \log \frac{1010}{0.2} = 4.65 \text{ m}$$

So, the steady-state drawdown in a single well due to radial flow amounts to 4.65 m. A similar drawdown can be expected when wells are placed in a well field with a triangular pattern. The distance between the wells in Figure 9.1 should then be $L= 1010 \times 1.73 = 1747$ m; it is then assumed that there will be no mutual interference between the various wells. If the wells are located at shorter distances or in a rectangular pattern, the drawdown in the wells will increase due to well interference; this phenomenon will be numerically illustrated in Section 9.5.

In well-field design for drainage purposes, there is a trade-off between the capacity and the number of wells required, as is illustrated in Table 9.1. If larger pumps are installed, fewer pumps will be required, which generally results in lower investment costs. On the other hand, larger capacity pumps may result in considerably higher drawdowns and thus higher energy costs. Determining pump capacities on a purely economic basis could lead to very high pumping rates. There are, however, several practical constraints to these high pump capacities for drainage purposes. A well with a very high pump capacity may serve a very large area that exceeds the spacing determined by other factors. If such a well were to be out of order for a prolonged period, the neighboring wells would be overburdened, and proper drainage of the area would be impossible. Moreover, if the water is also used for irrigation, pump capacities are often limited by the requirements of the farmers.

TABLE 9.1 Comparison Between Pumping Capacities, Drawdown Due to Aquifer Losses, and Distances Between Wells (Triangular Pattern) for a Net Recharge of 1.5 mm/day

Pumping Capacity (m³/hr)	Radius of Influence (m)	Aquifer Losses (m)	Distance Between Wells(m)	Area per Well (ha)
100	714	2.23	1236	160
200	1010	4.65	1747	320
300	1236	7.14	2141	480

9.2.1 Partial Penetration

Equation (10) was derived under the assumption that the well fully penetrates the pumped aquifer. Some aquifers are so thick, however, that installing a fully penetrating well would not be just)fied. In these cases, the aquifer has to be pumped by a partially penetrating well. This partial penetration causes the flow velocity in the immediate vicinity of the well to be higher than it would otherwise be, leading to an extra loss of head. According to Hantush (1964), the effect of partial penetration in an unconfined aquifer is similar to that in a confined aquifer, because the drawdown is usually small in relation to the saturated thickness of the aquifer. To calculate the effect of partial penetration, the following equation can be used

$$\Delta h_r = \frac{Q}{4\pi T} F \tag{11}$$

and

$$F = 2\frac{D}{p}\left[\left(1-\frac{p}{D}\right)\ln\left(\frac{2p}{r_w}\right) - \frac{p}{D}\ln\frac{2D}{p} - 0.423\frac{p}{D} + \ln\frac{2D+p}{2D-p}\right] \tag{12}$$

where Δh_p is the additional head loss due to partial penetration effects in m, Q is the well discharge in m³/day, T is the transmissivity of the aquifer in m²/day, D is the thickness of the aquifer in m, and I is penetration depth of the well into the aquifer in m, assuming screening over the full depth of the well. So, when the well only partially penetrates the aquifer, the additional head loss calculated from Equation (11) should be added to the drawdown calculated by Equation (10).

9.3 Aquifer and Well Losses

The drawdown in a pumped well consists of two components: the aquifer losses and the well losses (Figure 9.3). Aquifer losses are the head losses that occur in the aquifer where the flow is laminar. They are time dependent and vary linearly with the well discharge. The drawdown s_1 corresponding to this linear aquifer loss can be expressed as

$$s_1 = B_{1(r_w,t)}Q \tag{13}$$
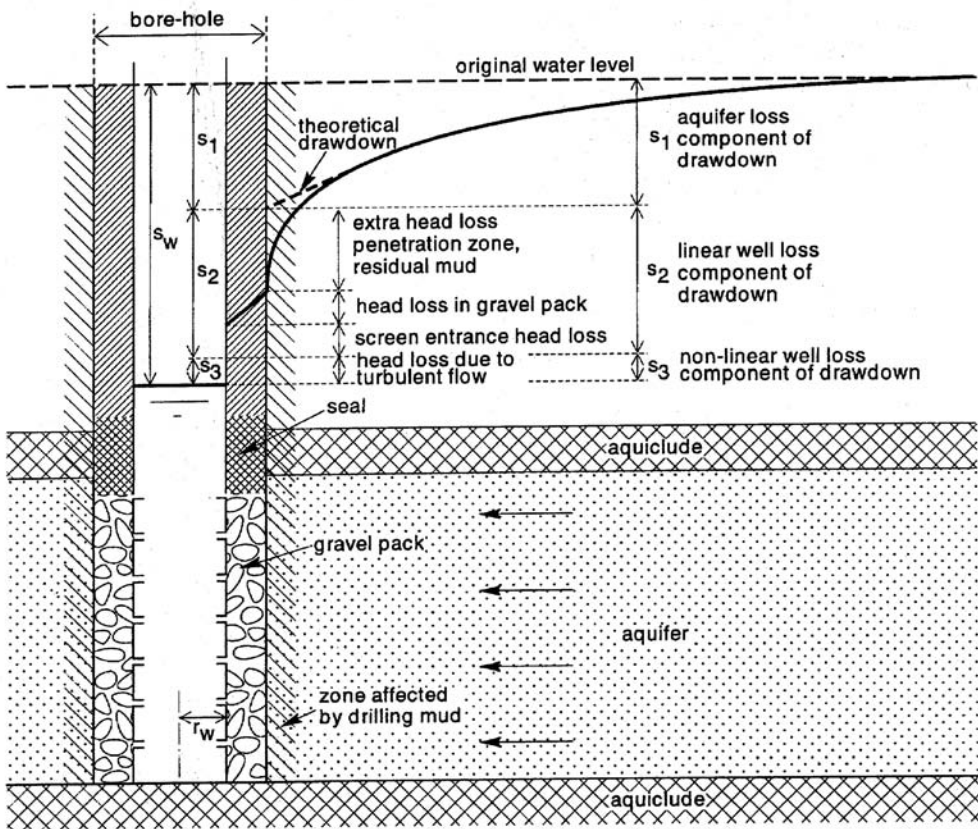
FIGURE 9.3 Various components of head losses in a pumped well.

where B_1 is the linear aquifer loss coeffficient in day/m^2. This coeffficient can be calculated from the wellflow equations presented in Chapter 8. For confined aquifers for example, it can be expressed using Equations (1) and (2) in Chapter 8 as

$$B_{1(r_w,t)} = \frac{W(u)}{4\pi T}$$

where $u = (r_w^2 S)/(4Tt)$. From the results of aquifer-test analyses, the values for transmissivity T and storativity S can be used to calculate B_1 values as function of r_w and t.

Well losses are divided into linear and nonlinear head losses. Linear well losses are caused by damaging the aquifer during drilling and completion of the well. They comprise, for example, head losses due to

the compaction of the aquifer material during drilling; head losses due to plugging of the aquifer with drilling mud, which reduces the permeability near the bore hole; head losses in the gravel pack; and head losses in the screen. The drawdown s_2 corresponding to this linear well loss can be expressed as

$$s_2 = B_2 Q \tag{14}$$

where B_2 is the linear well loss coefficient in day/m^2.

Among the nonlinear well losses are the friction losses that occur inside the well screen and in the suction pipe where the flow is turbulent, and head losses that occur in the zone adjacent to the well where the flow is usually also turbulent. All these losses responsible for the drawdown inside the well are much greater than one would expect on theoretical grounds. The drawdown s_3 corresponding to this nonlinear well loss can be expressed as

$$s_3 = CQ^P \tag{15}$$

where C is the nonlinear well loss coefficient in day^P/m^{3P-1}, and P is an exponent. The general equation describing the drawdown in a pumped well as function of aquifer/well losses and discharge rate thus reads

$$s_w = (B_1 + B_2)Q + CQ^P = BQ + CQ^P \tag{16}$$

where $s_w = s_1 + s_2 + s_3$. Jacob (1947) used a constant value of 2 for the exponent P. According to Lennox (1966), the value of P can vary between 1.5 and 3.5; from personal experience, its value may be even higher in fractured rock aquifers. The value of $P = 2$ as proposed by Jacob is, however, still widely accepted. Values of the three parameters B, C, and P in Equation (16) can be found from the analysis of so-called step-drawdown tests.

9.3.1 Well Efficiency

The relationship between drawdown and discharge can be expressed as the *specific capacity* of a well, Q/s_w, which describes the productivity of both the aquifer and the well. The specific capacity is not a constant but decreases as pumping continues and also decreases with increasing Q. The well efficiency, E_w, is defined as the ratio of the aquifer head loss to the total head losses; it reads when expressed as a percentage

$$E_w = \left\{ \frac{B_1 Q}{BQ + CQ^P} \right\} \times 100\% \tag{17}$$

The well efficiency according to Equation (17) can be assessed when both the results of a step-drawdown and those of an aquifer test are available. The former are needed for the values of B, C, and P and the latter for the value of B_1.

In practice, only the results of a step-drawdown test are usually available. The substitution of the B, C, and P values into Equation (17) would overestimate the well efficiency, because $B > B_1$. For these cases, Driscoll (1986) introduced another parameter, L_p, being the ratio of the laminar head loss to the total head losses; it reads when expressed as a percentage

$$L_p = \left\{ \frac{BQ}{BQ + CQ^P} \right\} \times 100\% \tag{18}$$

It should be noted, that Equation (18) is sometimes erroneously used to calculate the well efficiency.

9.4   Step-Drawdown Tests

A step-drawdown test is a single-well test in which the well is pumped at a low constant-discharge rate until the drawdown within the well stabilizes. The pumping rate is then increased to a higher constant-discharge rate and the well is pumped until the drawdown stabilizes once more (Figure 9.4). This process is repeated through at least three steps, which should all be of equal duration, say, a few hours each.
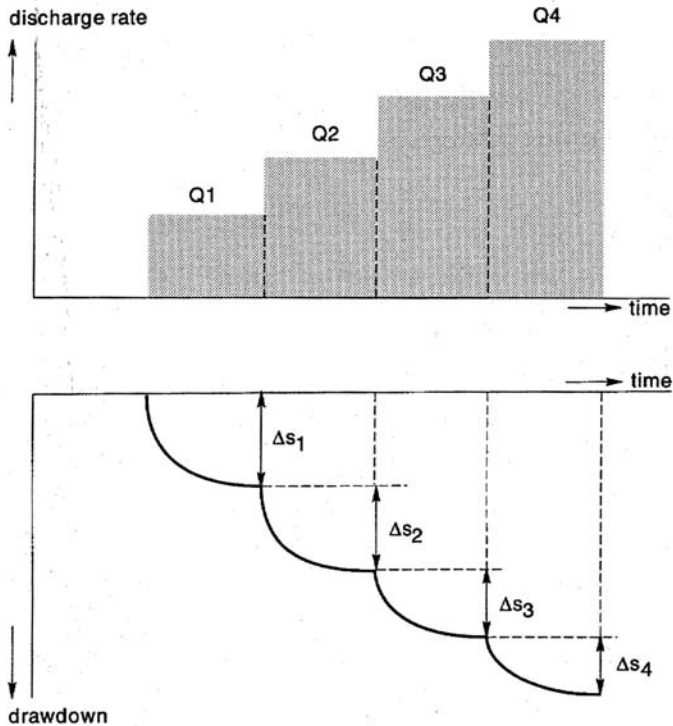
FIGURE 9.4 Principles of a step-drawdown test.

In step-drawdown analyses, use is made of so-called diagnostic plots. Values of s_w/Q versus Q are therefore plotted on arithmetic paper, where s_w represents the drawdown at the end of each step. Various configurations of diagnostic plots are then possible:

- The points fall on a horizontal line. This implies that $s_w/Q = B$. Equation (16) reduces to

$$s_w = BQ \tag{19}$$

    Hence, there are no nonlinear well losses. This situation is only encountered with very low pumping rates. The well will act differently if the pumping rates are increased.

- The points fall on a straight line under a slope. This means that $s_w/Q = B + C\,Q$. Equation (16) then reduces to

$$s_w = BQ + CQ^2 \tag{20}$$

    Equation (20) is known as the Jacob's equation. Based on this equation, Jacob (1947) developed an analysis method to calculate the values of B and C.

- The points fall on a curved line, i.e., $P \neq 2$ in Equation (16). When a concave curve can be drawn through the points, it implies that $P > 2$ and for a convex curve that $P < 2$. For these cases, Rorabaugh (1953) developed an analysis method to calculate the values of B, C, and P. Both analysis methods may be applied to confined, unconfined, and leaky aquifers.

9.4.1 Jacob's Method

The values of B and C can be found directly from the diagnostic plot of s_w/Q versus Q itself; it will yield a straight line whose slope is equal to C; the value of B can be found by extending the straight line until it intercepts the $Q = 0$ axis.

This procedure will now be illustrated with field data. A step-drawdown test was made in an unconfined aquifer. The well was pumped with a step-wise incmased discharge rate with a duration of each step of 60 minutes. Table 9.2 shows the observed drawdowns as a function of time. Figure 9.5 shows the timedrawdown plot on semi-log paper; from this figure it can be observed that the drawdowns were not yet stabilized at the end of each step. This implies that the observed drawdown values at the end of each step need to be corrected. This can be done using the following procedure as developed by Hantush-Bierschenk (Hantush, 1964):

- Extrapolate the curve through the plotted data of each step to the end of the next step
- Determine the increments of drawdowns $\Delta S_{w(i)}$ for each step by taking the difference between the observed drawdown at a fixed interval Δt, taken from the beginning of each step, and the corresponding drawdown on the extrapolated curve of the preceding step
- Determine the values of $s_{w(n)}$ corresponding to the discharge Q_n from $s_{w(n)} = \Delta s_{w(1)} + \Delta s_{w(2)} + \ldots + \Delta s_{w(9n)}$.

TABLE 9.2 Time-Drawdown Values Observed in the Pumped Well of a Step-Drawdown Test

Time in each step (min)	Step 1 Q:2765m³/day(m)	Step 2 Q:3364m³/day(m)	Step 3 Q:3853m³/day(m)	Step 4 Q:4306m³/day(m)
2	2.476	3.614	4.541	5.301
3	2.601	3.611	4.491	5.304
4	2.657	3.629	4.511	5.299
5	2.636	3.650	4.527	5.317
6	2.669	3.641	4.517	5.339
7	2.624	3.630	4.546	5.371
8	2.677	3.683	4.509	5.391
9	2.684	3.709	4.554	5.377
10	2.691	3.714	4.539	5.300
12	2.699	3.691	4.553	5.319
14	2.704	3.710	4.559	5.289
20	2.769	3.721	4.574	5.369
30	2.814	3.762	4.613	5.404
40	2.841	3.786	4.609	5.404
50	2.850	3.794	4.637	5.440
60	2.893	3.822	4.639	5.464

From Figure 9.5 the drawdown differences for each step were determined using a time interval $\Delta t = 60$ min. Based on these values, the specific drawdown values s_w/Q at the end of each step were calculated (see Table 9.3). Figure 9.6 shows the resulting diagnostic plot of s_w/Q versus Q, the data plot clearly exhibits a straight line. This implies that $P = 2$ (Equation [20]), so the Jacob's method can be applied. A best-fitting straight line was drawn through the points; its slope was calculated as $0.16 \times 10^{-3}/1276 = 1.25 \times 10^{-7}$ day²/m⁵. The interception point of the straight line with the $Q = 0$ axis can be calculated as follows: $1.133 \times 10^{-3} - 3500 \times 1.25 \times 10^{-7} = 7.0 \times 10^{-4}$ day/m². Hence, the drawdown equation of this well can be written as
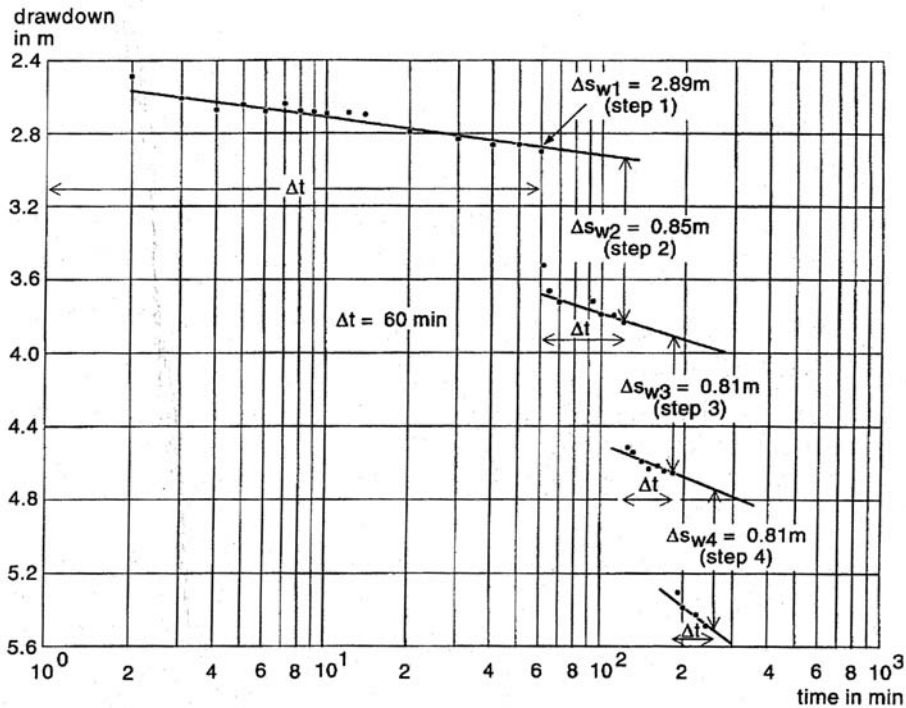
FIGURE 9.5 Time-drawdown plot of field data of a step-drawdown test in an unconfined aquifer.

TABLE 9.3 Specific Drawdown Determined from Semi-Log Plot

Step	Δs_w(m)	s_w(m)	Q(m³/day)	s_w/Q(day/m²)
1	2.89	2.89	2765	1.05×10^{-3}
2	0.85	3.74	3364	1.11×10^{-3}
3	0.81	4.55	3853	1.18×10^{-3}
4	0.81	5.36	4306	1.24×10^{-3}

$$s_w = 7.0 \times 10^{-4} Q + 1.3 \times 10^{-7} \times Q \tag{21}$$

This step-drawdown test was followed by a single-well test. For that test, the discharge rate of the third step was adopted, being 3853 m³/day. According to Equation (15), the nonlinear well loss s_3 can then be calculated as

$$s_3 = CQ^P = 1.3 \times 10^{-7} \times 3853^2 = 1.93 \text{ m}$$

This nonlinear well loss was subtracted from the drawdowns observed in the pumped well (see Chapter 8 under the heading Single-Well Tests). The analysis of these corrected drawdowns resulted in a transmissivity of 1410 m²/day and a specific yield of 0.025 for the tested aquifer. Based on these values, the theoretical drawdown according to the Theis equation (Equation [1] in Chapter 8) at $t = 60$ min can be calculated as 2.06 m. The linear aquifer loss coefficient B_1 is then calculated according to Equation (13)

$$B_1 = \frac{s_1}{Q} = \frac{2.06}{3853} = 5.35 \times 10^{-4} \text{ day/m}^2$$
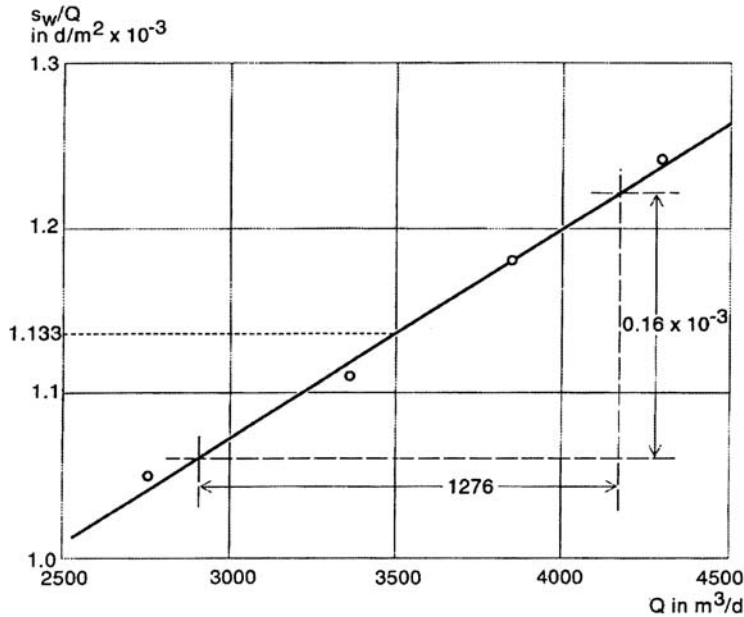
FIGURE 9.6 Diagnostic plot of s_w/Q versus Q of field data of a step-drawdown test with tbe Jacob's analysis method.

and the linear well loss coefficient B_2 is

$$B_2 = B - B_1 = 7.0 \times 10^{-4} - 5.35 \times 10^{-4} = 1.65 \times 10^{-4} \ \text{day/m}^2$$

The well efficienq for dhe discharge rate of 3853 m^3/day can then be calculated as 45% according to Equation (17). If Equation (18) had been used as representative of the well efficiency, it would have resulted in a value of 58%, i.e., an overestimation of the true well efficiency.

9.4.2   Rorabaugh's Method

The values of B, C, and P cannot be found direcHy from the diagnostic plot of s_w/Q versus Q itself, when $P \neq 2$. Equation (16) then reads

$$\frac{s_w}{Q} = B + CQ^{P-1} \tag{22}$$

Rearranging Equation (22) and taking the logarithms, Equation (22) can also be written as

$$\log\left[\frac{s_w}{Q} - B\right] = \log C + (P-1)\log Q \tag{23}$$

Equation (23) implies that a plot of (s_w/Q - B) versus Q on log-log paper would yield a straight line under a slope. The slope of this straight line is equal to P - 1, while the value of C can be found by extending the straight line until it intercepts the Q =1 axis.

This procedure will now be illustrated with field data from a step-drawdown test. Table 9.4 shows the drawdowns observed at the end of each step (Sheehan, 1971). Figure 9.7 shows the resulting diagnostic plot of s_w/Q versus Q, this plot clearly shows that the points do not lie on a straight line. This implies

that $P \neq 2$, so the Jacob's method cannot be applied. For different values of B (see Table 9.5), the values of $(s_w/Q - B)$ versus Q were plotted on log-log paper. Figure 9.8 shows that for a B value of 1×10^{-3} d/m^2 the plotted points almost exhibit a straight line. Through these points a best-fitting straight line was drawn; its slope was calculated as

$$\frac{\Delta[s_w/Q - B_3]}{\Delta Q} = \frac{\log(0.2 \times 10^{-3}/3.4 \times 10^{-3})}{\log(2180/9811)} = 1.88$$

TABLE 9.4 Step-Drawdown Data

Step	s_w(m)	Q(m^3/day)	s_w/Q(day/m^2)
1	2.62	2180	1.2×10^{-3}
2	6.10	3815	1.6×10^{-3}
3	17.22	6540	2.6×10^{-3}
4	42.98	9811	4.4×10^{-3}

Data from Sheahan, N. 1 i971. Type-curve solution of step drawdown test. Ground Water, 9, 25-29.
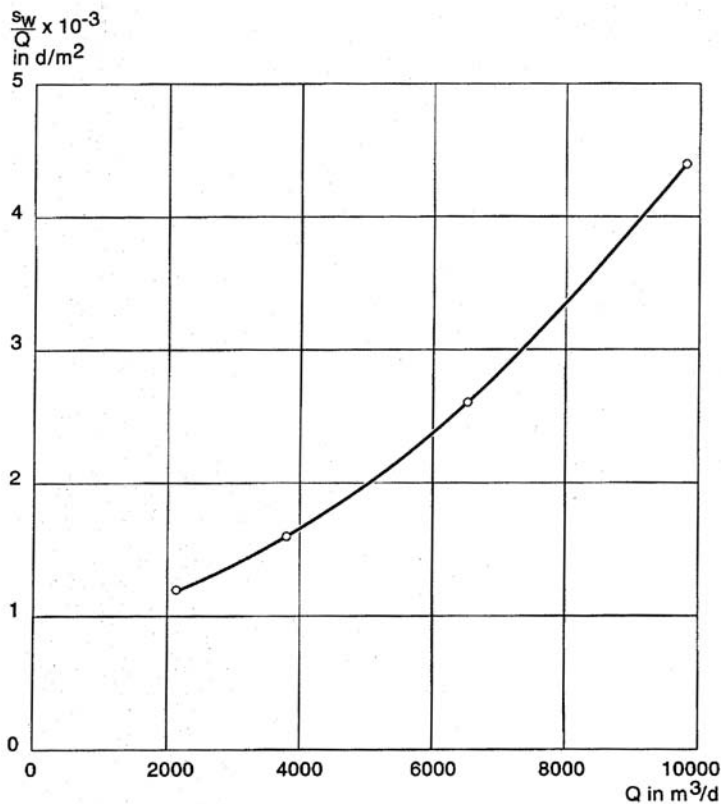
FIGURE 9.7 Diagnostic plot of s_w/Q versus Q of field data of a step-drawdown test.-

Because the slope of this line equals $(P - 1)$, it follows that $P = 2.88$. The interception point of the straight line with the $Q = 1$ axis can be calculated as follows. Figure 9.8 shows that the value of $(s_w/Q - B)$ for $Q = 10^4$ m^3/d is 3.55×10^{-3} day/m^2. Hence, the intersection of the line with the $Q=1$ axis is four log cyccles to the left. This corresponds with $4 \times 1.88 = 7.52$ log cycles belowthe point $(s_w/Q - B) = 3.55 \times 10^{-3}$ day/m^2.

TABLE 9.5 Values of $[s_w/Q - B_i]$ as Function of B_i as Used with Rorabaugh's Method

Step	$B_1=1.1 \times 10^{-3}$(day/m²)	$B_2=1.1 \times 10^{-3}$(day/m²)	$B_3=1.1 \times 10^{-3}$(day/m²)	$B_4=1.1 \times 10^{-3}$(day/m²)
1	1.2×10^{-3}	0.4×10^{-3}	0.2×10^{-3}	0.1×10^{-3}
2	1.6×10^{-3}	0.8×10^{-3}	0.6×10^{-3}	0.5×10^{-3}
3	2.6×10^{-3}	1.8×10^{-3}	1.6×10^{-3}	1.5×10^{-3}
4	4.4×10^{-3}	3.6×10^{-3}	3.4×10^{-3}	3.3×10^{-3}
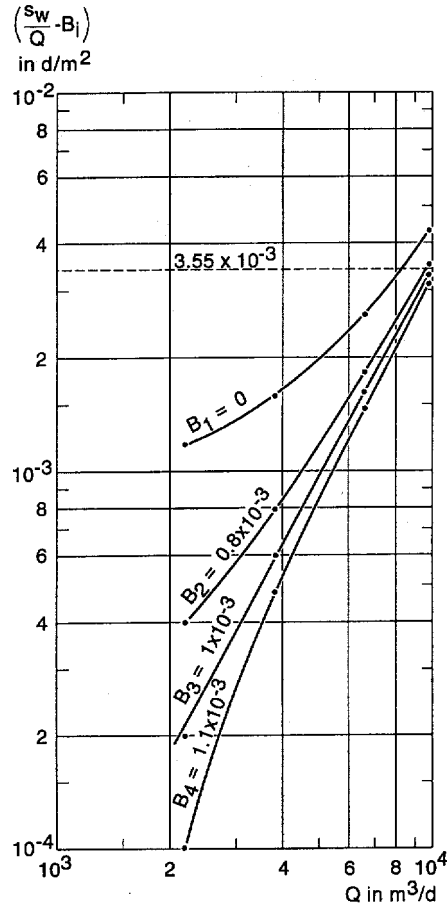
FIGURE 9.8 Log-logplot of $(s_w/Q - B_i)$ versus Q to determine the parameters B, C, and P according to the Rorabaugh's method.

The intersection point $(s_w/Q - B)$ is then calculated as follows: log $(s_w/Q - B)$ = log (3.55×10^{-3}) - 7.52 = -2.45 - 7.52 = -9.97, thus $(s_w/Q - B) = 1.1 \times 10^{-10} = C$. Hence, the drawdown equation of this well can be written as

$$s_w = 1.0 \times 10^{-3} Q + 1.1 \times 10^{-10} Q^{2.88}$$

Based on the above relationship, the optimum discharge rate can be determined, when this well would be converted to an exploitation well.

9.5   Well Design

Exploitation wells are drilled either for water supply for municipal, industrial, and irrigation purposes or for water table control for drainage purposes. The principal objectives of a properly designed well are: (1) pumping water at the lowest cost; (2) pumping water that is free of sand and silt; (3) minimum operation and maintenance costs; and (4) a long and economical lifetime. For water supply wells, good quality water with proper protection from contamination is an important additional objective.

Hydrogeologic information required for a proper design includes: stratigraphic information concerning the aquifer and overlying sediments, aquifer-test analyses of the physical properties of the aquifer, water-balance analyses of the sustainable yield, grain-size analyses of unconsolidated aquifer materials, and groundwater quality.

Well design is the process of specifying the physical materials and dimensions for a well. A good well design depends on many factors, some of which are discussed below. More detailed information on technical well design can be found in reference books such as those by Driscoll (1986) and Huisman (1975). Figure 9.9 shows the various well sections of a typical well design. The purpose and design of these well sections, and their position in the well, will be discussed below.

9.5.1   Casing Section

The pump housing is the upper section of blind casing that supports the well against collapse, and in which the pump is installed. The length of the pump housing should be chosen so that the pump remains below the water level in the well, for the selected discharge rate, under all conditions, and over the total lifetime of the well. A pump housing is always required when submersible pumps are used. No special pump housing is required in the case of a shallow water table with little drawdown where suction pumps can be used; both pump and engine are then installed at the surface beside the well. The diameter of the pump housing should be large enough to accommodate the pump with enough clearance for installation and efficient operation. It is recommended that the pump housing be two pipe sizes larger than the nominal diameter of the pump; the diameter of the pump depends on the selected discharge rate and the pump type.

The production casing is the lower section of blind pipe between the bottom of the pump housing and the top of the aquifer. The production casing is not required in unconfined aquifers at shallow depth where the pump housing reaches aufficiently deep into the top section of the aquifer. The length of the production casing depends on the thickness of the aquitard overlying the pumped aquifer. To minimize the head losses in the production casing itself, the upward velocity of the pumped water should be less than 1.5 m/s. Based on this criterion, Table 9.6 shows casing sizes recommended for various pumpin`n rates; for the pipe sizes and pumping rates shown in this table, the head losses will be small. Moreover, the diameter of the production casing should be smaller than the diameter of the pump housing and should be larger or equal to the diameter of the underlying screen section.

9.5.2   Calculation Example for the Length of Pump Housing

In this calculation example, the results of the calculations presented under the heading *Well-Field Equations* and the results of the step-drawdown test presented under the heading *Jacob'S Method* will be combined with the results of the aquifer test presented under the heading *Single-Well Tests* in Chapter 8.

The actual length of the pump housing is primarily determined by the required depth of the pump. The location of the pump depends on the expected depth to which the water level inside the well will drop for the selected design discharge rate. The procedure to determine the maximum expected waterlevel depth inside the pumped well is as follows.

The first step is to determine the aquifer losses. The aquifer losses are the head losses due to the laminar flow of water to the well and are determined by the hydraulic conductivity and thickness of the aquifer. Equation (10) can be used to estimate these losses when the wells are located in a triangular well field.
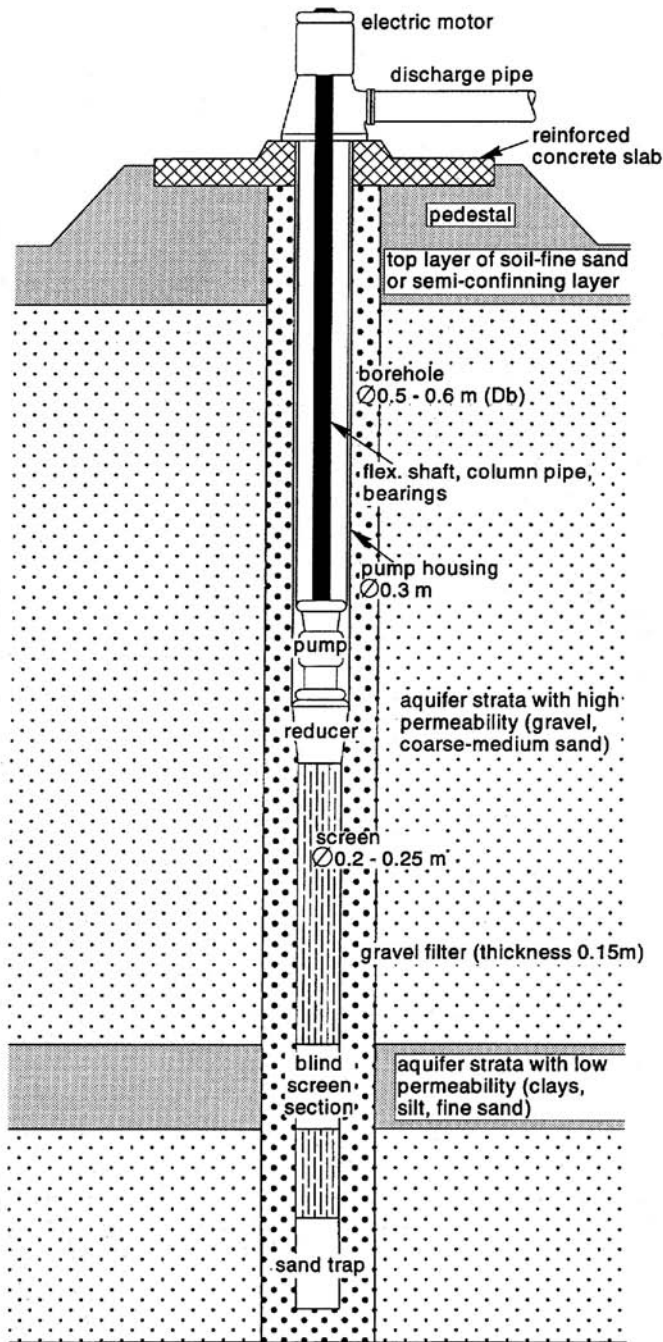
FIGURE 9.9 Typical well design.

Based on a constant recharge of 1.5 mm/day, a transmissivity of the aquifer of 1400 m²/day, and an outer radius of the well screen of 0.2 m, the resulting aquifer losses were presented in Table 9.1, assuming that there will be no interference effects of neighboring wells. When the recharge is intermittent rather than continuous because of seasonal effects, the aquifer may perform without recharge for periods of several months or even longer. The aquifer losses should then be calculated using the equations presented in

TABLE 9.6 Maximum Pumping Rates for Certain Diameters of Standard-Weight Casing, Based on an Upward Velocity of 1.5 m/s

Casing Size		Maximum Pumping Rate
(inch)	(mm)*	(m³/day)
4	102	1090
5	127	1690
6	152	2450
8	203	4250
10	254	6700
12	305	9590
14	337	11700
16	387	15500
18	438	19800
20	489	24700
25	591	36100

* Actual inside diameter.
Data from Driscoll, F. G.1986. *Groundwater and Wells*.
St. Paul, Johnson Division, Minnesota, 1089 p.

Chapter 8. Based on a specific yield of the aquifer of 0.025 together with the above transmissivity and well screen data, Equation 3 (Chapter 8) yields the following aquifer loss when it is assumed that the aquifer is not replenished during half a year

$$\Delta h_r = \frac{2.3 \times 200 \times 24}{4 \times 3.14 \times 1400} \log \frac{2.25 \times 1400 \times 182}{0.20^2 \times 0.025} = 5.50 \text{ m}$$

The effect of well interference should now be considered. For a discharge rate of 200 m³/hr, the well spacing was calculated as 1747 m (Table 9.1). The effect of well interference from one neighboring well can then be calculated as

$$\Delta h_r = \frac{2.3 \times 200 \times 24}{4 \times 3.14 \times 1400} \log \frac{2.25 \times 1400 \times 182}{1747^2 \times 0.025} = 5.55 \text{ m}$$

When the well field consists of seven wells as indicated in Figure 9.1, the aquifer loss of the center well can be calculated as $5.50 + 6 \times 0.55 = 8.8$ m. Table 9.7 shows these aquifer losses for discharge rates and distances between the wells according to Table 9.1. With partially penetrating tubewells, the additional head loss according to Equation (11) should be added to these aquifer losses.

The second step is to determine the well losses from the results of the step-drawdown test. Based on the values of C and P in Equation (21), Equation (15) gives the following nonlinear loss component

$$s_3 = CQ^P = 1.3 \times 10^{-7} \times (200 \times 24)^2 = 3.00 \text{ m}$$

Table 9.7 shows the contribution of the nonlinear well losses to the total head losses as a function of the discharge rate. Such calculations of the nonlinear well loss based on the results of a step-drawdown test can only be applied to wells with designs similar to that of the tested well. For wells with different designs, Huisman (1975) presented methods involving rather complicated calculations to estimate their expected well losses.

The third step is to assess the maximum depth to the water table; this information can be found from analyzing groundwater hydrographs showing the seasonal fluctuations of the water table throughout the years. Suppose that this maximum water table depth is 8 m below land surface. The required minimum

TABLE 9.7 Various Components of the Total Head Loss Based on Data of Table 9.1 after Half a Year without Recharge

Pumping Rate (m³/hr)	Aquifer Loss Components (m)		Nonlinear Well Loss (m)	Total Head Loss (m)
	Single Well	Well Interference		
100	2.75	2.22	0.75	5.72
200	5.50	3.30	3.00	11.80
300	8.24	3.96	6.74	18.94

length of the pump housing for the center well of the well field as depicted in Figure 9.1 would be 11.80 + 8.00 = 19.80 m for a discharge rate of 200 m³/hr. Usually an additional length of several meters is added for safety. So the total length of the pump housing can be taken as 23 m for this well.

When wells will be used for drainage purposes, the required design depth to the water table should be used instead of the maximum expected depth to the water table. In leaky aquifers, the difference in head between the free water table in the overlying aquitard and the piezometric head of the pumped aquifer should then also be added; this head loss depends on the hydraulic resistance of the aquitard and the drainable surplus. Representative values of the former can be found from aquifer-test analyses, while values of the latter can be found from water-balance analyses.

9.5.3 Screen Section

Important properties of the screen are that it prevents sand and fine material from entering the well during pumping, has a large percentage of open area to minimize the head loss and entrance velocity, supports the wall of the well against collapse, and is resistant to chemical and physical corrosion by the pumped water.

PVC and fiberglass screens are lighter and more resistant to corrosion by chemically aggressive water, but have a lower collapse strength than steel screens and casings. In practice, PVC and fiberglass-reinforced screens and casings will be technically and economically attractive for wells in alluvial aquifers, where wells are placed at moderate depths of up to 400 m. Steel screens are required in deep wells drilled in hardrock aquifers. Stainless steel screens combine both strength and resistance to corrosion and chemically aggressive water, but are more expensive.

The selection of the screen slot size depends on the type of aquifer and the use of a gravel pack. The screen slot size must be selected to ensure that most of the finer materials in the formation around the borehole are transported to the screen and removed from the well by bailing and pumping during the well-development period immediately after the borehole has been constructed and the screen and casing have been installed.

In wells without an artificial gravel pack, well development creates a zone of graded formation materials extending about 0.5 m outward from the screen. Driscoll (1986) and Huisman (1975), among others, give detailed procedures for selecting the correct slot size. They report that with good quality water and the correct slot opening, 60% of the material will pass through the screen and 40% will be retained. With corrosive water the 50%-retained size should be chosen, because even a small enlargement of the slot openings due to corrosion could cause sand to be pumped.

The screen length should be chosen so as to ensure that the actual screen entrance velocity is in accordance with the prescribed entrance velocities as listed in Table 9.8. From these screen entrance velocities, the minimum length of the well screen can be calculated from

$$Q = 86400 v_e l_{min} A_0 \tag{24}$$

where Q is the discharge rate of the well in m³/day, v_e is the screen entrance velocity in m/s, 1_{min} is the minimum screen length in m, and A_0 is the effective open area per meter screen length in m²/m. In determining the effective open area per meter screen length, it is often assumed that 50% of the actual

TABLE 9.8 Recommended Screen Entrance Velocities

Hydraulic Conductivity of Aquifer (m/day)	Screen Enterance Velocities (m/s)
>250	>0.03
250-120	0.03
120-100	0.025
100-40	0.02
40-20	0.015
<20	<0.01

Data from u.s. Environmental Protection Agency. 1975. *Manual of Water Well Construction Practices*. EPA-570/9-75-001. Office of water Supply, Washington D.C.156 p.

open area is clogged by gravel particles (Huisman, 1975). The actual open area per meter screen length depends on the type and diameter of the selected screen type. Conventional slotted screens have open areas not exceeding 10% in order not to weaken the column strength, whereas more expensive continuous slot screens of stainless steel or modern PVC screens have an open area of 30 to 50%. So the minimum total screen length is determined by the maximum screen entrance velocity and the actual screen type.

The optimum length of the screen may differ from its minimum length. Determining the optimum screen length is rather complex; it depends on (1) all the cost factors that determine the costs of pumping the required discharge or draining 1 ha; (2) the total thickness of the aquifer. In very thick aquifers, the deeper penetration of the well will result in a smaller drawdown, which reduces the pumping costs but increases the investment costs in the borehole; and (3) the selected pumping rate.

The total length of the required screen section is found by adding to the actual screen length, as outlined above, the total length of sections of blind (unperforated) pipe used to case off unproductive layers in the aquifer. The total length of blind pipe depends on the distribution of hydraulic conductivity in the aquifer (i.e., the distribution of layers of higher and lower hydraulic conductivity). This stratification can be determined from the driller's log, geophysical logs, and sieve analysis.

9.5.4 Calculation Example for the Well Screen

In this calculation example, the same data will be used as for the calculation of the length of the pump housing. The productive layers as delineated from the well logs had a total thickness of 56 m. The corresponding hydraulic conductivity of the aquifer is then 25 m/day. From Table 9.8, the screen entrance velocity should then not exceed 0.015 m/s. For a well screen with an open area of 20% and a diameter of 0.25 m, the effective open area per meter screen length can be calculated as, bearing in mind a clogging percentage of 50%

$$A_0 = 3.14 \times 0.25 \times 0.5 \times 0.20 = 0.08 \text{ m}^2/\text{day}$$

Substituting the above values into Equation (24) for a pumping capacity of 200 m^3/h = 4800 m^3/day yields the following minimum screen length

$$l_{\min} = \frac{4800}{86400 \times 0.015 \times 0.08} = 47 \text{ m}$$

Table 9.9 shows how this minimum screen length varies for different screen diameters (0.15, 0.20, and 0.25 m), and for different types of screens: cheap well screens with an open area of 10%, medium-priced well screens with an open area of 20%, and expensive, modern, continuously slotted well screens with an open area of up to 40%. The total available thickness of the productive layers was 56 m, so according to Table 9.9 only the expensive well screens for all the listed diameters and the medium-priced well screen with a diameter of 0.25 m can be used for a well with a discharge rate of 200 m^3/hr.

TABLE 9.9 Minimum Screen Lenghts for Different Types of Screens and Pump Discharges

	Open Area		Minimum Screen Lenght (m)		
Diameter (m)	(%)	(m²/m)	Q = 100 m³/hr	Q = 200m³/hr	Q = 300 m³/hr
0.15	10	0.024	77	154	231
0.15	20	0.047	39	79	118
0.15	40	0.094	20	39	59
0.20	10	0.031	60	119	179
0.20	20	0.063	29	59	88
0.20	40	0.126	15	29	44
0.25	10	0.039	47	95	142
0.25	20	0.079	23	47	70
0.25	40	0.157	12	24	35

For this discharge rate (= 4800 m³/day), the diameter of the production casing connecting the pump housing with the well screen should be 0.25 m according to Table 9.6. Because its diameter should be larger than or equal to the diameter of the underlying screen section, the above selected well screens can all be applied.

9.5.5 Gravel Pack

The effect of gravel-packed wells is to ensure that the zone around the well screen is made more permeable by removing some formation material and replacing it with specially graded material. This relatively narrow zone separates the screen from the formation material and increases the effective hydraulic diameter of the well. A gravel pack is chosen to retain most of the formation material; a well screen opening is then selected to retain about 90% of the gravel pack after development. Gravel pack material should ideally be clean, rounded, siliceous sands or gravels; carbonate material, shale particles, or soluble material such as gypsum should not exceed 5% of the total. Gravel pack material should be well sorted to assure good porosity and hydraulic conductivity of these materials around the screen. The application of a gravel pack is recommended in the following formations:

- Fine sandy alluvium and aeolian sand aquifers. In these formations, gravel packing should be considered so that larger slot openings can be used to increase the hydraulic efficiency of the well. In general, if a slot opening based on natural development is smaller than 0.25 mm, gravel packing may be more desirable because the screen's transmitting capacity may not be great enough to supply the desired yield.

- Alternating formations of fine, medium, and coarse sediment. In these formations, it is often cliffficult to determine precisely the position and thickness of each individual layer and to design a multiple-slot screen corresponding to the stratification.

- Poorly cemented sandstone aquifers. These formations may continuously lose fine material, resulting in a sand-pumping well. Another reason for gravel packing a sandstone aquifer is that the formation material usually provides little or no lateral support for the screen.

The gravel pack is designed on the basis of sieve analyses of aquifer samples. If aquifer samples from different depths show considerable variation in gradation, the gravel-pack design should be based to be stable against the finer-grade samples. Numerous investigators and agencies have experimented to develop formulae or criteria that will result in a stable gravel-pack gradation. According to Anderson (1995), the following criteria have generally been found satisfactory in actual practice:

- Aquifer material with uniformity coethcient less than 2.5: Use uniform gravel-pack material with a uniformity coefficient less than 2.5 and with the D_{50} of the gravel pack 4 to 6 times the D_{50} of the aquifer (Figure 9.10). If uniform gravel pack is not available, use a gravel pack with uniformity
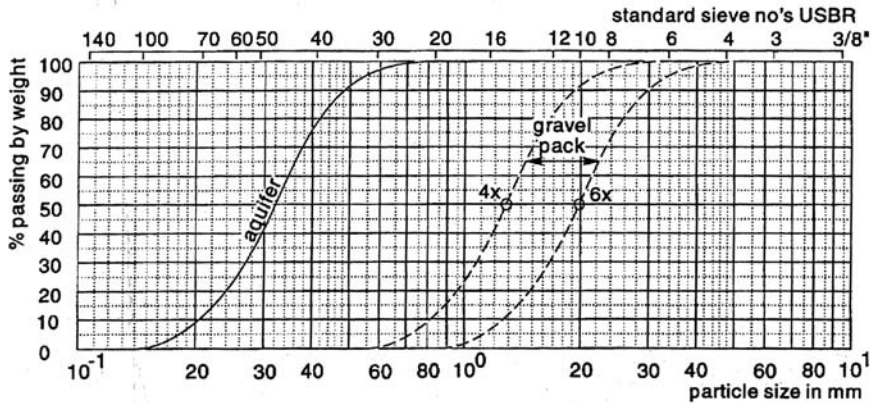
FIGURE 9.10 Aquifer and filter pack gradation.

coeffficient between 2.5 and 5 and with the D_{50} of the gravel pack not more than 9 times the Dso of the aquifer.

- Aquifer material with uniformity coeffficient between 2.5 and 5: Use uniform gravel-pack material with uniformity coeffficient less than 2.5 and with the D_{50} of the gravel pack not more than 9 times the D_{50} of the aquifer. If uniform gravel pack is not available, use a gravel pack with uniformity coefficient between 2.5 and 5 and with the D_{50} of the gravel pack not more than 12 times the D_{50} of the aquifer.
- Aquifer material with uniformity coefficient greater than 5: Multiply D_{50} of the aquifer by 6 and 9 and locate these points on the sieve analysis graph. Draw two parallel lines through these points having a uniformity coeffficient of 2.5 or less, and specify gravel-pack material that will fall between these lines.

Although smaller thicknesses already furfill the objective of a gravel pack, the thickness of a gravel pack should at least be 76 mm to ensure that a continuous layer of filter material will surround the entire screen. Under most conditions, the upper limit of gravel-pack thickness should be about 200 mm because the energy created by the development procedure must be able to penetrate the pack to repair the damage done by drilling, break down any residual drilling fluid on the borehole wall, and remove finer particles near the borehole.

9.5.6 Sand Trap

The sand trap is the section of blind pipe at the bottom of the screen section. Its function is to store sand and silt entering the well during pumping; this will occur even if the well has been properly developed. The length of the sand trap is usually of the order of a few meters (2 to 6 m). The diameter of the sand trap is usually the same as that of the screen section.

9.5.7 Pump

Centrifugal pumps are by far the most important class of pumps. Originally designed as a pump to be located at or near land surface for suction lift or booster service, it soon was adapted to installation under water in wells, first by long shaft extensions in large caissons (vertical turbine), and later in compact form as the familiar deep-well (submersible) turbine pump. A typical line-shaft pumping unit comprises a driver (power unit), discharge head, column (discharge) pipe and shaft, and one or more stages of impellers in a bowl assembly. The driver is commonly a vertical hollow-shaft electric motor, while the shaft is centered in the column pipe. A typical submersible pumping unit comprises a submersible electric motor directly coupled to the bowl assembly and a column (discharge) pipe.

The major factors determining the selection of a pump are the required discharge rate, the required head to be delivered by the pump, and the pump efficiency; the total dynamic head is made up of (1) the water-level depth inside the pumped well, as discussed earlier in this section; (2) the above-ground lift, and (3) head losses due to friction and turbulence in the discharge pipelines. The required brake horsepower at the pump shaft is defined as

$$P_s = \frac{pgQH}{\eta} \tag{25}$$

where P_s is the power to be delivered to the shaft of the pump in watts, Q is the pump discharge in m³/s, H is the total dynamic head delivered by the pump in m, ρ is the density of water in kg/m³, g is the acceleration due to gravity in m/s², and η is the dimensionless pump efficiency.

Most pump manufacturers provide performance curves for each individual impeller design and for various rotation speeds. Generally, these curves are drawn for a single stage (pumping unit). These curves typically include (1) a head-capacity curve relating the total dynamic head to the discharge rate, (2) a curve relating efficiency to discharge rate, and (3) a curve relating brake-horsepower requirement to discharge rate. From the performance curves, a particular impeller is chosen that can provide the desired yield for the total dynamic head at reasonable efficiency. The performance curve shown in Figure 9.11 indicates that a single-stage turbine pump with a 254 mm impeller operates at a peak efficiency of about 80% within a total head of 23 to 30 m and delivers from 6400 to 3700 m³/day, respectively. If the required head fluctuates within this range, the pump operates efficiently and the power required varies by 7 brake horsepower.

In selecting a turbine pump, it is usually desirable to pick an impeller type that has a steep rather than flat head-capacity curve. This is particularly true if the pumping lift will change sign)ficantly during the seasons of the year, or if the pumping lift is expected to increase over future years. With a steep curve, a moderate increase in pumping lift will not result in a large decrease in the discharge rate, whereas with a flat curve, even a slight increase in pumping lift can reduce the discharge by 50% or more. With steep head-capacity curves, the peak efficiency might be lower, but the average pumping conditions would occur at a higher efficiency.

9.5.8 Design Optimization

The design of a well is rather straightforward when the water transmitting properties of the aquifer system are limited with respect to the required discharge rate. The well will be drilled to the bottom of the aquifer and all the productive layers of the aquifer will be screened using screens with the largest possible slots and open area. When the required pumping capacity is not restricted by the water transmitting properties of the aquifer system, i.e., with highly productive, thick aquifers and/or modest pumping requirements, the design becomes more complex, especially when well fields are considered, as is usually the situation with wells for irrigation and drainage purposes. Some of the optimizing options are then:

- If larger pumps are installed, fewer pumps will be required, which generally results in lower investment costs; on the other hand, larger capacity pumps may result in considerably higher drawdowns and thus higher energy costs.
- Water can be brought to the land surface by a well with a short screen or a long screen. A short screen involves low investment costs and high energy costs because of greater drawdowns in the wells, while a long screen entails relatively higher investment costs but lower energy costs.

The optimization procedure involves examining the different well configurations that satisfy the design criteria, and, for each of these, calculating the investment costs and the annual costs of operation and maintenance. The present value of these costs is determined by applying an annual rate for discounting costs and an interest rate. The configuration that yields the lowest present value is then selected. Boehmer
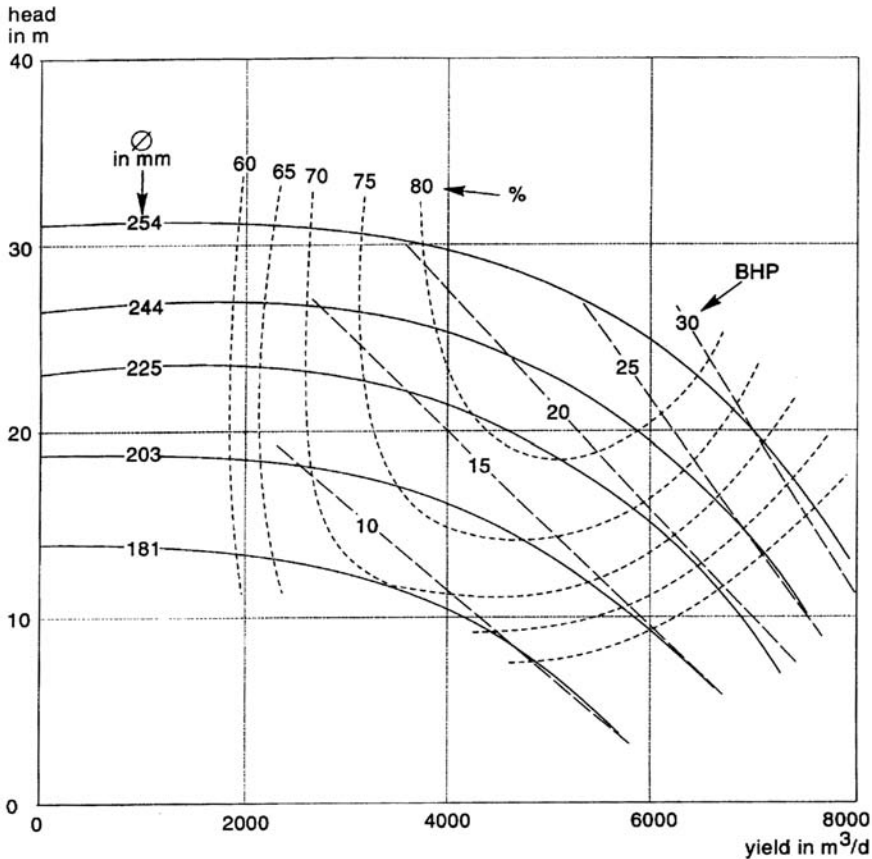
FIGURE 9.11 Representative performance curve for a vertical turbine pump operating at 1800 rpm.

and Boonstra (1994) provide a checklist of how investment costs can be calculated for a well with a prefixed pumping capacity and located in a given well-field layout.

With increasing screen lengths, investment costs rise, but drawdown and pumping costs fall. The cost per m³ of water first decreases, owing to the decreasing head losses caused by the decreasing partial penetration of the aquifer, and leads to lower pumping costs. Having reached a minimum, the price of water rises because the decreasing energy costs in the borehole no longer compensate for the higher investment costs. The calculations are repeated for different screen diameters, screen types, pump engine, and types of energy. Figure 9.12 shows the relation between the cost of water, investment and reinvestment costs, energy costs, and operation and maintenance costs.

Finally, the design with the lowest costs per m3 drainage water is selected. Obviously, the more types of screens, engines, pumps, and energy available, the better the results of the optimization procedure will be. The number of calculations required to arrive at a final result is large and complex and can best be handled by an optimum well-field-design computer program as advised by Boehmer and Boonstra (1994).

9.6   Well Construction and Maintenance

Various well drilling methods have been developed because geologic conditions range from hard rock, such as granite and dolomite, to completely unconsolidated sediments, such as alluvial sand and gravel. Particular drilling methods have become dominant in certain areas because they are the most effective
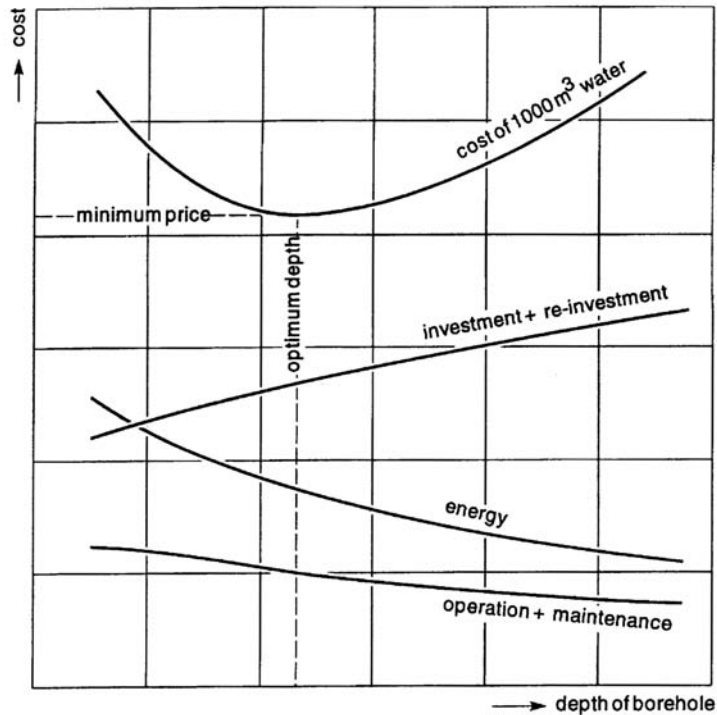
FIGURE 9.12 Costs of 1000 m³ water as a function of borehole depth.

in penetrating the local aquifers and thus offer cost advantages. In many cases, however, the drilling contractor may vary the usual drilling procedure depending on the depth and diameter of the well, type of formation to be penetrated, sanitation requirements, and principal use of the well. It is obvious, then, that no single drilling method is best for all geologic conditions and well installations. In most cases, the drilling contractor is best qualified to select the particular drilling procedure for a given set of construction parameters. Successful drilling is both an art developed from long experience and the application of good engineering practices.

Well construction usually comprises four of five distinct operations: drilling, installing the casing, placing a well screen and filter pack, if required, grouting to provide sanitary protection, and developing the well to ensure sand-free operation at maximum yield. Well drilling and installation methods are so numerous that only some of the basic principles and their applications can be described in this section. More detailed information on well construction methods can be found in reference books such as those by Driscoll (1986) and Huisman (1975).

9.6.1   Construction Methods

In the various construction methods, a distinction can be made between shallow and deep wells. Shallow wells, generally less than 15 m in depth, are constructed by boring, driving, or jetting. These methods are now briefly described and their applications are given in Table 9.10. Bored wells are constructed with hand-operated or power-driven augers. The drive shaft is built from rods; its upper part, the so-called kelly, has a square cross section to receive the necessary torque from a rotary table. At the bottom, the auger is provided with a cutting face; for emptying the auger, the drive shaft should be dismantled and reassembled afterward. A continuous-flight power auger has a spiral extending from the bottom of the hole to the surface. Cuttings are carried to the surface as on a screw conveyer, while sections may be added to the auger as depth increases. Driven wells are constructed by driving a pointed screen with

TABLE 9.10 Well Construction Methods and Applications

Method	Materials	Depth (m)	Diameter (m)	Yield (m³/day)
Bored	Unconsolidated	10-25	5-90	15-500
Driven	Silt, sand	15	3-10	15-200
Jetted	Silt, sand	15	4-8	15-150
Percussion	All	450	8-60	15-15000
Direct rotary	All	450	8-45	15-15000
Reverse rotary	Unconsolidated	60	40-120	2500-20000
Air rotary	Consolidated	250	15-45	2500-15000

Adapted from U.S. Soil Conservation Service 1969. *Engineering Field Manual for Conservation Practices*. Department of Agriculture, Washington D.C. 995 p.

attached pipe directly into the waterbearing formation. Water enters the well through a drive point at the lower end of the well. This consists of a screened cylindrical section protected during driving by a steel cone at the bottom. Driving can be done with a sledge, drop hammer, or air hammer. As driving proceeds and the well point sinks into the ground, succeeding sections of pipe are screwed on top of the screen. Jetted wells are constructed by the cutting action of a downward-directed stream of water. The high-velocity stream washes the earth away, while the casing, which is lowered into the deepening hole, conducts the water and cuttings up and out of the well. The soil material is subsequently removed in a settling basin after which the water is picked up again by the jetting pump.

Most deep, high-capacity wells are constructed by cable-tool percussion drilling or by one of the several rotary methods. These methods are now briefly described and their applications are given in Table 9.10. Cable tool drilling machines, also called percussion rigs, operate by repeatedly lifting and dropping a heavy string of drilling tools into the borehole. The drill bit breaks or crushes consolidated rock into small fragments, whereas the bit primarily loosens the material when drilling in unconsolidated formations. In both instances, the reciprocating action of the tools mixes the crushed or loosened particles with water to form a slurry at the bottom of the borehole. If little or no water is present in the penetrated formations, water is added to form a slurry. Slurry accumulation increases as drilling proceeds and eventually it reduces the impact of the tools. When the penetration rate becomes unacceptable, slurry is removed at intervals from the borehole by a sand pump or bailer. In water well drilling, the depth capability for cable rigs ranges from 100 to 1500 m.

The direct rotary drilling method was developed to increase drilling speeds and to reach greater depth in most formations. The borehole is drilled by rotating a bit, and cuttings are removed by continuous circulation of a drilling fluid as the bit penetrates the formation. The bit is attached to the lower end of a string of drill pipe which transmits the rotating action from the rig to the bit (Figure 9.13). In the direct rotary system, drilling fluid is pumped down through the drill pipe and out through the ports or jets in the bit; the fluid then flows upward in the annular space between the hole and drill pipe, carrying the cuttings in suspension to the surface. The fluid is cleaned in a settling pit before it is recirculated; drilling fluids include air, clean water, and scientifically prepared mixtures of special purpose materials.

In direct rotary drilling, the viscosity and uphole velocity of the drilling fluids are the controlling factors in removing cuttings effectively. Because of limitations in pump capacity and therefore effective cuttings removals, most direct rotary machines used to drill water wells are limited to boreholes with a maximum diameter of around 600 mm. This size may not be sufficient for high-capacity wells, especially those that are to be filter packed. Also, as hole diameters increase past 600 mm, the rate of penetration by direct rotary machines becomes less satisfactory. To overcome the limitation on hole diameter and drilling rate, reverse circulation machines were designed. The design of a reverse circulation rig is essentially the same as that of the direct rotary rig except most pieces of equipment are larger. For unconsolidated formations, the reverse circulation rig is probably the most rapid drilling equipment available. It requires a large volume of readily available water.
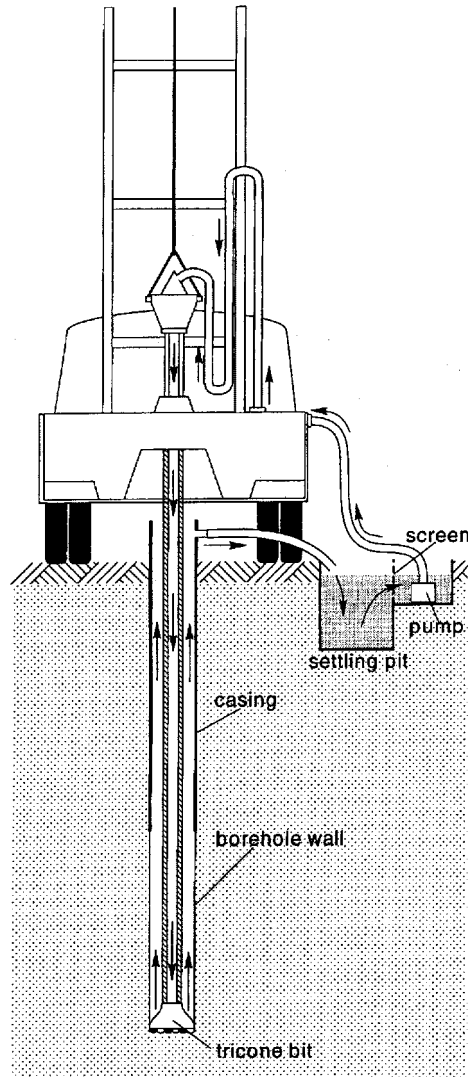
FIGURE 9.13 Schematic diagram of a direct rotary circulation system.

Rotary drillings can also be accomplished with compressed air instead of drilling mud. The technique is rapid and convenient for small-diameter holes in consolidated formations where a clay lining is unnecessary to support the wall against caving in. An important advantage of the air rotary method is its ability to drill through fissured rocks with little or no water.

Finally, it should be noted that the depth, diameter, and yield of wells as presented in Table 9.10, are only indications of the order of magnitude of what can be expected. In reality, larger diameters and depth may be obtained, whereas the yield depends primarily on the actual geology and availability of ground-water.

9.6.2 Well Development

The principal purpose of well development is to remove the fine materials adjacent to the well bore, to increase porosity and hydraulic conductivity of the aquifer and gravel pack, to remove any mud cake or compacted zone that results from the actual drilling, and to minimize or eliminate sand pumping.

Upon completion of drilling, most wells require development to reach maximum effficiency. This is particularly true of wells producing from unconsolidated aquifer materials and those in which an artificial gravel pack has been placed around the well screens. In addition, many wells may require periodic redevelopment to restore production capacity that has been lost as a result of such factors as encrustation of screens, clogging of screens by bacterial deposits, or migration of fine particles into a gravel pack. The following discussion summarizes some developments and procedures.

The simplest method of removing finer material from water-bearing formations is by overpumping, that is, pumping at a higher rate than the well will be pumped during exploitation. Overpumping, by itself, seldom produces an effficient well because most of the development action takes place in the most permeable zones closest to the top of the screen. The same applies to a certain extent to surging/back-washing. It consists of pumping a well at a high rate for a short period, shutting down the pump to allow water in the column to fall and backwash the screen, and then repeating the process until the discharge is clear. Although overpumping and backwashing techniques are widely used, and in certain situations may produce reasonable results, their overall effectiveness in high-capacity wells is rather limited when compared to other development methods.

Another method of development is to force water to flow into and out of the screen by operating a plunger up and down the casing, similar to a piston in a cylinder. The tool normally used is called a surge block. The block is usually made of discs of belting bolted between steel plates, or of wood. The outside diameter of the block should be only slightly less than the inside diameter of the screen. Before development, the well should be bailed to make sure that water will flow into it. Surging should start above the screen and move progressively downward to reduce the possibility of sandlocking the surge block. The initial motion should be relatively gentle with long strokes and at a slow rate. After surging above the screen, the hole should be cleaned and surging should start at the lower end of the screen, gradually working upward until the entire screen has been developed. The process is then repeated with a faster stroke until little or no sand can be pulled into the well. Surging procedures produce good results for screen installations in zones having good porosity and hydraulic conductivity.

Many drillers use compressed air to develop wells in consolidated and unconsolidated formations. The practice of alternatively surging and pumping with air became feasible with the great number of rotary drilling rigs equipped with large air compressors. In air surging, air is injected into the well to lift the water to the surface. As it reaches the top of the casing, the air supply is shut off, allowing the aerated water column to fall. The effectiveness is often enhanced, particularly with long screens, by using a double-packer tool that confines the pumping to a short length of the screen. Equipment can also be designed to provide a stronger backwash effect in addition to the pumping. As with surge blocks, the well should be periodically cleaned.

High-velocity jetting is a very effective way to develop sand and gravel aquifers and art)ficial gravel packs. In practice, jetting with water is almost always accompanied by simultaneous air-lift pumping so that clogging of the formation does not occur (Figure 9.14). A jetting tool is lowered on a drop pipe and water pumped down at high pressure to produce nozzle velocities of 30 m/s or more, with a minimum desirable velocity of about 45 m/s. Driscoll (1986) provides data for nozzles of several sizes at different operating pressures. The tool should have 2 to 4 nozzles, spaced equally around the circumference and directed horizontally toward the inside of the well screen. The nozzles should extend to within 25 mm of the inside of the screen. In operation, the jetting tool is lowered into the lower part of the screen, rotated slowly, and raised slowly throughout the length to be developed. The process is then repeated until development is complete.

Various acids and/or chemicals can be very effective in well development under certain conditions, both upon completion of drilling and for maintenance. Acid and chemicals can be used alone or in combination with other methods such as surge blocks or water jetting. Acids can be very effective in limestone and dolomite aquifers by enlarging fractures adjacent to the well bore and thereby sign)ficantly reducing drawdown at a given pumping rate. Acids can also be effective in removing encrusting deposits on well screens (see section on maintenance). Acids frequently used in well development and rehabili-tation are hydrochloric acid and sulfamic acids. The usual procedure is to introduce a volume of acid
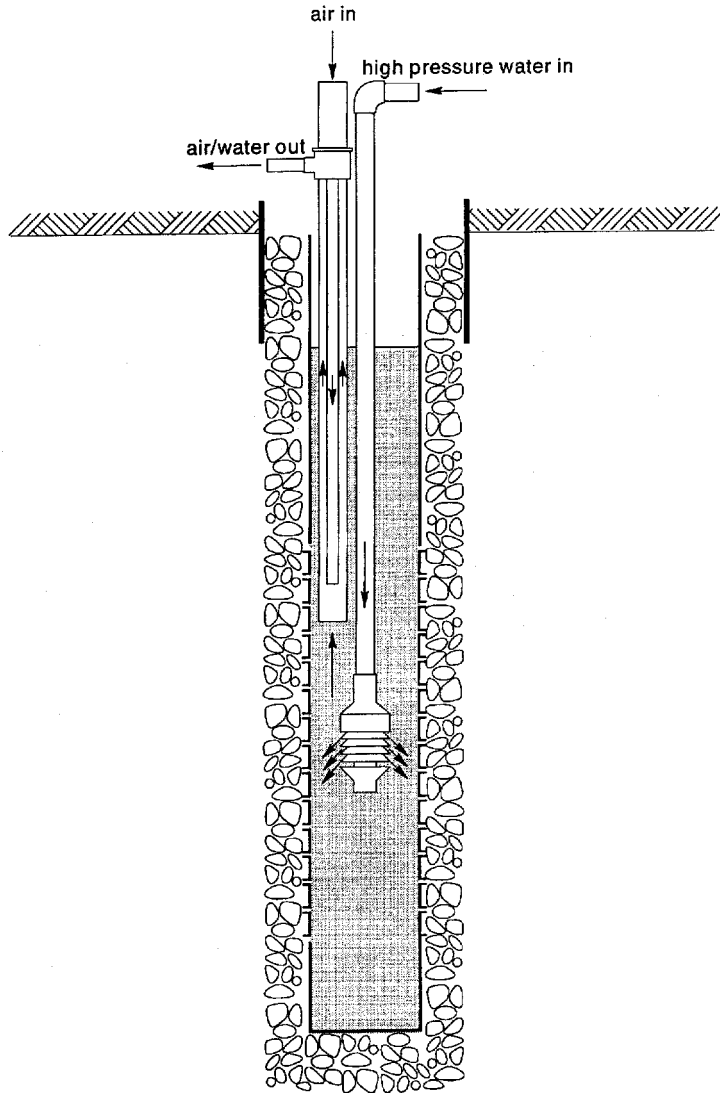
FIGURE 9.14 Schematic diagram of high-velocity jetting with air-lift pumping system.

several times that of the section of screen to be treated, followed by frequent agitation, and then removed by pumping after 6 to 8 hours. Chemicals frequently used in well development are polyphosphates and surfactants (wetting agents). The polyphosphates act to deflocculate and disperse clays and drilling muds. They are typically introduced as a solution of from 0.5 to 1.5 kg polyphosphate in 100 1iters of water in conjunction with surge-block or water jetting development. The addition of wetting agents aids in allowing the polyphosphates to penetrate clays.

9.6.3   Maintenance

The performance of a well usually declines after some years of operation, resulting in higher drawdowns and higher pumping costs. The well is in need of rehabilitation when the specific capacity of the well becomes so small that the pumping Costs increase or the discharge rate of the well can no longer be maintained. Before that time, the well needs to be rehabilitated. An effective well-maintenance program begins with good records being kept of the well's construction, including good records of the geological

conditions, the position and types of aquifers and aquicludes, water quality, and the specific capacity of the well, determined during well testing.

Every type of well requires its own maintenance program. Driscoll (1986) provides a checklist to evaluate the performance of a well. A sign)ficant decrease in specific capacity or an increase in the pumping of sand indicates that the well needs rehabilitation or restoration to its original performance. In general, rehabilitation measures are most successful when the well performance has not deteriorated too badly, or the specific capacity has not decreased too much. If the specific capacity of the well has declined by 25% of its original value, it is time to carry out a rehabilitation program. In order to determine the right moment for well rehabilitation, periodic monitoring of well performance should be done in the term of short standard tests. Complete well records can be kept at relatively minor expense, and these are indispensable in determining the causes of well failure and selecting the maintenance and rehabilitation program.

The major causes of a reduction in well performance are: (1) A reduced well yield due to chemical encrustation or clogging of the screen due to bacteriological activity; (2) plugging the formation around the well screen by fine particles of clay and sand in the pores; (3) pumping of sand due to poor well design or corrosion of the well screen; and (4) collapse of the well screen due to chemical or electrolyte corrosion of metal well screens.

Chemical and biological encrustation are major causes of well failure. Water quality and flow velocity through the screen openings determine the occurrence of encrustation. Chemical encrustation is caused by the precipitation of carbonates, mainly calcium carbonate, or iron hydroxides, which block the screen openings. Carbonate precipitation is caused by the release of carbon dioxide from the water owing to a pressure decline in the water caused by the drawdown in the well. Iron dissolved in groundwater may precipitate from the water on the well screen because oxygen is introduced into the water when the well is pumped. Another reason for the precipitation of iron may be the presence of iron bacteria in the water. Chemical encrustation can best be removed by treating the well with a strong acid solution that chemically dissolves the encrusting materials so that they can be removed from the well by pumping. Hydrochloric acid and sulfuric acid can be used. Chlorine, a strong oxidizing agent, inhibits the growth of iron bacteria. The use of hypochlorite is a relatively safe and convenient alternative to chlorine gas. The occurrence of iron bacteria in wells can be prevented by disinfecting the well and the pump immediately after installation.

Physical plugging by clay and silt particles can best be prevented by proper well development after the well screen has been installed. The removal of fine particles from the formation immediately around the screen can best be achieved by washing and brushing the screen with dispersing compounds such as sodium tripolyphosphate (STP) and other types of polyphosphates.

Sand pumping causes the abrasion of pump bowls, which leads to failure of the pump. Sand pumping results from over-sized slots in screens, over-sized gravel pack, corrosion of the well screen, inadequate development of the well, or too-high entrance velocities, causing the transport of sand from the formation toward the well. One of the above conditions, or a combination of them, results in sand from the formation entering the well. Remedying this problem may be uneconomical: it may be better to drill a new well. The best alternative, if possible, is to replace the screen or to place an inner screen inside the original well screen.

Corrosion of well screens can severely reduce the lifetime of a well. Chemical corrosion occurs especially when metal well screens are used in aggressive and saline water loaded with gases like hydrogen sulfide, carbon dioxide, and oxygen. Corrosion can be prevented by applying nonmetal screens or, when the water is not aggressive, only metal screens of stainless steel and low-carbon steel.

Finally, to pump water from a well in the most economical way, proper maintenance of pumps and engines is a prerequisite. Pump and engine manufacturers prescribe periodical maintenance of their products. Maintenance procedures depend on the pump type. They indude the adjustment and replacement of impellers, bearings, stufffing boxes, and bowl assemblies. A complete analysis of pump and engine maintenance is beyond the scope of this chapter, so readers are referred to the maintenance procedures specified by manufacturers.

For Further Information

Driscoll (1986) provides a wealth of practical information on well hydraulics, well drilling, well design, well pumps, well maintenance and rehabilitation, and groundwater monitoring that cover a wider range of conditions than presented here.

Boonstra (1989) presents a software package which facilitates the analysis of step drawdown data, including both Jacob's method and Rorabaugh's method.

References

Anderson, K. E. 1995. *Ground Water Handbook.* National Ground Water, Dublin, Ohio, 401 p.

Boehmer, W. K. and Boonstra, J. 1994. Tubewell drainage systems, in *Drainage Principles and Applications,* H.P Ritzema (Ed.). ILRI-Publication 16, Wageningen, 931-964.

Boonstra, J. 1989. SATEM: *Selected Aquifer Test Evaluation Methods: A Microcomputer Program.* ILRI Publication 48, Wageningen, 80 p.

Driscoll, F. G. 1986. *Groundwater and Wells.* St. Paul, Johnson Division, Minnesota, 1089 p.

Hantush, M. S. 1964. Hydraulics of wells, in *Advances in Hydroscience.* 1, 281-432.

Huisman, L. 1975. *Groundwater Recovery.* 2nd Ed. MacMillan, London, 336 p.

Jacob, C. E. 1947. Drawdown test to determine effective radius of artesian well. *Trans. Amer. Soc. Civil. Engrs.*, 112, Paper 2321, 1047-1064.

Lennox, D. H. 1966. Analysis of step-drawdown test. *J. Hydr. Div., Proc. of the Amer. Soc. Civil Engrs.,* 92(HY6), 25-48.

Peterson, D. F., Israelson, O. W., and Hansen, V.E.1952. *Hydraulics of Wells.* Bulletin 351. Utah Agricultural Experimental Station, 48 p.

Rorabaugh, M. J. 1953. Graphical and theoretical analysis of step-drawdown test of artesian well. *Trans. Amer. Soc. CiviL Engrs.*, 79, separate no. 362, 23 pp.

Sheahan, N. T. 1971. Type-curve solution of step-drawdown test. *Ground Water.* 9, 25-29.

U.S. Environmental Protection Agency. 1975. *Manual of Water Well Construction Practices.* EPA-570/975-001. Office of Water Supply, Washington D.C. 156 p.

U.S. Soil Conservation Service 1969. *Engineering Field Manualfor Conservation Practices.* Department of Agriculture, Washington D.C. 995 p.

Glossary

Drainable Surplus The quantity of water that flows into a groundwater reservoir in excess of the quantity that flows out under natural conditions.

Impeller Devices which accelerate the water within the pump to build pressure.

Specific Capacity The yield of a well per unit drawdown, usually expressed as cubic meters per day per meter drawdown. It is a measure for the productivity of a well.

Uniformity Coefficient Ratio of the 40% retained particle size of the sediment and the 90% retained size. The lower its value, the more uniform the grading between these limits.

10

Aquifer Characterization

Yoram Rubin,
Susan S. Hubbard,
Amy Wilson, and
Mark A. Cushey
University of California

10.1 Introduction

The goal of this chapter is to provide practical tools for quantitative analysis of complex, heterogeneous aquifers under the common situation of data scarcity. Some of the analyses and data acquisition techniques described here can be carried out as reconnaissance studies, painting a picture of the aquifer using a wide paintbrush, while others can be applied at an advanced stage, resulting in a more detailed portrait. In many cases, detailed studies are not required to characterize a site, but they may be undertaken if deemed necessary from the preliminary assessment. For example, in applications such as resource management and regional mass balance analyses, characterization of the aquifer through its geometry and effective properties may be all that is necessary; in fact, this may be all that can be realistically achieved given the severe data limitations usually encountered. However, studies of contaminant transport generally require detailed mapping of the spatial distribution of the hydraulic conductivity. When a detailed

response analysis or confidence intervals are needed, systematic data acquisition followed by stochastic analysis and simulation is required.

This chapter covers methods of hydrogeological site characterization for both gross and detailed studies. We address both site investigation techniques and methods for determining aquifer response, generally proceeding from preliminary to advanced analyses.

In Section 10.2 we discuss geophysical techniques in a hydrogeological context. These methods, which are generally applied at the preliminary stage of site investigation, may be used to identify aquifer boundaries, stratigraphy, and faults, and in some cases they can also be used to estimate hydrological properties and to delineate contaminated areas and the extent of sea water intrusion. Our discussion is confined to practical use and limitations; geophysical theory is beyond the scope of this book and is left to more advanced texts on the subject.

The remainder of the chapter provides complete reviews of the concepts and techniques involved in the more advanced stages of site characterization. In Section 10.3, we present the basic concepts of geostatistical site characterization for heterogeneous aquifers, including a large database of geological materials and their respective geostatistical spatial variability models. In Section 10.4, we present the concepts, applications, and limitations of the effective hydraulic conductivity approach, which is used to compute the average response of an aquifer once a general model of spatial variability becomes available. Section 10.5 addresses the subject of Monte Carlo analysis. The starting point is a geostatistical model of spatial variability, and the goal is to provide an in-depth uncertainty analysis which incorporates the effects of spatial variability and estimation uncertainty into a comprehensive picture of prediction uncertainty. In Section 10.6, we review some practical techniques for characterization of an aquifer's vulnerability to contamination. Our guiding principle throughout the chapter is the need for a rational characterization method that recognizes the usual situation of data scarcity.

10.2 Geophysical Techniques in Aquifer Characterization

10.2.1 Introduction

Geophysical data complement direct characterization data by providing a denser grid of subsurface measurements than is obtainable from core point measurements or from volume-averaged pump test measurements alone. Geophysical data can be collected in a noninvasive manner and can be used to reduce the number of direct measurements needed to characterize a site. Geophysical methods that are currently used for near-surface site characterization or that have potential to aid in these investigations include:

- Electrical — Electrical resistivity, electromagnetic (EM) induction, and ground-penetrating radar (GPR)
- Seismic — reflection, crosshole transmission, and refraction
- Gravitational
- Magnetic
- One-dimensional borehole — caliper, electric, nuclear, and acoustic well logs

The trade-off between aquifer coverage and resolution of heterogeneities for different characterization tools is illustrated in Figure 10.1. At one end of the spectrum, hydrologists use measurements from wells such as pump and injection tests to estimate hydraulic conductivity and connectivity of subsurface units. While these techniques offer excellent coverage, they lack resolution. At the other end of the spectrum are core sample measurements which offer high resolution but extremely low coverage; in a typical case, the core volume fraction is less than 10^{-6} of the total aquifer volume. In terms of both resolution and volume of aquifer sampled, geophysical data bridge the information gap between the more traditional site characterization extremes.
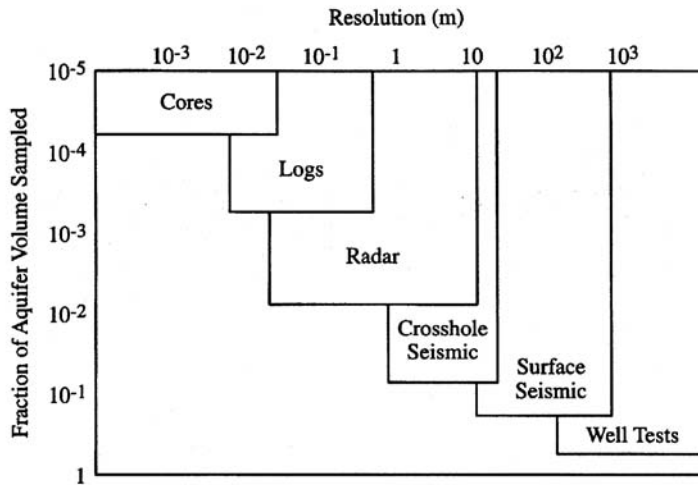
FIGURE 10.1 Comparison between resolution and volume of aquifer sampled for conventional core and well tests as well as for geophysical techniques.

Figure 10.1 also shows that the volume of aquifer sampled and the resolution of the measurement vary with each geophysical technique. Some techniques have, on average, higher resolution than others, and for each method there is a range of possible resolutions, which is primarily governed by the mechanics and acquisition parameters of the geophysical technique, such as source frequency, acquisition geometry, and sample spacing. In general, there is a trade-off between resolution and volume of aquifer sampled. Figure 10.1 also shows that in terms of both resolution and volume of aquifer sampled, geophysical data bridge the information gap between the more traditional site characterization techniques of core analysis and well tests.

Geophysical methods entail the measurement of subsurface physical properties or contrasts of these properties. From these measurements, we attempt to deduce the nature and distribution of subsurface materials. For example, subsurface variations in elastic moduli and density cause seismic waves to travel at different speeds. Information about these changes, and thus about the nature and distribution of the subsurface materials, can be deduced from analyzing the seismic arrival times. Similarly, spatial variations in electrical resistivity and dielectric properties of subsurface materials produce spatial variations in apparent resistivity measurements and ground-penetrating radar responses, respectively. The ability to deduce the nature and distribution of subsurface materials is a function of the magnitude of the physical property contrast and the spatial scale of these changes, and their relation to the resolution and sensitivity of the technique. In the geologic section shown in Figure 10.2, a saturated clay overlays a saturated sand layer, which overlays a heavily fractured bedrock surface. The contrasts in physical properties vary depending on which materials are juxtaposed, and the ability to detect these changes varies with the geophysical method employed. For example, the velocity contrast between the clay and sand layers may not be detected using surface seismic methods, but the electrical resistivity contrast is likely sufficient to be detected using electrical methods.

Geophysical data can be used to detect vertical and/or lateral changes in physical properties. The mode of data collection is governed by the mechanics of the particular geophysical technique, the acquisition geometry, and the acquisition parameters. In electrical method terminology, the term *profiling* refers to the measurement of lateral changes in electrical properties over a constant subsurface depth. These electrical property profiles can be used to infer lateral variations in subsurface materials such as the pinch-out of an aquifer, and a grid of profiles can be collected and interpreted to delineate the aquifer in three dimensions. Often we are interested in the changes in physical properties with depth, as shown in Figure 10.2. The term *sounding* in electrical methods refers to a collection of measurements which are
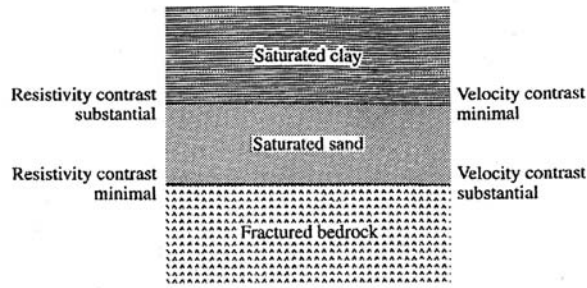
FIGURE 10.2 Schematic geological section and associated resistivity and velocity contrasts at interfaces. (From Burger, H. R. 1992. *Exploration Geophysics of the Shallow Subsurface.* Prentice Hall, Englewood Cliffs, NJ. With permission.)

associated with a single surface location and are made as a function of depth. Profiling and sounding acquisition modes are commonly used in both electrical and electromagnetic methods.

Surface seismic and ground-penetrating radar data are displayed as *wiggle-trace* profiles with distance on the horizontal axis and arrival time (which can be converted to depth) on the vertical axis. The vertical and lateral variations in arrival time, amplitude, and phase of the wiggles that comprise the vertical cross sections are indicative of subsurface physical property changes. Seismic and ground-penetrating radar profiles thus yield two-dimensional information about physical property changes.

Crosshole acquisition geometries are common to many geophysical methods. A typical seismic crosshole tomographic geometry consists of sources and detectors located in separate boreholes as shown in Figure 10.3. Direct energy from a source in one borehole is received by several geophones that are connected by a cable in the other borehole. The vertical source position is changed, and the recording is repeated until the source has occupied all positions in the source borehole. For seismic crosshole tomographic data reduction, the interwell area is discretized into a grid composed of cells or pixels, and inversion algorithms are used to transform the recorded travel time and amplitude information into estimations of velocity and sometimes attenuation at each pixel (Peterson et al., 1985). Crosshole acquisition geometry has also been used with electrical methods; electrodes occupy the wellbores when collecting crosshole electrical resistivity data, borehole induction coils occupy the wellbores when collecting crosshole EM data, and radar antennas occupy the boreholes when collecting radar crosshole data. Crosshole tomographic data can produce high-resolution images and are usually used for detailed site investigation.

Both surface geophysical and crosswell geometries can be used for *dynamic process monitoring*, where data are collected at the same location as a function of time. Observing the data as "time difference" cross sections (data set collected at an earlier time subtracted from data set collected at a later time) enhances the image of subtle geophysical attribute changes due to dynamic processes such as stream flooding, hydraulic fracturing, and the spread of contaminant plumes.

The choice of which geophysical method to use for a particular investigation is made by considering the goals of the investigation, the necessary level of resolution, conditions of the site, funds available for the investigation, and availability of other data. Because there may be several plausible subsurface scenarios that could produce a given geophysical response, the combined use of different geophysical techniques is often recommended. Additionally, relations between geophysical measurements and hydrologic or lithologic parameters of interest are often non-unique. The use of two or more different geophysical methods can reduce the ambiguity associated with these relationships as well as providing crosschecks between the different types of measurements. Geophysical data are also often collected in an iterative fashion where lower resolution techniques, acquisition parameters, and acquisition geometries are used for reconnaissance investigation, and higher resolution techniques, acquisition parameters, and acquisition geometries are used to investigate smaller areas within the reconnaissance area that require
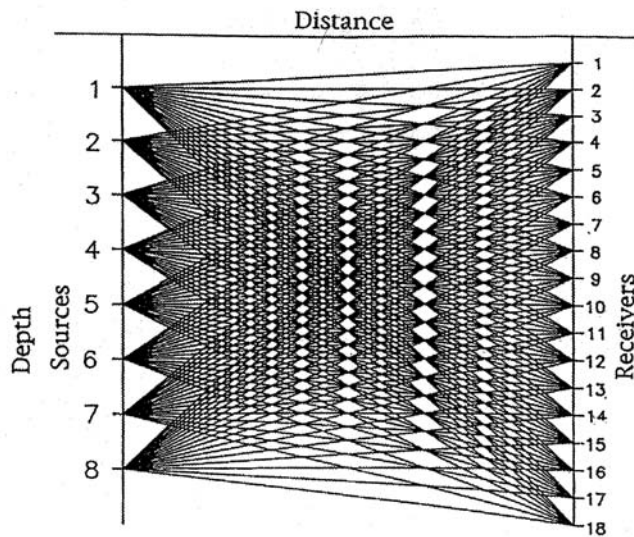
FIGURE 10.3 Crosshole tomography geometry for seismic and radar methods. Sources and receivers are located in separate boreholes, and energy from each source is received by all geophones. Crosshole acquisition geometries have also been used with electrical resistivity and EM methods.

more detailed information or that produced an anomalous response during the reconnaissance investigation.

The purpose of this section is to review the previously listed geophysical techniques, and provide a description of the underlying physical principles of each method and of each instrument, common acquisition geometries, and general data reduction and interpretation methods. Also presented are several petrophysical relationships that may permit the transfer of geophysical measurements into estimates of hydrological parameters such as porosity and water content. These relationships are a topic of current research and thus are presented primarily to illustrate the potential that geophysical methods have for aiding subsurface characterization investigations, not to provide a recipe for geophysical-to-hydrogeological data transfer. Several important components of this topic are not addressed here, including the issue of matching between the scales of the petrophysical model and the measurements. A detailed discussion of different approaches for combining geophysical and hydrological data is given in Chapter 11.

Potential applications and limitations of the different surface or crosshole techniques are listed in Tables 10.1 and 10.2, together with the frequency of use, relative cost, and relative resolution of each method. Tables 10.3 and 10.4 provide a similar summary for the borehole techniques. For detailed information about acquisition, reduction and interpretation of all geophysical methods, the reader is referred to Robinson and Coruh (1988) and Burger (1992).

10.2.2 Electrical Methods

Electrical methods involve the detection of effects produced by electric current flow in the ground. We focus in this section on electrical methods that are controlled using artificial energy sources. The two properties of primary interest are:

- The ability of a material to conduct electrical current (*electrical conductivity*), or its inverse (*electrical resistivity*)
- The *dielectric constant*, which is a measure of the separation (polarization) of opposite electrical charges within a material that has been subjected to an external electrical field.

TABLE 10.1 Surface and Crosshole Geophysical Methods

Method	Parameters Commonly Inferred from Measurements	Potential Applications	Limitations	Relative Cost[1]	Relative Resolution[2]	Frequency of Use for Ground Water Studies and Site Characterization
Electrical						
DC Resistivity	Electrical resistivity	Mapping of: gross stratigraphy, faults, depth to bedrock fresh-salt water interfaces, landfills, some contaminant plumes.	Cannot be used in paved areas, wires can be cumbersome, arrays can be long for deep investigations.	Moderate-high	Moderate	Very common
EM	Electrical conductivity	Mapping of: gross stratigraphy, salt-fresh water interfaces, depth to bedrock, faults, some contaminant plumes. Detection of buried tanks and pipes.	Difficult to resolve resistive targets, noise from power lines, fences, and pipes.	Low-moderate	Moderate	Very common
		Estimation of hydrologic properties using petrophysical relationships and tomographic or CMP acquisition geometries.	Ineffective in highly electrically conductive environments, interference from electrical noise such as power lines and fences.	Moderate	High	Very common (surface acquisition)
GPR	Dielectric constant	Mapping of: detailed stratigraphy, some contaminant plumes, cavities, depth to bedrock and water table. Detection of buried tanks and pipes.	For hydrological property estimation, more sophisticated data acquisition and reduction processes are required.			Uncommon (CMP or tomographic acquisition)
Seismic						
Reflection	P-wave reflectivity and velocity	Mapping of: gross and detailed stratigraphy, faults, water table.	Acquisition often difficult in unconsolidated environment. Sophisticated acquisition and processing system required, sensitive to cultural noise.	Moderate-high	High	Common
Crosshole	P-wave velocity and attenuation	Mapping of: detailed stratigraphy, faults, cavities, some contaminant plumes. Estimation of porosity and permeability using petrophysical relations.	Sophisticated acquisition and processing necessary, sensitive to cultural noise. Best in saturated sections.	High	High	Uncommon
Refraction	P-wave velocity	Mapping of: gross stratigraphy and velocity structure, depth to bedrock and water table, significant faults.	Gross feature identification only, cannot resolve layers that have lower velocities than overlying layers, sensitive to cultural noises.	Low	Low	Common
Gravity and Magnetics						
Gravity	Density	Mapping of: depth to bedrock, faults, landfills, and cavities.	Gross feature identification only, requires extensive data reduction and accurate elevation information.	Low-moderate	Low	Uncommon
Magnetics	Presence of magnetic materials	Mapping of: depth to magnetic basement, locating buried drums and pipes, landfill delineation.	Interference from industrial and near surface magnetic features.	Low	Low-high depending on application	Very Common

[1]Using acquisition parameters that sample similar subsurface volumes. (Costs for some methods given by Goldstein, 1994.)

[2]Resolution is a function of geophysical method, acquisition parameters, and site conditions. This column presents relative resolutions for acquisition in environments favorable for each method using parameters optimal for a near-surface investigation for that method. See text for examples.

TABLE 10.2 Applications of Surface and Crosshole Geophysical Methods

Surface/Crosshole Method:	DC Resist-ivity	EM	GPR	Seismic Reflec-tion	Cross-hole Seismic	Seismic Refrac-tion	Gravity	Magnetics
Application:								
Depth to Water Table	2	2	4	2	0	4	1	0
Fresh/Salt Water Interface	4	4	2	1	0	2	0	0
Depth to Bedrock	4	4	4	4	0	4	4	2
Gross Hydrostratigraphy	4	4	4	4	1	4	1	1
Detailed Hydrostratigraphy	2	3	4	4	4	1	1	0
Significant Fault Detection	4	4	4	4	4	4	4	4
Cavity Detection	2	1	3	2	3	0	2	1
Porosity or Permeability Estimation[1]	3	3	3	3	3	1	1	0
Water Content Estimation[1]	3	3	3	1	0	1	1	0
Contaminant Detection[1]	3	3	3	1	3	0	0	0
Detection of Buried metallic Objects	2	4	4	2	1	0	1	4
Landfill Delineation	4	4	4	2	2	2	2	4

0 = not considered applicable; 1 = limited use; 2 = used, or could be used, but not the best approach or has limitations; 3 = has potential but not fully developed; 4 = generally considered an excellent approach, techniques are well developed.
[1] These applications are current topics of research.

The ability of a material to conduct electrical current can be investigated using electrical resistivity methods which measure the electrical potential distribution produced by an electric current, or by using electromagnetic methods which detect the electromagnetic field that is produced by an alternating electrical current introduced into the ground. The dielectric constant is a measure of the electrical polarization that occurs within a material when an external electric field is applied; variations in dielectric properties affect the response of ground-penetrating radar methods. In the following paragraphs we review the methods of electrical resistivity, electromagnetic techniques, and ground-penetrating radar. Although several other electrical methods exist, these particular techniques are most commonly used for groundwater and near-surface studies.

10.2.2.1 Electrical Resistivity

For groundwater studies, DC resistivity methods have been more frequently used than any other electrical method. Resistivity is a measure of the ability of electrical current to flow through materials; it is an intrinsic property of the material and is measured in units of Ohm-meters (Ω-m). Electrical resistivity methods involve the introduction of a time-varying direct current (DC) or very low frequency (<1 Hz) current into the ground between two current electrodes. In two-electrode systems, current flows from the positive to the negative current electrode. The currents set up equipotential surfaces, and current flow lines are perpendicular to these surfaces. Figure 10.4 is a cross section in a homogeneous and isotropic formation through a field set up by a two-current electrode configuration which illustrates the relative positions of the equipotential lines and the current lines. The labels in Figure 10.4 indicate the percent of total current that penetrates to the depth of the line. As the distance between the electrodes increases, the percent of the current flowing at depth increases. For a homogeneous and isotropic formation, 70% of the current flow is above a plane at a depth equal to the electrode spacing (Burger, 1992), which in Figure 10.4 is the distance between C_1 and C_2. When an interface is present in the subsurface, the positions of the equipotential lines are altered from those in a homogeneous formation. For a two-layer system, the fraction of current that penetrates a particular depth is a function of the electrode spacing, depth of the interface, and the resistivities of materials above and below the interface (Van Nostrand and Cook, 1966).

TABLE 10.3 Borehole Geophysical Methods

Logging Tool	Physical Property Measured	Applications	Limitations, Sources of error	Frequency of Use for Groundwater Applications	Borehole Conditions	Radius of Investigation
Caliper	Borehole diameter	Borehole diameter measurement, lithology estimation, and fracture detection.		Common	Cased or uncased, saturated or unsaturated.	Borehole wall only
Electrical						
Single-Point Resistance	Electrical resistance	Lithological identification, water level detection.	Qualitative interpretation, sensitive to drilling fluid, hole rugosity, and power line noise.	Common	Uncased hole saturated with conductive fluid.	Small: 5–10 times diameter of electrode
Normal Resistivity	Electrical resistivity	Water quality, lithological identification, depth to water table, porosity estimation.	Anomalous readings in thin beds, noise from power lines.	Common	Uncased hole saturated with conductive fluid.	40–160 cm
Self or Spontaneous potential (SP)	Electrochemical, electrokinetic or redox potential	Lithological identification, water quality, depth to water table, and shale content estimation.	Sensitive to: salinity difference between formation and drilling fluid, mud cake and temperature, noise from electrical cable currents, power lines, and railroad tracks. Usually used in conjunction with resistivity logs. Qualitative interpretation.	Common	Uncased hole saturated with conductive fluid.	Variable
Dipmeter	Electrical resistivity	Primary sedimentary structure orientation, fracture detection.	Computer analysis most effective.	Uncommon	Uncased hole saturated with conductive fluid.	Borehole wall only
Nuclear						
Gamma	Natural gamma	Lithological identification, clay content, and permeability estimation.	Affected by casing, cement, mud, borehole diameter, and position of probe in well.	Common	Cased (PVC or steel) or uncased, saturated or unsaturated.	15–30 cm from borehole wall
Gamma-Spectrometry	Natural gamma radiation	Lithological identification, identification of radioisotopes in groundwater.	Affected by casing diameter, fluid, casing, annular material and instrument drift. Requires computer analysis.	Uncommon	Cased (PVC or steel) or uncased, saturated or unsaturated.	15–30 cm from borehole wall
Gamma-Gamma or Density	Electron density	Lithologic and depth to bedrock identification, density, porosity, and moisture content estimation.	Affected by borehole diameter, gravel pack, casing, mud cake, cement and position of probe in hole. Source requires license.	Common	Saturated or unsaturated, optimal in uncased.	15 cm from borehole wall

TABLE 10.3 (continued) Borehole Geophysical Methods

Logging Tool	Physical Property Measured	Applications	Limitations, Sources of error	Frequency of Use for Groundwater Applications	Borehole Conditions	Radius of Investigation
Neutron	Hydrogen content	Moisture content, saturated porosity, lithology, and depth to water table.	Affeted by borehole diameter, thickness of mud cake, casing, cement, and weight, temperature and pressure. Source requires license.	Common	Saturated or unsaturated, optimal in uncased.	15–25 cm from borehole wall
Acoustic						
Acoustic, Sonic, or Velocity	P-wave transit time	Lithology and fracture identification, depth to bedrock, estimation of grout integrity and porosity.	Spikes occur when amplitude of P-wave is less than detection level of tool.	Common	Uncased and saturated.	25 cm for sediments, 120 cm for rocks
Acoustic Waveform	P-wave transit time, amplitude and frequency	Lithology identification, fracture characterization, cement bond logging.	Complex analysis required.	Uncommon	Uncased for all situations except for cement bond logging. Saturated.	25 cm for sediments, 120 cm for rocks
Acoustic Televiewer (ATV)	Borehole wall acoustic reflectivity	High resolution identification and orientation of fractures and bedding, assessing casing condition.	Reflection from other logging tools or borehole wall can be greater than features in soft rock, sensitive to gain settings and hole deviations, expensive.	Uncommon	Uncased for all situations except for assessing casing condition. Saturated.	Borehole wall only

Adapted from Keys, S. W. 1989. *Borehole Geophysics Applied to Ground-Water Investigations*. National Water Well Association, Dublin, OH.

TABLE 10.4 Applications of the Most Common Borehole Geophysical Methods

Borehole Method:	Single Point Resistance	Resistivity	Self Potential	Gamma	Gamma-Gamma (Density)	Neutron	Acoustic (Sonic or Velocity)
Application:							
Groundwater Exploration	4	4	4	2	2	3	4
Lithology	4	4	4	4	4	4	4
Fault Displacement	0	0	1	0	0	0	2
Fracture Detection	2	0	1	0	0	0	3
Porosity Estimation	1	4	0	1	4	4	4
Moisture Content	0	0	0	0	4	4	0
Permeability Estimation	0	0	0	3	2	2	0
Water Quality	2	4	2	2	0	0	0
Density Estimation	0	0	0	0	4	3	2

0 = not considered applicable; 1 = limited use; 2 = used, or could be used, but not the best approach or has limitations; 3 = excellent potential, but not fully developed; 4 = generally considered an excellent approach, techniques are well developed.
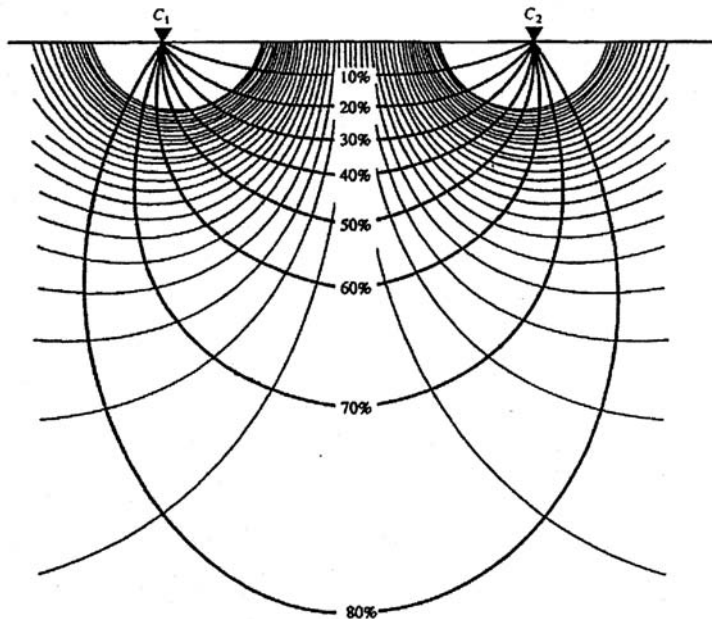
FIGURE 10.4 Cross section in a homogeneous and isotropic earth through an electric field which is set up by a two-current electrode configuration indicated by C_1 and C_2. The equipotential lines are perpendicular to current flow lines. The labels on the current flow lines indicate the percent of total current that penetrates to the depth of the line. (From Burger, H. R. 1992. *Exploration Geophysics of the Shallow Subsurface.* Prentice Hall, Englewood Cliffs, NJ. With permission.)

To obtain a value for subsurface resistivity, two potential electrodes are placed between the current electrodes shown in Figure 10.4, and the difference in potential or voltage (ΔV) is measured. This measurement, together with the known current (I) and a geometric factor (K) that is a function of the particular electrode configuration, can be used to calculate resistivity (ρ) following Ohm's law:

$$\rho = \frac{\Delta V}{I} K \tag{1}$$
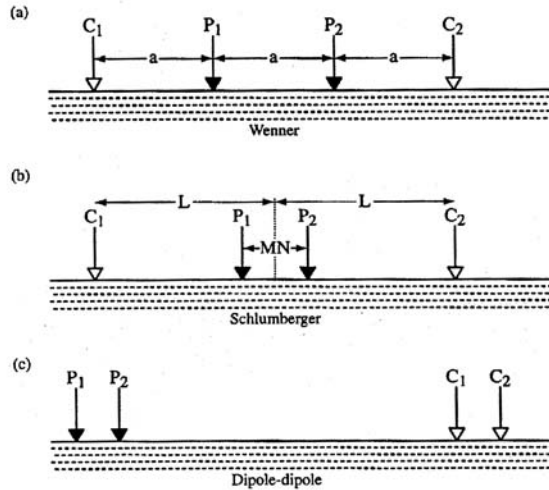
FIGURE 10.5 Common electrode configurations used to measure apparent resistivity of the subsurface. C_1 and C_2 are the current electrodes and P_1 and P_2 are the potential electrodes. (a) Wenner array. The outer two current electrodes apply a constant current, and the inner two potential electrodes measure the voltage difference created by the current. The electrodes maintain a constant spacing of "a". (b) Schlumberger array. The spacing (MN) between the potential electrodes is much smaller than the spacing (2L) between the current electrodes. (c) Dipole-dipole array. The potential electrode pair functions independently from the current electrode pair which enables a sampling in both the lateral and vertical direction of subsurface apparent resistivity. (From Burger, H. R. 1992. *Exploration Geophysics of the Shallow Subsurface.* Prentice Hall, Englewood Cliffs, NJ. With permission.)

The expression for uniform ground shown in Equation (1) is also applied when the resistivity of the ground is heterogeneous, but the more general term "apparent resistivity" (ρ_a) is used.

Four-electrode configurations are commonly used to measure the apparent resistivity of the subsurface. The simplest configuration is the Wenner array shown in Figure 10.5a, where the outer two current electrodes, labeled C_1 and C_2, apply a constant current, and the inner two potential electrodes, labeled P_1 and P_2, measure the voltage difference created by this current. The electrode spacing is a fixed value a, and the apparent resistivity of the subsurface sampled by this array can be computed using:

$$\rho_a = \frac{\Delta V}{I} 2\pi a \qquad (2)$$

Another commonly used electrode configuration is the Schlumberger array (Figure 10.5b), where the spacing (MN) between the potential electrodes (P_1, P_2) is much smaller than the spacing (2L) between the current electrodes (C_1, C_2).

In the dipole-dipole array, illustrated in Figure 10.5c, the current electrode pair and the potential electrode pair are closely spaced, and there is a significant distance between the sets. Unlike the Wenner and Schlumberger arrays, which collect data in either the *profiling* or the *sounding* mode depending on electrode array geometry, the potential electrodes and the current electrodes of the dipole-dipole array function independently of each other, permitting simultaneous profiling and sounding. The advantages and disadvantages of the different electrical resistivity array configurations are discussed by Zodhy et al. (1974).

For illustrative purposes, consider the Wenner array. Profiling is undertaken by moving the entire array laterally along the surface by a fixed distance after each reading to obtain apparent resistivity measurements as a function of distance. For each reading, the value of apparent resistivity is assigned to the geometric center of the electrode array. The interpretation of horizontal profiling data is generally

qualitative, and the primary value of the data is to locate geologic structures such as buried stream channels.

The best interpretive results are generally obtained by using a combination of profiling and sounding data (Zodhy et al., 1974). Soundings are obtained with the Wenner array by "expanding" the array along a straight line such that the spacing between the individual electrodes remains equal but takes on different values after each measurement. The depth of the measurement is a function of the electrode spacing as well as the subsurface resistivity contrasts; the effects that both the electrode spacing and the presence of an interface have on apparent resistivity are illustrated in Figure 10.6a–10.6c. The dashed lines in this figure represent current flow lines in a homogeneous environment and the solid lines represent actual current flow in the presence of a single interface that separates units with different resistivities. In Figure 10.6a, the electrodes are spaced close together and the current only traverses the upper (lower resistivity) unit. As the electrode spacing is increased (Figure 10.6b), the current penetrates to greater depths, and the apparent resistivity is higher than that measured in the situation of Figure 10.6a due to the influence of the lower (higher resistivity) layer. Also note that because the lower layer is more resistive, less current flows at the depth of the interface than if the interface were absent. Lastly, in Figure 10.6c, the electrodes are still farther apart, which permits a substantial amount of current to flow through the higher resistivity layer.

The qualitative variations in apparent resistivity with increasing electrode spacing, a, are shown in Figure 10.6d. The curve of Figure 10.6d is called a sounding curve, and it illustrates, for the geology shown in Figures 10.6a–10.6c, a general increase in resistivity with depth. The sounding curves can be quite complex when the subsurface geology is not laterally homogeneous and when there are more than two subsurface layers. The interpretation of multilayer sounding curves is not unique. To interpret electrical resistivity sounding data, various curve-fitting or computer inversion schemes are used, or the measured data are compared with model computations (Zodhy et al., 1974).

An example of the use of both profiling and sounding acquisition modes using the Wenner array is shown in Figures 10.7a and 10.7b. The map shown in Figure 10.7a is a contour map produced from the resistivity measurements of several profiles collected near San Jose, CA, using an a-spacing (see Figure 10.5a) of 6.1 m (Zodhy et al., 1974; Burger, 1992). Contours of equal apparent resistivity delineate an approximately east-west trending apparent resistivity high. The geologic cross section (BA) shown in Figure 10.7b is based on four expanding-spread traverses (soundings), the apparent resistivity profile information, and information from three boreholes whose locations are indicated on the cross section. The cross section reveals that the high-resistivity area shown on the apparent resistivity map of Figure 10.7a is a zone of gravel and boulders that defines the location of a buried stream channel.

In addition to mapping subsurface structure and stratigraphy, electrical resistivity measurements can be used to infer information about lithology and hydrogeological parameters. Electrical conduction usually takes place in fluids in connected pore spaces, along grain boundaries, and within fractures; the matrix of the material is generally considered nonconductive. Material texture, porosity, presence of clay minerals, moisture content, and the resistivity of the pore fluid all affect electrical resistivity. In general, all of the following reduce resistivity: increased water content, increased salinity of water, increased clay content, and decreased grain size (Burger, 1992). Assuming that water fills the voids, resistivity is also lowered by an increase in porosity, an increase in the number of fractures, and by increased weathering. Because of the myriad of influences on apparent electrical resistivity measurements, it is not possible to correlate resistivity values with lithology directly without other constraining information.

Electrical resistivity values for common materials are shown in Table 10.5. From this table, it is evident that there are wide ranges of electrical resistivity for some materials and also that resistivity values of different materials overlap. Values commonly vary over 12 orders of magnitude and have a maximum range of 24 orders of magnitude (Telford et al., 1990). However, the following general statements can be made regarding electrical resistivity:

- Resistivity is sensitive to moisture content; thus unsaturated sediments usually have higher resistivity values than saturated sediments.
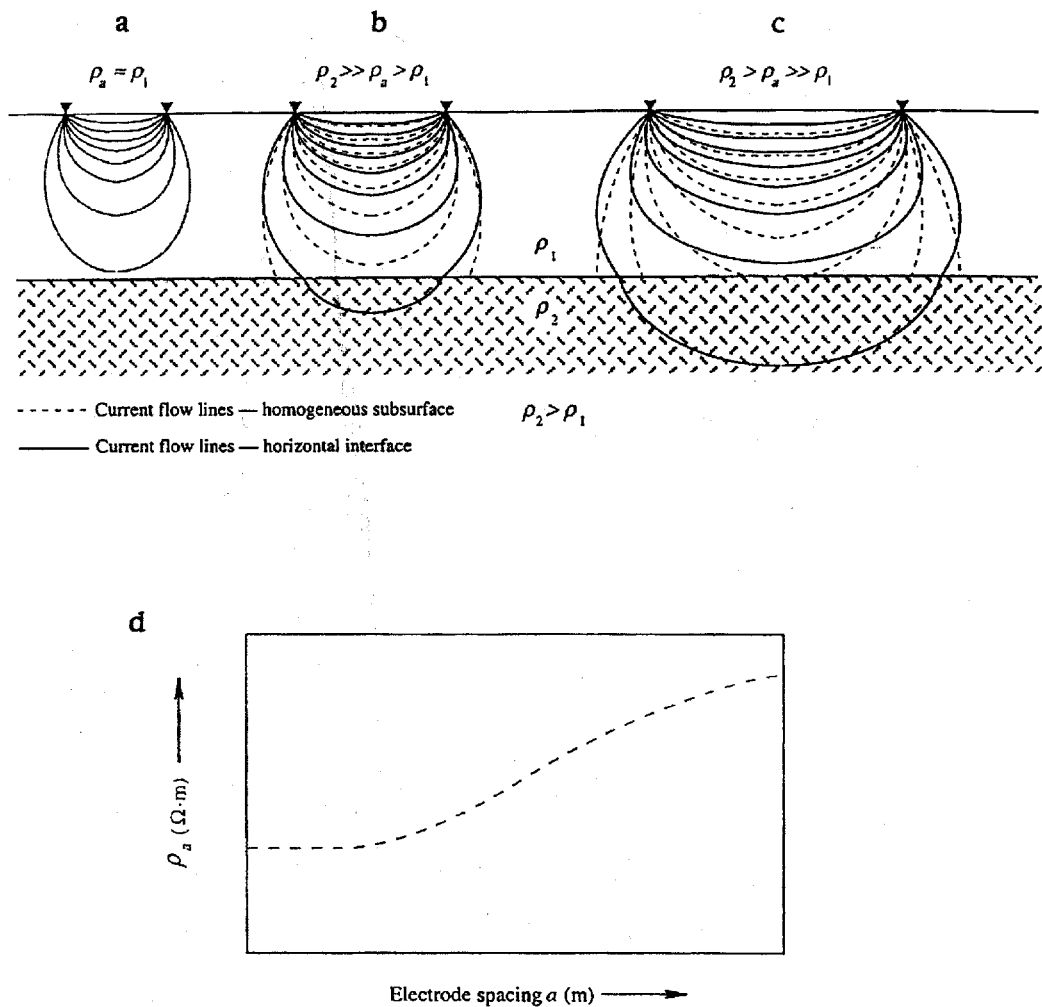
FIGURE 10.6 Effects of electrode spacing and presence of an interface on apparent resistivity measurements. The dashed lines represent current flow lines in the absence of the interface, and the solid lines represent actual current flow lines. (a–c) As the current electrode spacing is increased, the current lines penetrate deeper and the apparent resistivity measurements are influenced by the lower (more resistive) layer. (d) The qualitative variations in apparent resistivity as a function of electrode spacing are illustrated by the two-layer sounding curve. (From Burger, H. R. 1992. *Exploration Geophysics of the Shallow Subsurface*. Prentice Hall, Englewood Cliffs, NJ. With permission.)

- Sandy materials generally have higher resistivity values than clayey materials.
- Granitic bedrock generally has a higher resistivity value than saturated sediments and frequently offers a large apparent resistivity contrast when overlain by these sediments.

Empirical formulae as well as site-specific calibration curves have been used to relate apparent resistivity measurements to hydrological parameters of interest. For example, the electrical resistivity of a porous material varies with the volume and arrangement of the pores and even more with the resistivity and volume of pore fluid. An empirical formula that relates these parameters was developed in the laboratory and is given by Archie (1942):

$$\rho_r = a\phi^{-m}S^{-n}\rho_w \tag{3}$$
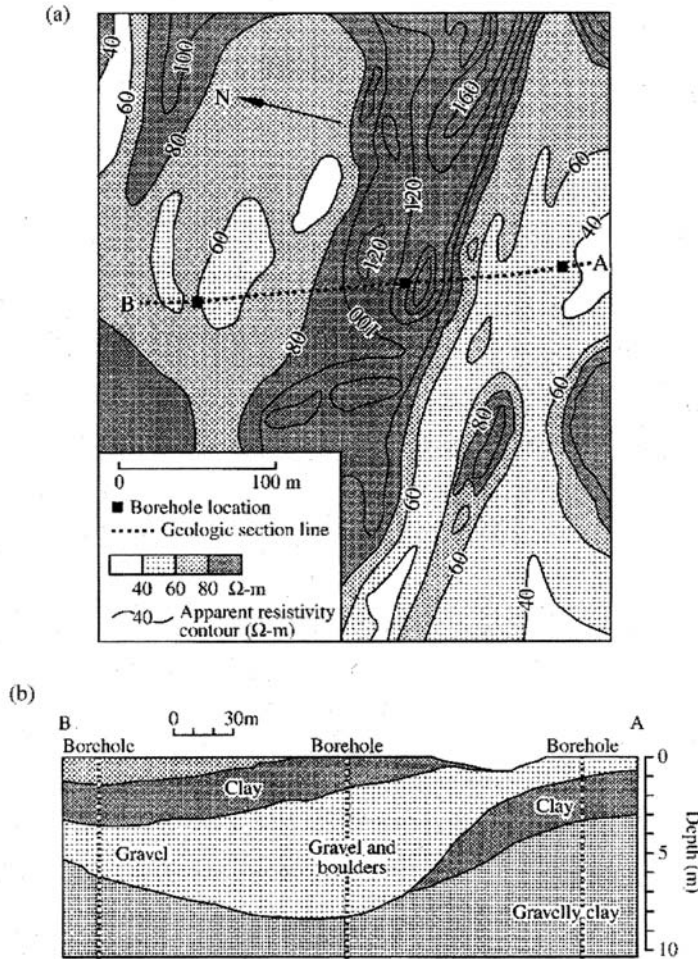
FIGURE 10.7 Resistivity survey used to delineate lateral and vertical variations in subsurface stratigraphy. (a) Contour map produced from resistivity measurements collected using many constant-spread Wenner profiles with an "a" spacing of 6.1 m. (From Zodhy A. A., Eaton, G. P., and Mayhey, D. R. 1974. Applications of surface geophysics to groundwater investigations. Techniques of Water Resources Investigations of the U.S. Geological Survey, Book 2, Chapter D1 and Burger, H. R. 1992. *Exploration Geophysics of the Shallow Subsurface.* Prentice Hall, Englewood Cliffs, NJ. With permission.) The contours of equal apparent resistivity delineate an approximately east-west trending apparent resistivity high. (b) A geologic cross section (BA) based on four electrical resistivity soundings, the apparent resistivity information from the profiles, and from three boreholes whose locations are shown above the cross section. The cross section reveals that the high-resistivity trend shown in (a) is a zone of gravel and boulders that define the location of a buried stream channel.

where ρ_r is the electrical resistivity of the rock, ρ_w is pore water resistivity, ϕ is fractional porosity, S is fractional water saturation, and *n, a,* and *m* are constants (n ≈ 2, $0.6 \leq a \leq 1.0$, and $1.4 \leq m \leq 2.2$; Ward [1990]). Although Archie's law was obtained using lithified materials, Jackson et al. (1978) showed that it is sufficiently accurate for unconsolidated materials as well. Equation (3) is generally used for well log interpretation, but if ρ_r, ρ_w, and ϕ can be independently measured, and a and m reasonably estimated, then the fractional water saturation can be found using electrical surveys (Ward, 1990). This concept was used by Pfeifer and Anderson (1995) to monitor the migration of tracer-spiked water through the subsurface using a resistivity array.

TABLE 10.5 Resistivity and Dielectric Constants for Typical Near-Surface Materials

Material	Resistivity (ohm-m)	Dielectric Constant
Sand (dry)	10^3–10^7	3–6
Sand (saturated)	10^2–10^4	20–30
Silts	10^2–10^3	5–30
Shales	10–10^3	5–15
Clays	1–10^3	5–40
Humid soil	50–100	30
Cultivated soil	200	15
Rocky soil	1000	7
Sandy soil (dry)	7100	3
Sandy soil (saturated)	150	25
Loamy soil (dry)	9100	3
Loamy soil (saturated)	500	19
Clayey soil (dry)	3700	2
Clayey soil (saturated)	20	15
Sandstone (saturated)	25	6
Limestone (dry)	10^9	7
Limestone (saturated)	40	4–8
Basalt (saturated)	100	8
Granite	10^3–10^5	4–6
Coal	10^4	4–5
Fresh water	30–10^4	81
Permafrost	10^2–10^5	4–8
Dry snow	10^5–10^6	1
Ice	10^3–10^5	4–12

Data from Davis, J. and Annan, A. 1989. Ground-penetrating radar for high-resolution mapping of soil and rock stratigraphy. *Geophysical Prospecting.* 37, 531-551; Telford, W. M., Geldart, L. P., and Sheriff, R. E. 1990. *Applied Geophysics.* Second Edition, Cambridge University Press; Ulriksen, P. F. 1982. *Application of Impulse Radar to Civil Engineering.* Ph.D. Thesis: Lund University, Sweden.

Additionally, using apparent resistivity measurements and hydrological measurements collected in the field, relationships have been developed between resistivity and hydraulic conductivity/transmissivity in porous aquifers (Kelly, 1977; Mazac et al., 1985; Huntley, 1986), resistivity and hydraulic conductivity in fractured aquifers (Mazac et al., 1990), and resistivity and saturated hydraulic conductivity in the vadose zone (Mazac et al., 1990). Most of these studies involved using linear regression techniques with site-specific field data. The use of petrophysical relations to estimate near-surface hydraulic parameters continues as a topic of research.

Significant progress has been made in the development and use of crosshole electrical resistance tomography (ERT); this technology has recently become commercially available. ERT involves using a crosshole geometry similar to that illustrated in Figure 10.3, where electrodes are placed in wellbores and at the surface, and the differences in voltage between potential electrodes are measured. ERT data acquisition is automated and capable of recording 3500 measurements per hour, requiring only one technician. This method has proven to be very successful in imaging electrical resistivity contrasts and for dynamic process monitoring. For example, Ramirez et al. (1993) used electrical resistance tomography to map the subsurface distribution of a stream flood as a function of time, and Daily and Ramirez (1995) used electrical resistance tomography to monitor an *in situ* remediation process for removal of volatile organic compounds from subsurface water and soil.

In summary, the complexities in the interpretation of sounding curves and the non-uniqueness of the solutions suggest that surface resistivity methods are best suited to simple and flat-layered subsurface geology. The field effort itself can be cumbersome because the electrodes are connected by wires that

must be relocated after readings, and the length of the spread can get quite long for deep investigations. However, the wide range of values of electrical resistivity in natural materials and the sensitivity of electrical resistivity to parameters such as moisture content have made electrical resistivity methods a useful tool in hydrological investigations. Reviews of the resistivity method are given by Ward (1990), Van Nostrand and Cook (1966), and Zodhy et al. (1974).

10.2.2.2 Electromagnetic Induction

The use of electromagnetic techniques for groundwater and environmental site assessments has increased dramatically in recent years. Controlled-source inductive electromagnetic (EM) methods use a transmitter to pass a time-varying current through a coil or dipole on the earth's surface. This alternating current produces a time-varying magnetic field which interacts with the conductive subsurface to induce time-varying eddy currents. These eddy currents give rise to a secondary EM field (Figure 10.8). Attributes of this secondary magnetic field, such as amplitude, orientation, and/or phase shift, can be measured by the receiver coil and, by comparing these attributes with those of the primary field, information about the presence of subsurface conductors, or the subsurface electrical conductivity distribution, can be inferred.
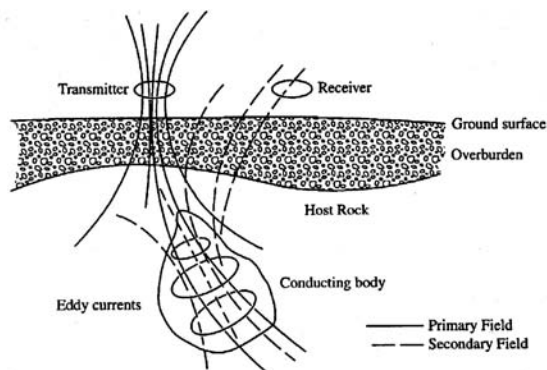
FIGURE 10.8 Electromagnetic Induction Technique. Changes in the primary magnetic field with time induce eddy currents in the subsurface which give rise to a secondary EM field. (From Best, M. E. 1992. *Geological Association of Canada Short Course Notes Volume 10.* Wolfville, Nova Scotia, May 28-29. With permission.)

Recall that electrical conductivity is the inverse of electrical resistivity; resistivities in Ohm-meters (Ω-m) can be divided into 1000 to convert to conductivity in milli-Siemans per meter (mS/m). As such, electrical conductivity measurements made using EM methods are also affected by material texture, porosity, presence of clay minerals, moisture content, and the electrical resistivity of the pore fluid. EM methods tend to require less time to acquire and achieve greater penetration depth with shorter arrays than DC resistivity methods. However, the EM equipment can be more expensive, and the interpretational methods necessary to extract qualitative information can be more complicated than those used with resistivity methods. Because a conductive subsurface environment or target is required to set up the secondary field measured with inductive EM methods (Figure 10.8), EM methods are best suited for use when attempting to detect the presence of high-conductivity subsurface targets such as salt water saturated sediments or clay layers. It is difficult to detect the presence of electrically resistive targets with EM methods.

EM instrumentation can take on many different forms. Most EM systems consist of a source that transmits a time-varying EM field, one or more receivers to measure the components of the total (primary and secondary) magnetic field, and sometimes the components of the electric field, and the necessary electronic circuitry to process, store, and display the signals (Best, 1992). As with resistivity data, EM induction data can be collected in profile or sounding mode. The mode of acquisition and the resolution and penetration of the data are governed by the subsurface conditions, the domain (frequency or time)

under which the system operates, and the coil configuration. For frequency domain systems, high transmitter frequencies permit high-resolution investigation of subsurface conductors at shallow depths while lower transmitter frequencies permit deeper observations but at a loss in resolution. Time domain systems measure the secondary magnetic field as a function of time, and early-time measurements yield information about the near-surface, while later-time measurements are increasingly influenced by the electrical properties of deeper layers. The depth of penetration and resolution are also governed by coil configuration; the measurements from larger coil separations are influenced by electrical properties at greater depths, while smaller coil spacings sample from the near-surface. Thus, to collect an EM profile, the coil spacing must be constant *and* the frequency or time of measurement fixed. To collect an EM sounding, measurements are collected at the same surface location using different frequencies, different observation times, *or* different coil spacings. All coil configurations require minimum contact with the ground, enabling noninvasive data collection as well as investigation of subsurface electrical conductivity data beneath paved areas. Although still a research topic, EM crosshole geometries have been used for dynamic process monitoring (Wilt et al., 1995).

Because electrical conductivity is the inverse of electrical resistivity, the discussion regarding the relation of electrical resistivity to lithology or hydrological parameters given in Section 10.2.2.1 can be applied in an inverse manner to electrical conductivity measurements. For example, electrical conductivity is higher for saturated sediments than for unsaturated sediments, and clayey materials generally have higher electrical conductivities than sandy materials. Examples of investigations that have used EM inductive methods to investigate hydrogeological parameters include those of Kachanoski et al. (1988, 1990) and Sheets and Hendricks (1995), who used EM induction methods to estimate soil water content, and Kachanoski et al. (1988), who used EM induction methods to investigate the spatial variations of soil texture and pore fluid. Additionally, McNeill (1990) discusses the relation between electrical conductivity and the hydrogeological parameters of porosity and saturation.

The introduction of contaminants into an aquifer may alter the electrical conductivity of the pore fluid. EM methods can be used to detect contaminants if:

- The contrast between the electrical conductivity of the contaminant and the original pore fluid is at least 10%.
- The plume or contaminated aquifer occupies a significant fraction of the zone of influence of the measurement.
- Variations in electrical conductivity due to other factors that affect electrical conductivity, such as lithology, are smaller than or separable from variations in electrical conductivity due to the contaminant (Greenhouse and Harris, 1983).

Organic contaminants generally reduce electrical conductivity and are more difficult to detect over background variations than inorganic contaminants, which typically increase the electrical conductivity, but EM has been successful at detecting both organic and inorganic groundwater contamination plumes (Buselli et al., 1990). Figure 10.9a shows an apparent conductivity contour map obtained from interpretation of frequency domain Geonics EM-34 profiling measurements over a brine evaporation pit (Hoekstra et al., 1992). Several measurements were collected, as shown by the dots in Figure 10.9a, using a single frequency and a fixed coil spacing of 20 m. An apparent electrical conductivity high trending from the brine pond to the southwest is interpreted to be the main direction of contaminant migration. Five EM soundings at the locations indicated by the numbered boxes in Figure 10.9a were collected with the time domain Geonics EM-47 and inverted to yield electrical conductivity estimations as a function of depth. The geologic cross section shown in Figure 10.9b was obtained from lateral interpolation between the five soundings. This cross section suggests that there is a high-conductivity layer beneath stations 2 and 3, interpreted to be the area of brine contamination from the brine pond shown in Figure 10.9a.

Reviews of the instrumentation available for EM induction systems, their applicability for environmental site characterization, and EM interpretational methods are given by Frischknecht et al. (1991), Hoekstra and Blohm (1990), Goldstein (1994), Mac Lean (1992), and McNeill (1990).

(a)

N

0   100   200 Feet
0   50 Meters

A

20 mmhos/m
40
60
80

BRINE
POND

EM34
MEASUREMENT
LOCATIONS

1

2

3

EM47
SOUNDING
LOCATIONS

4

A'

(b)

A   TDEM EM 47 STATIONS   A'

1   2   3   4   5

0

DEPTH (feet)

42(26)   35(29)   23(43)   21(48)   24(29)

64(16)   24(42)   25(40)   85(12)   73(14)

100

9(11)

8(125)

200

0   100   200 Feet
0   50 Meters

8 Resistivity, Ohm-m
(125) Conductivity, mmhos/m
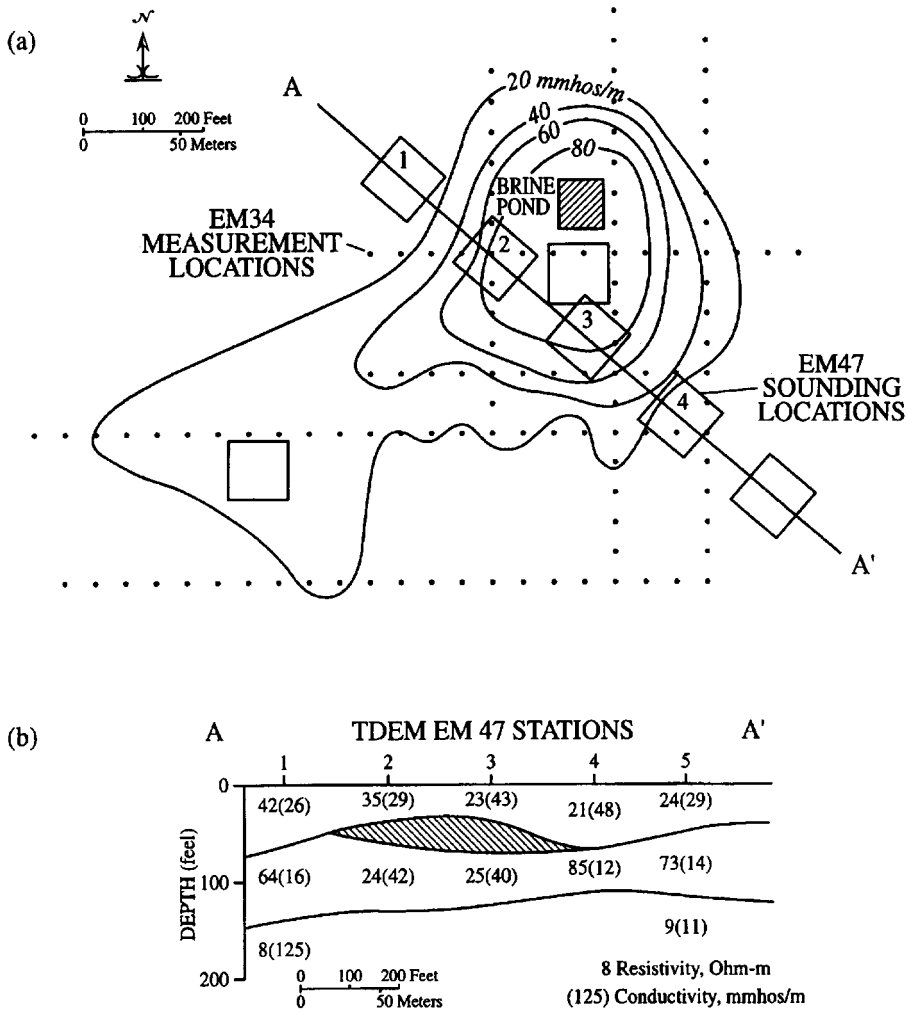
FIGURE 10.9 (a) Apparent conductivity contour map over a brine evaporation pit interpreted from EM profiles. (b) Geoelectric cross section along A-A' traverse interpreted from five EM soundings whose locations are indicated by the five numbered boxes shown in (a). (From Hoekstra, P., Lahti, R., Hild, J., Bates, R. C., and Phillips, D. 1992. Case histories of shallow time domain electromagnetics in environmental site assessment. *Groundwater Monitoring Review.* 12(4), 110-117. With permission.)

10.2.2.3   Ground-Penetrating Radar

Ground-penetrating radar (GPR) is a relatively new geophysical tool that has become increasingly popular with the need to better understand near-surface conditions. GPR methods use electromagnetic energy at frequencies of 10 to 1000 MHZ to probe the subsurface. At these frequencies, the separation (polarization) of opposite electric charges within a material that has been subjected to an external electric field dominates the electrical response. The dielectric constant or relative permittivity, which is the ratio of the electric permittivity of the material relative to the electric permittivity of free space, is used to describe these high-frequency electric properties. In general, GPR performs better in unsaturated coarse or moderately coarse textured soils; GPR performance is often poor in electrically conductive environments such as saturated systems or systems dominated by the presence of clays. However, GPR techniques have been successful in both saturated environments (Fisher et al., 1992) and in subsurface environments that have nonexpanding clays such as at the Savannah River site in South Carolina (Wyatt et al., 1996).

GPR systems consist of an impulse generator which repeatedly sends a particular voltage and frequency source to a transmitting antenna. A signal propagates from the transmitting antenna through the earth and is reflected and scattered primarily by subsurface dielectric contrasts; the modified signal is subsequently recorded by the receiving antenna. Dielectric constants are affected by material saturation, porosity, material constituency, temperature, and pore fluid composition. Approximate dielectric constants for common near-surface materials are given in Table 10.5. This table shows that most dry materials have dielectric constants of 3 to 7, which are much lower than the dielectric constants of saturated materials.

The most common GPR acquisition mode is surface common-offset reflection, which involves the collection of one trace per surface location from a transmitter-receiver antenna pair (Figure 10.10a). With this acquisition mode, data are displayed as two-dimensional wiggle trace profiles in real time during data collection and can be stored digitally for subsequent data processing. Uses of common-offset reflection GPR data include mapping the top of bedrock as well as locating objects whose electrical properties are different from those of the host environment, such as buried waste tanks. GPR data have been used in hydrogeological investigations to locate the water table and to delineate shallow, unconsolidated aquifers (Beres and Haeni, 1991). Figure 10.11 is an example of a 100 MHZ GPR profile that reveals the water table, interfaces between different hydrostratigraphic units, and top of bedrock.

GPR resolution, defined as the ability of the system to distinguish two signals that are close together in time, varies as a function of antenna frequency and the electrical properties of the host material. Davis and Annan (1989) suggest that for a wet soil, vertical resolution on the order of 7 m is possible using 10 MHZ radar antennas, 0.5 m resolution is possible using 100 MHZ radar antennas, and less than a centimeter resolution is possible using 1000 MHZ radar antennas. Increasing the frequency increases the
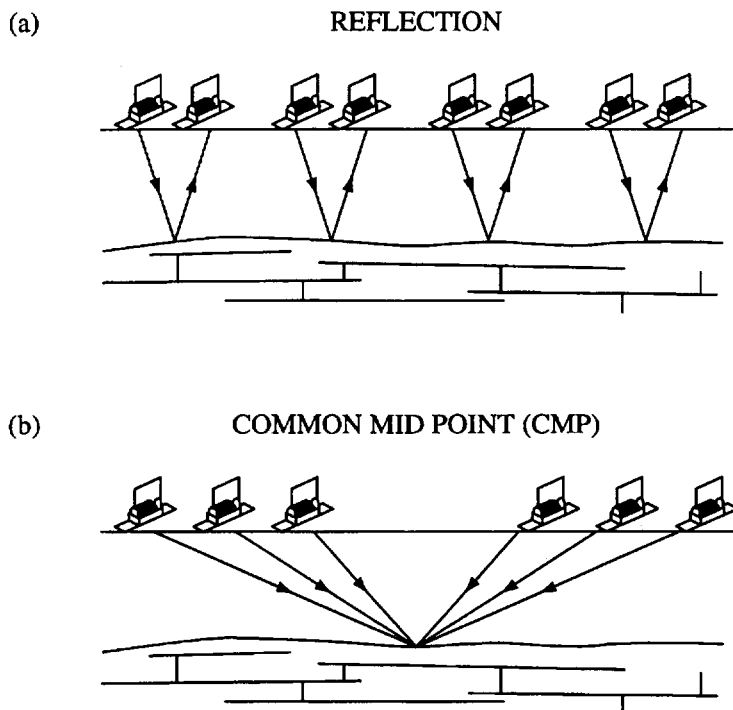
FIGURE 10.10 GPR acquisition geometry: (a) common-offset reflection mode and (b) common-midpoint mode. (From Annan, A. P. and Cosway, S. 1992. Ground penetrating radar survey design. *Proc. of the Symposium on the Applications of Geophysics to Engineering and Environmental Problems (SEGEEP)*. Chicago. With permission.)
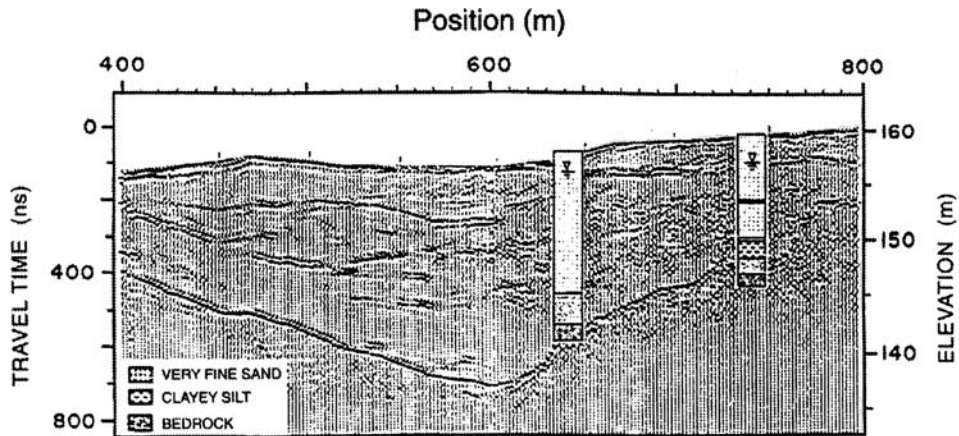
FIGURE 10.11 Ground penetrating radar profile over a shallow valley revealing water table, interfaces between different hydrostratigraphic units, and top of bedrock. (From Annan, A. P. and Cosway, S. 1992. Ground penetrating radar survey design. *Proc. of the Symposium on the Applications of Geophysics to Engineering and Environmental Problems (SEGEEP)*. Chicago. With permission.)

resolution but decreases the penetration depth. Also affecting the radar signal penetration depths are the radar system performance, reflection properties at boundaries where electrical properties vary, and attenuation which is controlled by scattering, electrical properties of surface and subsurface materials, and geometrical spreading losses. Generally, one chooses a radar center frequency that yields both sufficient penetration and resolution; for field applications this is often between 50 to 200 MHZ. It is, however, possible to collect radar data at the same location using several different frequency antennas. The use of multiple frequency antennas enables sampling over a greater depth range with optimal resolution for all depths (Smith and Jol, 1992).

Estimation of the dielectric constant is necessary to infer quantitative hydrogeological information from GPR data. For the high frequency range of interest used with GPR methods, the propagation phase velocity (V) in a low electrical conductivity material can be related to the dielectric constant (κ) as:

$$\kappa = \left(\frac{c}{V} \right)^2 \qquad (4)$$

(Davis and Annan, 1989), where c is the propagation velocity of electromagnetic waves in free space (3×10^8 m/s).

Information about the dielectric structure of the shallow subsurface is potentially available from velocity analysis of GPR data acquired with surface common-midpoint (CMP) or crosshole acquisition geometry. GPR CMP geometry results in gathers of traces where the midpoint between the transmitting and receiving antennas is common to several antenna pairs as shown in Figure 10.10b. GPR CMP data have been processed using a standard seismic data processing sequence (Fisher et al., 1992; Wyatt et al., 1996). Crosshole tomographic GPR techniques consist of measuring direct electromagnetic energy from a transmitter in one borehole to a receiving antenna in another borehole. The acquisition geometry for crosshole radar is similar to that shown in Figure 10.3; for radar acquisition the data from each transmitting-receiving antenna pair must be collected individually. Although CMP and crosshole data are time consuming to acquire, and the acquisition and processing of these data are still topics of research, they permit velocity analysis which can be used to estimate dielectric constants. Dielectric constants can then potentially be used to distinguish between lithologies (Table 10.5) or to infer hydrogeological properties.

Petrophysical models are necessary to link dielectric estimates that are potentially available from GPR data to hydrogeologic parameters. For example, Topp et al. (1980) used a regression analysis to obtain an empirical relation between dielectric constant and volumetric moisture content (θ):

$$\theta = -5.3 \times 10^{-2} + 2.92 \times 10^{-2} \kappa - 5.5 \times 10^{-4} \kappa^2 + 4.3 \times 10^{-6} \kappa^3 \tag{5}$$

This relation suggests that the dielectric constant increases with water content. Greaves et al. (1996) used the Topp equation to transfer interval velocity estimates obtained from surface CMP GPR data into subsurface water content estimates. Comparison of their estimates with published hydrogeologic and geologic descriptions of the survey area suggests that the water content section obtained from radar data was reasonable. The Topp equation was also used by Hubbard et al. (1997b) to transfer dielectric constants obtained from crosshole radar measurements into water content estimates. Comparison of the radar-assisted water content estimates with those obtained at the same location using a borehole neutron probe (see Section 10.2.6.3) revealed that crosshole radar data can be extremely useful for estimating water content.

For a sand-clay system, a mixture formula was used by Wharton et al. (1980) to express the bulk dielectric constant as a function of the dielectric constants of air (κ_a), water (κ_w), sand (κ_s), and clay (κ_{cl}), as well as the porosity (ϕ), fractional water saturation (S_w), and the volumetric clay fraction (V_{cl}):

$$\kappa = \left[(1-\phi) V_{cl} \sqrt{\kappa_{cl}} + (1-\phi)(1-V_{cl}) \sqrt{\kappa_s} + S_w \phi \sqrt{\kappa_w} + (1-S_w) \phi \sqrt{\kappa_a} \right]^2 \tag{6}$$

Knoll et al. (1995) found under laboratory conditions that Equation (6) fits the observed dielectric behavior of a constructed unconsolidated sand-clay mixture. They converted the relation between effective dielectric constant, saturation, porosity, and clay fraction shown in Equation (6) to a relation between effective dielectric constant, saturation, and intrinsic permeability using petrophysical relationships. Their petrophysical formulas, which were developed theoretically and verified in the laboratory, are shown in Figure 10.12. These curves reveal that the dielectric constant is a function of both saturation and intrinsic permeability and that the dielectric constant is much more sensitive to saturation state than to the intrinsic permeability when the permeability values are high. Using the petrophysical relations given by Knoll et al. (1995), Hubbard et al. (1997a) developed a technique to estimate permeability and saturation in bimodal systems using CMP or crosshole GPR data. Crosshole radar data have also been used for dynamic process monitoring (see Section 10.2.1) to monitor water infiltration in fractured systems as well as the geometry of a condensation front caused by the application of heat to the subsurface (Hubbard et al., 1997b).

The petrophysical relations discussed above were all based on a pore fluid of water; however, contaminants may have a different dielectric constant than water and may be detectable in these cases using radar methods. GPR has been successful in imaging PCE contamination (Greenhouse et al., 1993) and TCE contamination (Brewster and Annan, 1994). Because of the high frequencies employed, GPR offers the best resolution of any surface geophysical technique when operating under optimal conditions. An excellent overview of the GPR method is given by Davis and Annan (1989).

10.2.3 Seismic Methods

Seismic reflection, crosshole transmission, and refraction methods use artificially generated, high-frequency (100 to 5000 Hz) pulses of acoustic energy to probe the subsurface. These pulses can be delivered to the subsurface with hammer blows, weight drops, blasting caps, rifle or shotgun slugs, piezoelectric devices, or vibrators. These disturbances are produced at a point and propagate outward as a series of wavefronts. The passage of the wavefront creates a motion which can be detected by a sensitive geophone and recorded on the surface; the path that the energy follows is conceptualized by a ray drawn normal to the wavefront.
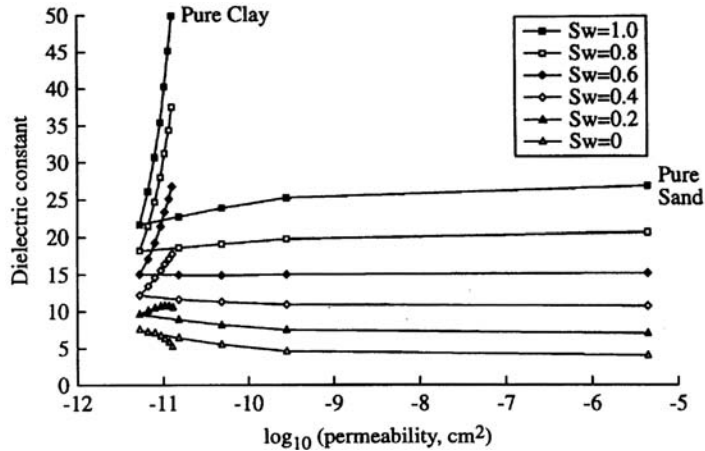
FIGURE 10.12 Petrophysical curves of Knoll et al. (1995) relating dielectric constant to fractional water saturation (S_w) and $\log_{10}$ (permeability) where the intrinsic permeability is in cm². (From Knoll, M., Knight, R., and Brown, E. 1995. Can accurate estimates of permeability be obtained from measurement of Dielectric properties? *SAGEEP Annual Meeting Extended Abstracts.* Orlando, FL, Apr. 23-26. With permission.)

According to the theory of elasticity upon which seismic wave propagation is based, four different waves are produced by a disturbance, and these waves travel with different propagation velocities governed by the elastic constants and density of a material. One type of wave is transmitted by a back-and-forth particle movement in the direction of the propagating wave. This type has the greatest speed and is thus the first to arrive at a geophone; for this reason it is called the P (primary)-wave. Other terms for this wave are the compressional or longitudinal wave. The particle motion of transverse waves is perpendicular to the direction of the wavefront propagation. Transverse waves, also called S (secondary or shear)-waves, have lower velocities than the P-wave and thus arrive later in the recording. In contrast to both P- and S-waves, which are transmitted though the interior of a solid and are thus classified as body waves, surface waves, including both Raleigh and Love waves, are confined to a region near the free surface of a medium. Because it is easiest to detect the P-wave energy, most seismic exploration techniques have focused on the use of this wave, and in the following discussion, we also concentrate exclusively on information available from P-waves.

Figure 10.13 shows the major paths of P-wave energy upon disturbance in a horizontally-layered, two-medium system where the *seismic velocity* of the upper layer is less than that of the lower layer. The first arrival at the detector or geophone is the (compressional) air wave, which is followed by the direct compressional wave that propagates along a path from the disturbance to the geophone. P-wave energy is reflected from the interface and also refracted from the head wave that travels along the interface. In addition to these compressional arrivals, ground motion from Raleigh surface waves, or ground roll, also contributes to the signal recorded at a geophone. The events commonly used for interpretation are the reflected P-waves, the direct P-waves, and the refracted P-waves. Different seismic methods focus on different seismic wave arrivals — reflection techniques focus on reflected arrivals, crosshole methods mainly utilize direct waves, and refraction methods are designed to detect refracted head waves. The principles of each of these methods are discussed below.

10.2.3.1 Reflection

The surface reflection technique is based on the return of reflected P-waves from boundaries where velocity contrasts exist. Seismic reflection data are usually collected as common-shot or common-receiver gathers which are sorted during processing into common-midpoint or CMP gathers (Yilmaz, 1987). These gathers of traces represent reflections from a subsurface location (the midpoint) that has been sampled by several source-receiver pairs in a fashion similar to that shown in Figure 10.10b.
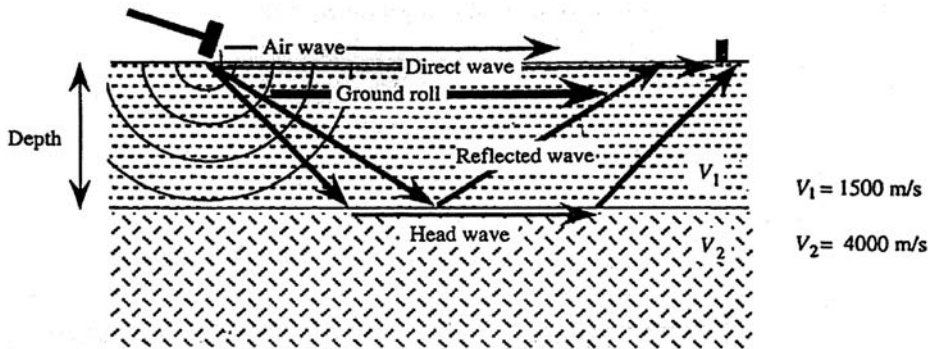
FIGURE 10.13 Major ray paths of P-wave energy. (From Burger, H. R. 1992. *Exploration Geophysics of the Shallow Subsurface.* Prentice Hall, Englewood Cliffs, NJ. With permission.)

Due to the lack of well-defined velocity contrasts in unconsolidated and unsaturated materials and to the similar arrival times of the direct wave, reflected wave, and ground roll (Figure 10.13), seismic reflection data acquisition is often difficult in the near-surface. Computer-based processing of seismic reflection data generally produces a wiggle-trace profile that resembles a geologic cross section; Figure 10.14 is an example of a seismic reflection section over a steep-sided valley that is floored by bedrock. An interpreted borehole log has been inserted into the section at the borehole location and reveals that the seismic reflections correspond to interfaces between sands, clays, gravels and bedrock. The vertical resolution of a few meters shown in this figure is an example of good quality surface seismic data collected in an area favorable for seismic data acquisition with careful attention to acquisition parameters.

In addition to obtaining structural and stratigraphic information about the subsurface from the P-wave reflection amplitude and arrival times, information about seismic velocity is available through seismic data processing (Yilmaz, 1987). Seismic P-wave velocity is affected by porosity, permeability, pore fluid type, depth of burial, consolidation, and temperature. Table 10.6 lists approximate P-wave velocities for common near-surface materials. This list reveals that unique relations between seismic
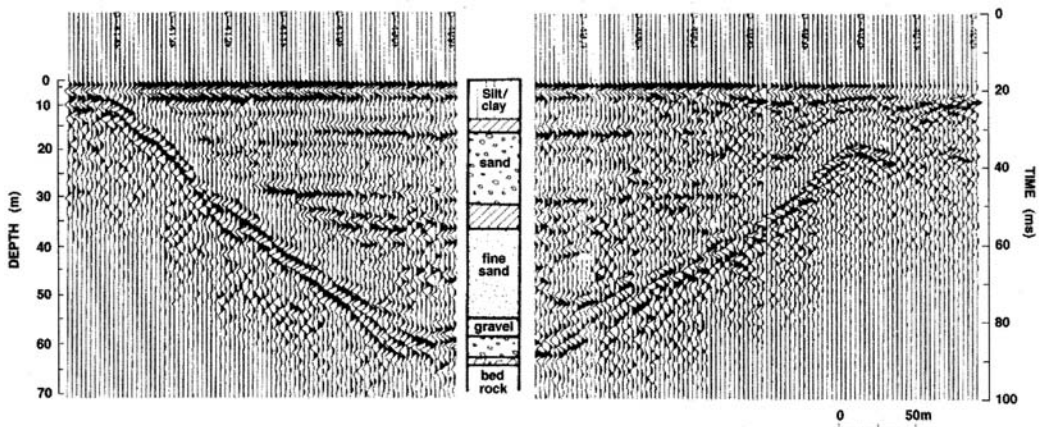
FIGURE 10.14 High resolution near-surface seismic section over a steep-sided valley. Interpreted borehole log inserted at the borehole location reveals that reflections imaged are interfaces between different hydrostratigraphic units. (From Pullan, S. E. and Hunter, J. A. 1990. Delineation of buried bedrock valleys using the optimum offset shallow seismic reflection technique, in *Geotechnical and Environmental Geophysics: Vol. 3, Geotechnical, S.E.G. Investigations in Geophysics 5.* Ed. Stanley Ward. 75-87. With permission.)

TABLE 10.6 Seismic P-Wave Velocities for Typical Near-Surface Materials

Material	P–Wave Velocity (m/s)
Unconsolidated materials	
Weathered layer	300–900
Soil	250–600
Alluvium	500–2000
Clay	1100–2500
Unsaturated sand	200–1000
Saturated sand	800–2200
Unsaturated sand/gravel	400–500
Saturated sand/gravel	800–2200
Unsaturated glacial till	400–1000
Saturated glacial till	1700
Consolidated Materials	
Sandstone and shale	2000–4500
Chalk	2300–2600
Limestone	2000–6000
Coal	2200–2700
Basalt	5400–6400
Granite	5000–6400
Metamorphic	3500–7000
Other	
Water	1400–1600
Ice	3400–3800
Air	331.5

(Adapted from Press, F. 1966. Seismic velocities, in *Handbook of Physical Constants*, rev. ed., Clark, S. P., Jr., Ed. Geologic Society of America Memoir. 97: 97-173.)

P-wave velocities and lithology generally do not exist. However, some generalities can be made regarding seismic velocities and lithology (Burger, 1992):

- Velocities in unsaturated sediments are lower than in saturated sediments.
- Velocities in unconsolidated sediments are lower than in consolidated materials.
- Velocities are very similar in unconsolidated, saturated sediments.
- Velocities in weathered rocks are lower than in similar but unweathered rocks.

In general, then, it is possible using seismic velocity to distinguish dry from saturated sediments and sediments from consolidated rocks.

Both empirical and theoretical approaches have been used to relate seismic P-wave velocity to hydrogeological parameters. For example, a velocity (V) – porosity (ϕ) relationship is given by the Wyllie time-average equation using estimated matrix velocity (V_m) and pore fluid velocity (V_f) values:

$$\frac{1}{V} = \frac{1-\phi}{V_m} + \frac{\phi}{V_f} \qquad (7)$$

Klimentos (1991) used regression analysis to obtain the following relation between velocity (V), porosity (ϕ), and clay content (C) in consolidated, sand-clay systems:

$$V = 5.59 - 6.93\phi - 2.18C \qquad (8)$$

For an investigation where other information is available, regression techniques are commonly used to relate site-specific velocity measurements to porosity or lithology. Marion et al. (1992) used a theoretical microgeometrical approach to relate seismic P-wave velocity to pressure and porosity within an unconsolidated, sand-clay system. This relation was verified empirically in the laboratory, and the model was extended by Rubin et al. (1992) to investigate the relationship between log permeability, effective pressure, and seismic P-wave velocity as shown in Figure 10.15. This figure suggests that a unique relationship between P-wave velocity and log permeability does not exist. A review of shallow seismic acquisition and processing techniques is given by Steeples and Miller (1990).
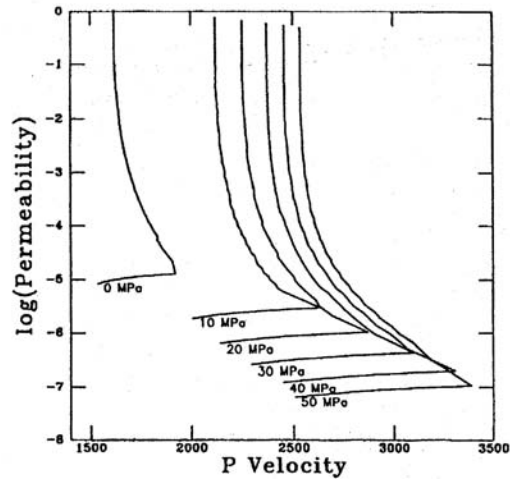
FIGURE 10.15 Petrophysical curves relating seismic P-wave velocities, log(permeability), and pressure in unconsolidated sand-clay systems. (Adapted from Marion, D., Nur, A., Yin, H., and Han, D. 1992. Compressional velocity and porosity in sand-clay mixtures. *Geophysics.* 57, 554-563 and Rubin, Y., Mavko, G., and Harris, J. 1992. Mapping permeability in heterogeneous aquifers using hydrological and seismic data. *Water Resources Research.* 28(7), 1192-1800.)

10.2.3.2 Crosshole Transmission

Crosshole tomographic data acquisition is possible using electrical resistivity, EM, radar, and seismic methods; of these, seismic methods are the most common. Although the majority of the crosshole tomographic seismic data sets have been collected for research projects, the extremely high resolution of up to 0.5 m offered by this method suggests that it will undoubtedly become more common for site characterization. A typical seismic crosshole topographic geometry is shown in Figure 10.3. The multiple sampling of the intrawellbore area permits very detailed estimation of the velocity structure (Hyndman et al., 1994). As seismic P-wave velocities can be related to lithological and hydrogeological parameters as discussed above, this extremely high resolution method is ideal for detailed stratigraphic and hydraulic characterization of interwell areas (Hyndman et al., 1996). Although not yet established as a field method, seismic imaging of organic contaminants in the laboratory has provided a fundamental step toward the application of seismic tomographic imaging of interwell contamination (Geller and Myer, 1995). The processes of crosshole seismic data inversion and incorporation of these data with hydrological data are discussed in detail in Chapter 11.

10.2.3.3 Refraction

With refraction methods, the incident ray is refracted along the target boundary before returning to the surface (Figure 10.13). The refracted energy arrival times are displayed as a function of distance from the source, and interpretation of this energy can be accomplished manually by using simple software or forward modeling techniques. The arrival times and distances can be used to obtain velocity information

directly. Refraction techniques are most appropriate when there are only a few shallow (<50m) targets of interest, or where one is interested in identifying gross lateral velocity variations or changes in interface dip (Lankston, 1990). Seismic refraction methods yield much lower resolution than seismic reflection and crosshole methods. However, because refraction methods are inexpensive and acquisition may be more successful in unsaturated and unconsolidated environments, it is sometimes chosen over reflection methods for applications such as determining the depth to the water table and to the top of bedrock, the gross velocity structure, or for locating significant faults. A review of the refraction method is given by Lankston (1990).

10.2.4 Gravitational Methods

Measurements of the changes in gravitational acceleration can be used to obtain information about subsurface density variations. As density is a bulk property of rocks and tends to be consistent throughout a geological formation, gravity methods are used to identify gross features based on density variations. Due to the lower resolution offered by this method, it is not commonly used for detailed site characterization. It does, however, provide an inexpensive way to detect targets such as the interface between sedimentary overburden and bedrock or locations of significant faults. Gravity methods have also been used to detect sinkholes, other subsurface voids, and landfill boundaries. The common measuring device is a gravimeter, an instrument which is portable and easy to use. A spring balance inside the gravimeter measures differences in the weight of a small internal object from location to location; the weight differences are attributed to changes in the acceleration of gravity due to lateral variations in subsurface density. Measurements can be collected at a regional or local scale depending on the station spacing, which is usually less than half of the depth of interest.

Gravity measurements must be adjusted prior to interpretation by comparing the observed gravity values with the calculated response of a homogeneous formation. The theoretical response to the gravitational field due to such factors as the datum, latitude, tidal, altitude, terrain, drift, and regional gradient are calculated and subtracted from the raw data to yield what is called a gravity anomaly field (Zodhy, 1974). A Bouguer anomaly refers to the value that is obtained after latitude corrections, elevation corrections, and (sometimes) terrain corrections have been applied to the gravity data. An example of the gravity data reduction and interpretation process is shown in Figure 10.16. In this study, gravity data were collected to determine the depths to bedrock in an urban area where the drilling of numerous wells was not feasible (Kick, 1985). Detailed gravity measurements were collected by siting gravity stations at 40-foot intervals and obtaining detailed elevation information which was necessary for the elevation correction at these stations. After corrections were applied to the gravity measurements, the Bouguer gravity values were plotted against distance (Figure 10.16a). A regional gravity trend, obtained from a smoothed version of the data in conjunction with the state gravity map, is also shown in Figure 10.16a. This regional trend was removed from the gravity curve to enhance the local effects such as those caused from soil thickness variations; the new data are called the residual gravity data. The interpretation of the residual gravity data is shown in Figure 10.16b. This figure demonstrates that, in all cases except one, the bedrock depths interpreted from the residual gravity data are close to those confirmed by borehole data.

Anomaly field interpretations are non-unique; a large number of hypothetical features with varying density, depth, and shape can produce the same gravity anomaly. Qualitative interpretation usually consists of constraining a profile or contoured anomaly map with other known geologic information. Forward modeling techniques and characteristic curve matching are also used for interpretation. Reviews of the gravity technique and applications to environmental studies are given by Hinze (1990) and by Butler (1991).
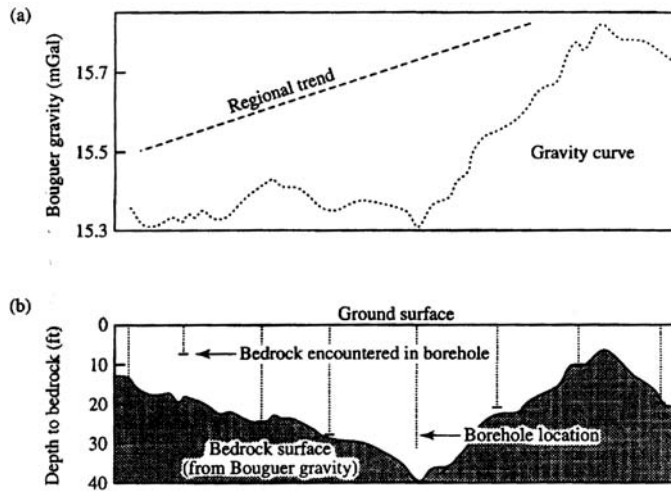
FIGURE 10.16 Interpretation of gravity data to estimate depth to bedrock (Kick, 1985). (a) Bouguer gravity curve obtained from reduced gravity measurements and regional gravity trend. (b) Interpretation of the local gravity data after removal of the regional trend showing that the bedrock depths interpreted from the gravity data are similar to those confirmed by borehole information. (From Kick, J. F. 1985. Depth to bedrock using gravimetry. *Geophysics: The Leading Edge of Exploration, Society of Exploration Geophysicists.* 4(4), 38-42. With permission.)

10.2.5 Magnetic Methods

Magnetic methods measure the direction, gradient, or intensity of the earth's magnetic field. The intensity of the magnetic field at the earth's surface is a function of the location of the observation point in the primary earth magnetic field as well as local or regional concentrations of magnetic material such as magnetite, the most common magnetic mineral. Magnetometers are used to measure the total geomagnetic field intensity or relative values of the vertical field intensity, and magnetic gradiometers measure the horizontal and vertical gradient of this magnetic field. After correcting for the effects of the earth's natural magnetic field, magnetic data can be presented as total intensity, relative intensity, and vertical gradient anomaly profiles or contour maps. Interpretation of magnetic surveys generally involves forward modeling or mapping of the anomalies and correlating them with other known geologic information. This interpretation is non-unique.

Magnetic methods are generally used to identify gross features at a resolution similar to that of seismic refraction and gravitational methods. As magnetic signatures depend to a large extent on magnetic mineral content, which is low in most sediments that comprise aquifers, magnetic field work is not commonly employed for near-surface investigations. It is used, however, to map the depth to basement rock, provided that it contains sufficient magnetic minerals. Magnetic methods with fine station spacing and high lateral resolution (a few meters) are now one of the most commonly used geophysical methods for site investigation due to their ability to locate metal objects such as drums and abandoned drill hole casings which are buried in near-surface sediments. A review of magnetic methods as applied to environmental problems is given by Hinze (1990).

10.2.6 Borehole Geophysical Methods

Borehole geophysics refers to the process of recording and analyzing physical property measurements made in wells. One-dimensional log information can be correlated between wells to extrapolate information such as test or core data and can be used in conjunction with surface geophysical data. The volume of material sampled by the borehole log, or the volume of investigation, is related to the log type,

source-detector spacing, and subsurface material, and thus varies with well site conditions and logging parameters employed. Interpretation of the recorded log data often involves comparing several different logs displayed side-by-side or by cross-plotting data from one log against data from other logs, core analyses, or tests. For hydrocarbon exploration, the decision to test and complete a well is based to a large extent on geophysical log information, and as a consequence, most of the interpretation guidelines for borehole geophysics have been developed for borehole and rock environments encountered in petroleum exploration. An excellent reference for borehole geophysics applied to groundwater investigations is given by Keys (1989) which is the reference for the following information where not otherwise cited. A review of the process of drilling, coring, and sampling water wells is given by Shuter and Teasdale (1981).

Borehole measurements are made by lowering a sonde into the borehole on the end of an electric cable. The sonde is a probe, 2.5 to 10.0 cm in diameter and 0.5 to 10.0 m in length, which encloses sources, sensors, and the electronics necessary for transmitting and recording signals. Measurements made in the borehole are recorded on the surface in analog form on chart paper or in digital form as shown in Figure 10.17, which is an example of several logs that have been collected in a shallow water well.

The following discussion focuses on the underlying physical principles of logging methods that are currently used in groundwater applications or that have potential for aiding these investigations, including caliper, electric, nuclear, and acoustic. Table 10.3 lists the applications, limitations, and borehole conditions required for each method as well as the frequency of use of the method for groundwater applications. Although some borehole techniques have potential for aiding near-surface studies, they are still not commonly used, usually because of the sophistication and thus higher cost of the method. Table 10.4
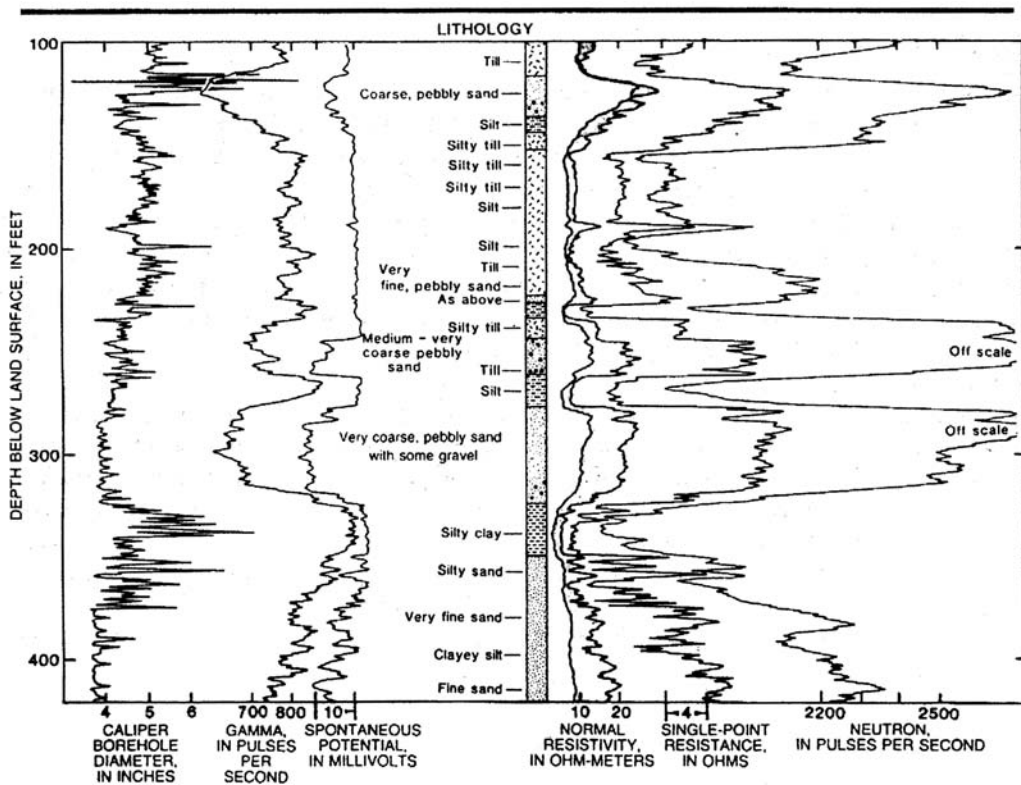
FIGURE 10.17 Borehole geophysical and lithologic logs collected at a typical shallow water well. (From Keys, S. W. 1989. *Borehole Geophysics Applied to Ground-Water Investigations*. National Water Well Association, Dublin, OH. With permission.)

summarizes the applications of the most commonly used borehole geophysical methods. These applications refer to information available from individual logs only; more information, for example, about fault displacement, can be obtained by correlating several logs or by integrating the well logs with surface geophysical or other data.

10.2.6.1 Caliper Logging

Caliper logs are mechanical or acoustic tools that measure the diameter of the borehole. The mechanical caliper tool is a single arm connecting between one and six caliper probes that are pressed against the borehole wall by spring pressure. As the tool is pulled up the borehole, the mechanical caliper probes move in response to changes in borehole diameter and the acoustic calipers measure the reflection transit time of an acoustic signal from the borehole wall. Changes in the diameter of the borehole affect the response of all geophysical tools, and thus a caliper log is generally collected in conjunction with all logging types to aid in interpretation and correction.

10.2.6.2 Electric Logging

Electric logs measure potential differences due to the flow of electric current in and adjacent to the well. There are many different types of electric logs, all of which must be run prior to casing and in saturated boreholes.

Single-point resistance is a very simple tool that is used to measure resistance by passing a constant current between an electrode in the well and another located on the surface. Electrical resistivity, measured in Ohm-meters (Ω-m), includes the dimensions of the material being measured and is an intrinsic property of the material. Resistance depends on the material composition as well as on the cross-sectional area and length of the current path through the medium. Electrical resistance (r), measured in Ohms (Ω), is related to resistivity (ρ) by

$$r = \frac{\rho L}{S} \tag{9}$$

where S is the cross-sectional area normal to the flow of current (m^2) and L is the length of the current path in meters.

The single-point resistance log shown in Figure 10.17 reveals the usefulness of this tool for high resolution lithological identification; the sand and gravel units characteristically give a higher resistance reading than the silts and clays. *Resistivity* measurements are made by sending a current through electrodes lowered down the borehole and recording the voltage across other electrodes located on the sonde. Resistivity calculations are performed using Ohm's law (Equation [1]). Resistivity logs are commonly used to calculate water quality, as well as saturation and porosity (Equation [3]). Short and long *normal*, *focused,* and *lateral* logs are all resistivity methods that differ by the electrode configuration on the sonde, and *microresistivity* logs employ electrodes that are forced to contact the borehole wall. All configurations measure the resistivity of the subsurface material, pore fluid, borehole fluid, or some combination of these which is used for interpreting the stratigraphic or hydrogeologic parameters as shown in Table 10.4. *Dipmeter* tools use three or four contact microresistivity probes connected to a single arm which press against the borehole walls to measure resistivity simultaneously at multiple orientations within the borehole. Interpretation involves correlating resistivity measurements made from each probe to estimate primary sedimentary structure orientation of the units penetrated by the wellbore. Computer analysis of dipmeter data is most effective. A *spontaneous potential (SP)* log measures the natural potential difference between a single electrode in the wellbore and a reference electrode at the surface. These potentials exist due to a combination of electrochemical effects created, for example, by migration of ions from concentrated to more dilute solutions. Less significant in sedimentary environments are additional effects caused by electrokinetic (streaming) potentials that result from fluid moving in or out of the hole or redox potentials caused by ionic charge accumulation between metallic mineral grains and the fluids adjacent to the grains (Daniel and Keys, 1990).

10.2.6.3 Nuclear Logging

Nuclear logging entails the detection of unstable isotopes near the borehole. The considerable advantage of nuclear logs over electric logs is that they can be run after casing has been installed. As isotopes decay, they emit radiation such as gamma photons or rays from the nucleus. Gamma proton and neutron radiation sources are often used in borehole applications because of their ability to penetrate dense material such as rock and casing. Borehole geophysical tools that measure radioactivity of nearby formations may be classified as those that detect natural gamma radiation, those that employ controlled gamma rays to induce radiation, and those that use neutron sources to induce nuclear processes. The radioactivity is measured as electronic pulses, and the quantity and amplitude of the pulses yield information about the surrounding formation. Logging tools that use artificial radioisotopes as sources are regulated by governmental agencies and require a license for use.

The *gamma log* uses a scintillation detector to measure total naturally occurring gamma radiation of the material penetrated by the borehole within a selected energy range. The three most common naturally occurring radioactive materials that affect the gamma log are potassium-40, uranium-238, and thorium-232. The usefulness of this log lies in the fact that these isotopes are generally more abundant in shales and clays and less common in sands and calcareous materials, and that measurements can be reliably made above the water table. The gamma log shown in Figure 10.17 reveals the characteristically high gamma log count rate associated with silts and clays. In addition to the count rate that is measured with the gamma log, the *gamma-spectrometry* method records the amplitude of the pulses over a wide energy range. Analysis of this energy permits more diagnostic information on lithology and also permits estimation of the type and quantity of radioisotopes that may be contaminating the groundwater.

Gamma-gamma or *density* logs record gamma radiation that originates from an artificial gamma source in the well which is backscattered by the borehole and surrounding material. The count rate of the backscattered gamma rays can be related to the electron density of the material which is in turn proportional to the bulk density of the material. If the fluid and grain densities are known, the bulk density measured with the gamma-gamma log can be used to calculate porosity:

$$\phi = \frac{\text{grain density} - \text{bulk density}}{\text{grain density} - \text{fluid density}} \tag{10}$$

Because moisture content affects the bulk density of materials, gamma-gamma logs can be used to record changes in moisture above the water table.

Neutron tools consist of an artificial low-energy wellbore neutron source and one or two neutron detectors. The neutrons emitted by the source lose energy upon collision with other elements in the vicinity of the borehole. Because hydrogen has a mass similar to the neutron, it is the most effective element at slowing the neutrons. The quantity of slowed neutrons is thus interpreted to be proportional to the quantity of hydrogen present, which is in turn interpreted to be proportional to the moisture content or saturated porosity. For most materials, resistivity and neutron logs have a similar character because of the relationship between saturated porosity and pore fluid resistivity (Figure 10.17).

10.2.6.4 Acoustic Logging

Acoustic (sonic or velocity) tools transmit an acoustic pulse from a source, through the fluid and material in the vicinity of the borehole, and to the detector. All acoustic logs require saturated borehole conditions for signal transmission. These tools emit an acoustic source at frequencies of 10 to 35 KHz (0.01 to 0.035 MHZ) which creates compressional or P-waves. As the waves travel, some of the energy is refracted back to two receivers located on the sonde. The difference in travel time between the receivers is used to calculate interval velocity, which is recorded as a function of depth in the wellbore. *Acoustic waveform* logging entails the recording and interpretation of the entire waveform rather than just the time of the first arrival of energy as is recorded with the acoustic log. Analysis of the amplitude, frequency, and other

components of the wave train may yield information about the quality of the bond between casing and cement and cement and borehole wall. Additionally, attributes of the signal can yield information about the formation penetrated by the borehole. For example, correlations have been found between attenuation and permeability of fractures.

The *acoustic televiewer (ATV)* emits high frequency (1.3 MHZ) acoustic energy toward the borehole wall. The transducer rotates clockwise at about three rotations per second. During rotation, the transducer sends signals toward and receives signals from the borehole wall. Fractures or other openings in the borehole wall force the signal to travel farther through the fluid and thus to be more attenuated. Attenuation causes amplitude variations which can be visually captured by adjusting the contrasts on the display. The display can be thought of as a cylinder that has been opened along the north side and flattened to provide a 360-degree ultrasonic picture of the borehole wall. Planar horizontal features that intersect the well will appear on this display as dark horizontal bands and planar dipping features that intersect the well will appear as dark sinusoids with the lowest point of the curve in the direction of dip.

10.3 Characterization of Heterogeneous Aquifers

The theories and approaches presented in this section can be used to interpret the data collected by the methods discussed in Section 10.2 and in Chapters 11 and 12.

Figure 10.18 is a schematic representation of a few types of heterogeneous geological formations. Each cross section shows several blocks, each homogeneous, with a large variability between them. In aquifers, these blocks represent a discretized image of the hydraulic conductivity or other hydrogeologic parameter values, which generally vary greatly over relatively short distances. The details of images such as those of Figure 10.18 are difficult to capture in the absence of a very large number of field measurements, an unwieldy goal that is generally abandoned in favor of describing the aquifer's hydrogeologic attributes as *space random functions* (Dagan, 1989), by which properties such as conductivity, porosity, and clay content are characterized through their statistical moments.
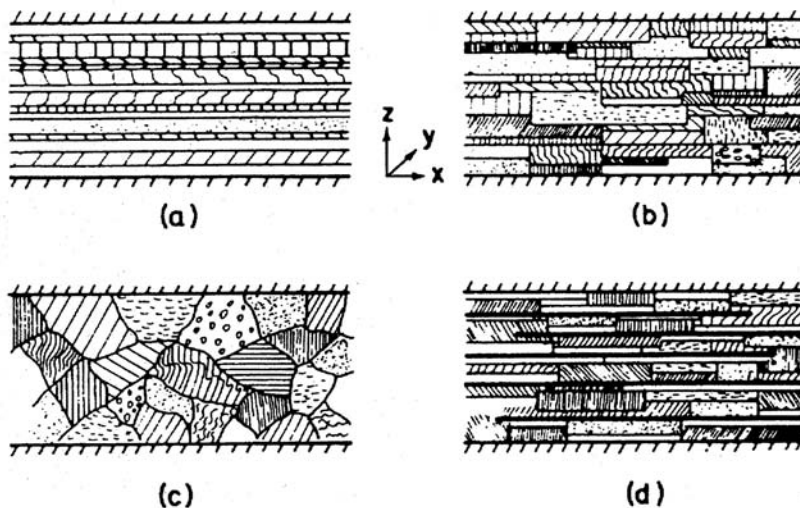
FIGURE 10.18 Schematic representation of a few types of heterogeneous structures at the local scale: (a) stratified formation with horizontal continuous bedding, (b) anisotropic formation with horizontal log conductivity correlation scale larger than the vertical, (c) isotropic three-dimensional structure, and (d) thin impervious lenses separating heterogeneous layers of the matrix; the horizontal extent of the lenses is much larger than their vertical separation distance and the horizontal correlation scale of the matrix. (From Dagan, G. 1989. *Flow and Transport in Porous Formations.* Springer Verlag, Berlin, 465 p. With permission.)

Consider an attribute $Z(\boldsymbol{x})$, where $\boldsymbol{x} = x_i$, $i = 1,2,3$ is a space coordinate and boldface letters denote vectors. As a space random function, Z can be characterized through its expected value $<Z(\boldsymbol{x})>$, where angled brackets denote the expected value operator, and a series of central moments μ_n:

$$\mu_n = <\left(Z(\boldsymbol{x}) - <Z(\boldsymbol{x})>\right)^n>; \quad n = 2,\dots,\infty \tag{11}$$

The expected value and the moments μ_n, computed from samples collected in the field, characterize the univariate probability distribution function (pdf) of Z. In the earth sciences, however, due to data limitations, Z is usually characterized through its expected value and μ_2 only. Aquifers where these two moments are invariant to translation are said to be weakly stationary.

Despite their relatively large variability, hydrogeological attributes show persistence over some distance. From a statistical point of view, such persistence is modeled via *spatial correlation*. An exhaustive statistical description requires the multivariate pdf of Z, but again, due to data limitations, the spatial correlation structure is summarized through the spatial covariance:

$$C_Z(\boldsymbol{x},\boldsymbol{x}') = <\left(Z(\boldsymbol{x}) - <Z(\boldsymbol{x})>\right)\left(Z(\boldsymbol{x}') - <Z(\boldsymbol{x}')>\right)> \tag{12}$$

The spatial correlation structure can also be summarized through the semivariogram function γ_Z:

$$\gamma_Z(\boldsymbol{x},\boldsymbol{x}') = \frac{1}{2} <\left(Z(\boldsymbol{x}) - Z(\boldsymbol{x}')\right)^2> \tag{13}$$

In stationary formations, γ_Z can be related, with one exception, to the spatial covariance as follows:

$$\gamma_Z(\boldsymbol{x},\boldsymbol{x}') = \sigma_Z^2 - C(\boldsymbol{x},\boldsymbol{x}') \tag{14}$$

where $\sigma_Z^2 = C_Z(\boldsymbol{x},\boldsymbol{x})$ is the population variance of Z. For large separation distances $|\boldsymbol{x} - \boldsymbol{x}'|$, $C_Z(\boldsymbol{x},\boldsymbol{x}')$ approaches zero and the semivariogram approaches the population variance σ_Z^2. The exception to Equation (14) is the case where a finite variance does not exist, for example in self-similar formations (Neuman, 1990).

Methods for identifying γ_Z (and hence C_Z, using Equation [14]) are described in detail in Chapter 12. The approach is to fit a parametric model of the general form:

$$C_Z(\boldsymbol{x},\boldsymbol{x}') = \sigma_Z^2 \rho\left(\frac{r_1}{a_1}, \frac{r_2}{a_2}, \frac{r_3}{a_3}\right) \tag{15}$$

to the data based on some optimality criteria. In Equation (15), ρ is the correlation function, and a_i are characteristic lengths used to normalize the lag distances $r_i = |\boldsymbol{x}_i - \boldsymbol{x}_i'|$. Note that the assumption of *stationarity* is implicit in Equation (15), hence the dependence on lags and not on the actual coordinates. In our subsequent derivations, we will assume that the coordinate system is aligned with the directions parallel and orthogonal to bedding. This assumption is particularly consequential for the discussion on *effective conductivity*, which can then be presented as a diagonal tensor.

Commonly used models for ρ are shown in Figure 10.19 (see also Dagan, 1989 and Gelhar, 1993; Chapter 12 contains several expressions for semivariogram models, which can be related to the covariances through Equation [14]). The significant difference between the models lies in the behavior of the short distance correlations. For example, the exponential covariance exhibits a "cusp"-like behavior near the origin, which is typical of sharp transitions between blocks, while the smooth behavior of the Gaussian model near the origin implies smoother transitions. The characteristic lengths a_i may have different

physical meanings when used in the different models. To unify the definition of the length scales, we define the integral scale:

$$I_i = \frac{1}{\sigma_Z^2} \int_0^\infty C_Z(r_i) \, dr_i = \int_0^\infty \rho(r_i, r_j = 0 \text{ for } j \neq i) \, dr_i; \quad i = 1, \ldots, 3 \tag{16}$$

where $r_i = |x_i - x_i'|$. The integral scale is a measure of the spatial persistence of z; larger I_i indicates that the effects of a z deviate will be experienced over longer distances in the ith direction. A summary of
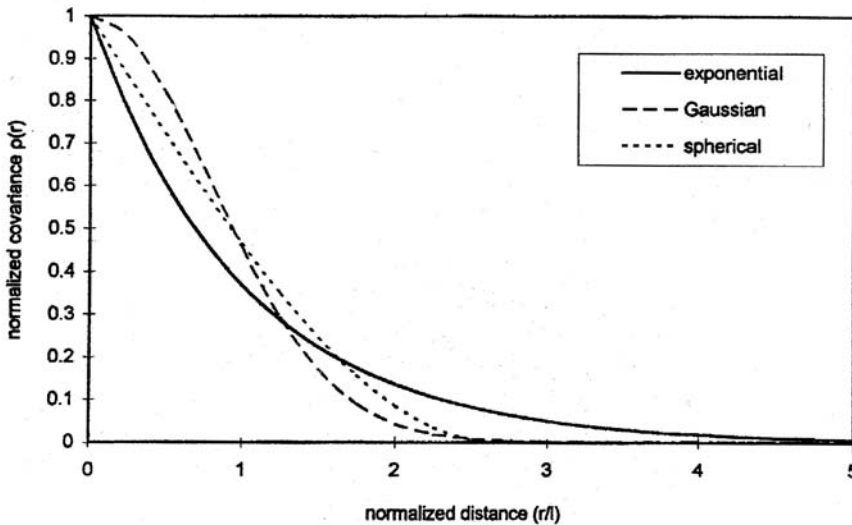
FIGURE 10.19 Commonly used spatial correlation models.

relationships between a_i and I_i (from Equations [15] and [16], respectively) is given in Dagan (1989, Part 3). Statistical heterogeneity can now be classified, based on Equation (16), as follows:

- Three-dimensional heterogeneity refers to the case where $I_1 \neq I_2 \neq I_3$.
- Axisymmetric heterogeneity is the case with $I_1 = I_2 \neq I_3$, where I_1 and I_2 are the integral scales in the plane of bedding. Taking I_3 and I_1 as the integral scales in the directions normal and parallel to bedding, respectively, allows us to define the ratio of anisotropy $e = I_3/I_1$. Figure 10.20a depicts the exponential correlation function for separation distances along the x_3 axis as a function of e. A smaller e implies a smaller I_3 and hence a lower correlation in the x_3 direction. Figure 10.20b shows the semivariograms which correspond to the covariances of Figure 10.20a, obtained using Equation (14). Lower correlation implies a faster approach of the semivariogram to its asymptotic value, referred to as the sill.
- Isotropic heterogeneity refers to the case where $I_1 = I_2 = I_3$. When this condition is not met, the aquifer is said to be statistically anisotropic. In the case of axisymmetric heterogeneity, the heterogeneity is said to be isotropic in the plane of bedding, but in general it is considered anisotropic.

Experience gained from field studies shows that the spatial correlation structure of the hydraulic conductivity, K, is in many cases best described by considering its natural log transform $Y = \ln(K)$. Y is modeled through the expected value $<Y>$, and the spatial covariance:

$$C_Y(x, x') = <Y'(x)Y'(x')> \tag{17}$$

where $Y'(x) = Y(x) - <Y>$.
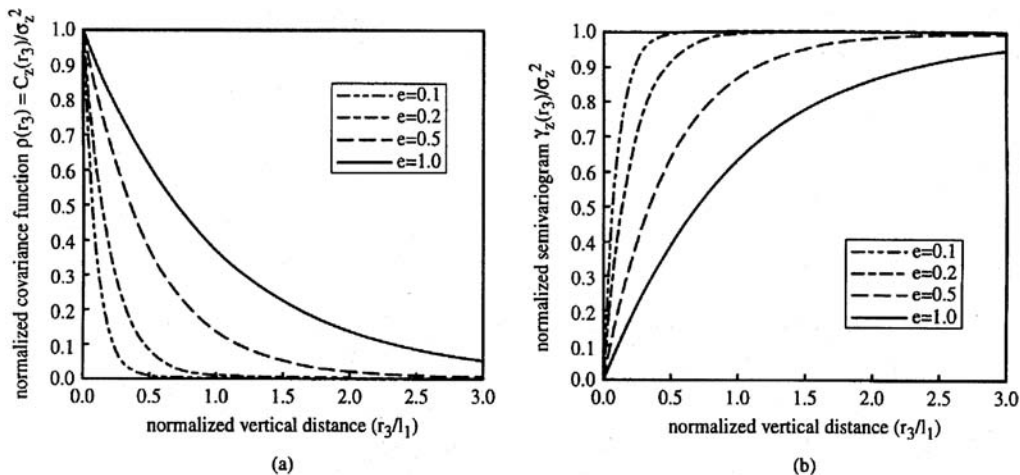
FIGURE 10.20 The exponential covariance function (a) and semivariogram (b) for various degrees of anisotropy.
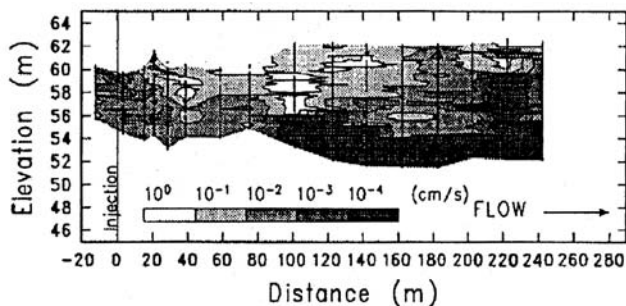
FIGURE 10.21 Vertical cross section of the hydraulic conductivity at a site in Columbus, Mississippi. (From Rehfeldt K. R., Boggs, J. M., and Gelhar, L. W. 1992. Field study in a heterogeneous aquifer 3. Geostatistical analysis of hydraulic conductivity. *Water Resources Research.* 28(12), 3309-3324. With permission.)

To illustrate the process of semivariogram (or covariance) identification, consider Figure 10.21, depicting the block structure at a Columbus, Mississippi site (Rehfeldt et al., 1992). The histogram of the measured Y values is depicted in Figure 10.22, showing that Y can be reasonably modeled with a Gaussian pdf. Figure 10.23 depicts the sampled and fitted semivariograms. The model fitted is axisymmetric and depicts the typical situation where the integral scale in the vertical direction (λ_v on the figure) is much smaller than in the horizontal (λ_h on the figure).

Table 10.7 summarizes site data gathered from field studies as they appeared in Gelhar (1993; the references in the table can be found there in full). Correlation scales in this table correspond to our definition of integral scales (Equation [16]). Table 10.7 is augmented by Table 10.8, which also includes horizontal hydraulic conductivity data. Further data on representative hydraulic conductivity values are given in Gelhar et al. (1992). These values may be borrowed for applications in poorly sampled, geologically similar aquifers, but should be regarded as approximate initial estimates only. The reader is referred to Section 10.4 for additional discussion on representative ranges of hydraulic conductivity values.
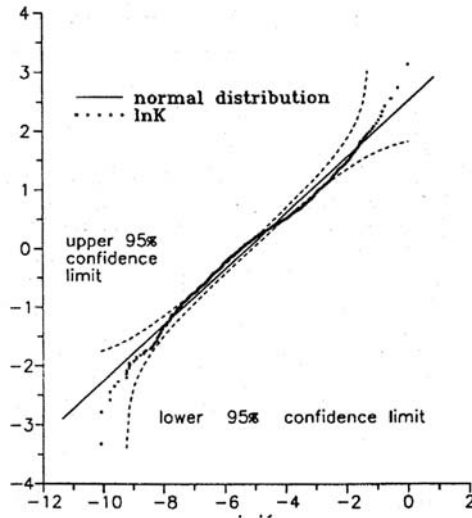
FIGURE 10.22 Normal probability plot of the log conductivity (Y = ln K) data from Columbus, Mississippi. (From Rehfeldt K. R., Boggs, J. M., and Gelhar, L. W. 1992. Field study in a heterogeneous aquifer 3. Geostatistical analysis of hydraulic conductivity. *Water Resources Research.* 28(12), 3309-3324. With permission.)

10.4 Effective Conductivity and Recharge in Heterogeneous Aquifers

10.4.1 Introduction and Definitions

Aquifers are heterogeneous, and hydraulic conductivity may vary greatly over short distances. Usually, it is impossible to map the exact spatial distribution of the hydraulic conductivity because such a task would require an enormous expenditure on data acquisition and in fact could alter the properties of the aquifer due to excessive drilling. The effort is also unnecessary in many applications where an average response is completely sufficient for management and decision making.

To define the effective conductivity, we start from Darcy's law for isotropic media:

$$q_i = -K \frac{\partial H}{\partial x_i}; \quad i = 1, 2, 3 \tag{18}$$

where q_i is the specific flux in the ith direction, K is the isotropic hydraulic conductivity, and H is the hydraulic head. Defining $<q_i>$ and $<\partial H/\partial x_i>$ as the average flux and average head gradient, respectively, the effective hydraulic conductivity is defined by:

$$<q_i> = -K_{ef} <\frac{\partial H}{\partial x_i}> \tag{19}$$

where angled brackets denote the expected value operator. Note that to qualify as an effective property, K_{ef} must be an intrinsic property, that is, it should be a function of the hydraulic properties of the aquifer and not of the head gradient, since in such a case the coefficient in front of the mean head gradient on the right-hand side of Equation (19) has no general applicability.
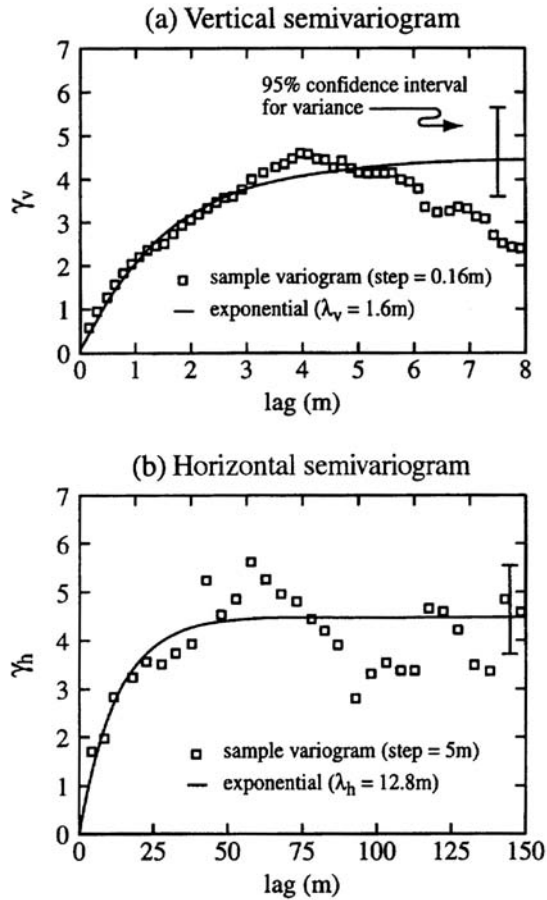
FIGURE 10.23 Vertical (a) and horizontal (b) isotropic semivariograms for the log conductivity (Y = ln K) data from Columbus, Mississippi. (From Rehfeldt K. R., Boggs, J. M., and Gelhar, L. W. 1992. Field study in a heterogeneous aquifer 3. Geostatistical analysis of hydraulic conductivity. *Water Resources Research.* 28(12), 3309-3324. With permission.)

The effective conductivity provides primary statistical information. Furthermore, in cases where the flux is stationary, the expected value <q_i> can be used to estimate the space average of the flux. Computation of the effective hydraulic conductivity in heterogeneous media requires knowledge of the statistical moments of the hydraulic conductivity and in some cases its spatial covariance. For definitions of effective properties in the following discussion, we shall refer to the flow dimensionality, which is equal to the number of space coordinates required to describe the flow field.

10.4.2 Exact Results and Absolute Bounds in Steady, Uniform-in-the-Average Flow

In the case of one-dimensional flow, K_{ef} is the harmonic mean of K: $K_H = (<K^{-1}>)^{-1}$. In the case of lognormal K (in fact for any symmetrical distribution of ln K) and for two-dimensional flow, $K_{ef} = K_G$, where K_G is the geometric mean, i.e., $\ln(K_G) = <\ln K>$. In practice, these relationships imply the form of averaging that needs to be applied to the measured hydraulic conductivities to obtain estimated values of the effective properties.

Additional closed-form results have been obtained by first-order analysis in σ_Y^2, the variance of the log conductivity. For isotropic heterogeneity, the effective conductivity is

TABLE 10.7 Standard Deviation and Correlation Scale of the Natural Logarithm of Hydraulic Conductivity or Transmissivity

Source	Medium	Type[†]	σ_Y	Correlation Scale (m) Horizontal	Vertical	Overall Scale (m) Horizontal	Vertical
Aboufirassi and Marino (1984)	Alluvial-basin aquifer	T	1.22	4000		30,000	
Bakr (1976)	Sandstone aquifer	A	1.5–2.2		0.3–1.0		100
Binsariti (1980)	Alluvial-basin aquifer	T	1.0	800		20,000	
Byers and Stephens (1983)	Fluvial sand	A	0.9	>3	0.1	14	5
Delhomme (1979)	Limestone aquifer	T	2.3	6300		30,000	
Delhomme (pers. communication)							
Aquitane	Sandstone aquifer	T	1.4	17,000		50,000	
Durance	Alluvial aquifer	T	0.6	150		5000	
Kairouan	Alluvial aquifer	T	0.4	1800		25,000	
Normandy	Limestone aquifer	T	2.3	3500		40,000	
Nord	Chalk	T	1.7	7500		80,000	
Devary and Doctor (1982)	Alluvial aquifer	T	0.8	820		5000	
Gelhar et al. (1983, Fig. 6.15)	Fluvial soil	S	1.0	7.6		760	
Goggin et al. (1988)	Aeolian sandstone outcrop	A	0.4	8	3	30	60
Hess (1989)	Glacial outwash sand	A	0.5	5	0.26	20	5
Hoeksema and Kitanidis (1985)	Sandstone aquifer	T	0.6	4.5×10^4		5×10^5	
Hufschmied (1986)	Sand and gravel aquifer	A	1.9	20	0.5	100	20
Loague and Gander (1990)	Prairie soil	S	0.6	8		100	
Luxmoore (1981)	Weathered shale subsoil	S	0.8	<2		14	
Rehfeldt et al. (1989a)	Fluvial sand and gravel aquifer	A	2.1	13	1.5	90	7
Russo and Bressler (1981)	Homra red Mediterranean soil	S	0.4–1.1	14–39		100	
Russo (1984)	Gravelly loamy sand soil	S	0.7	500		1600	
Sisson and Wierenga (1981)	Alluvial silty-clay loam soil	S	0.6	.1		6	
Smith (1978)	Glacial outwash sand and gravel outcrop	A	0.8	5	0.4	30	30
Sudicky (1986)	Glacial-lacustrine sand aquifer	A	0.6	3	0.12	20	2
Viera et al. (1981)	Alluvial soil (Yolo)	S	0.9	15		100	

[†] Types of data: T, transmissivity; S, soils; A, three-dimensional aquifer.

(From Gelhar, L. W. 1993. *Stochastic Subsurface Hydrology*. Prentice Hall, Englewood Cliffs, NJ, p. 390. **All studies are referenced in full at the source.**)

TABLE 10.8 Hydraulic Conductivity Statistics and Spatial Correlations at Selected Field Sites

	Medium	Mean Y or {mean K}[1]	σ_Y^2 or $\{\sigma_K^2\}$[1]	Nugget[2]	Integral Length/Observation Length Horizontal (m)	Vertical (m)	Reference
3D	glaciofluvial aquifer	–7.6 (m/s)	0.37	0.0	1.0–2.5/70	NR[3]	Bjerg et al. (1992)
	alluvial terrace deposit (stationary model)	–5.2 (cm/s)	4.5	0.0	12.8/150	1.6/8	Rehfeldt et al. (1992)
	alluvial terrace deposit (nonstationary model)		2.7		4.8/150	0.8/8	
	fluvial sands	–3.684(cm/s)	0.0263	NR[3]	3/50[(7)]	0.7/1.5	Moltyaner (1986)
	volcanic basin slope	6.02 (cm/d)	1.46	NR[3,4]	2.5/25	0.5/6	Wierenga et al. (1991)
	volcanic ash flow tuff (stationary model)	–4.93 (m/s)	3.61	NR[3]	NR[3]	NR[3]	Istok et al. (1994)
	volcanic ash flow tuff (nonstationary model)		0.6	PN[5]	25/600	0.8/24	
2D	carbonate aquifer	7.48 (m²/d)	1.6	0.0	10,000/200,000		Graham and Neff (1994)
	sandy gravel unit within fluvioglacial deposits	–2.2 (mm/s)	0.025	0.008	2.8/40		Jussel et al. (1994)
	alluvial aquifer[8]	10 (ft²/d)	0.07	0.04	9650/16000		Clifton and Neuman (1982)
		8.9 (ft²/d)	0.14				
	glacial till under no-tillage condition (*in situ* tests)	{1.41} cm/h	{2.359}[6]	{0.663}	60/85		Mohanty et al. (1991)
	glacial till under no-tillage condition (laboratory tests)	{3.15} cm/h	{1.563}[6]	{0.265}	46/85		
	unconsolidated glacial materials	1.14 (m/d)	1.0	NR[3]	15/325		McLaughlin et al. (1993)
	medium grained sand unit within glacial and aeolian deposits	{0.0385} cm/s	{2.97}	PN[5]	NR[3]/10		Healy and Mills (1991)
	sandy loam soil with potato cropping areas (*in situ* tests[9])	{1.09} cm/h	0.57	PN[5]	6/180		Banton (1993)
	sandy loam soil with potato cropping areas (laboratory tests[9])	{10.14} cm/h	2.94				
	alluvial fan (along a transect in the direction of alluvial transport, fine sand layer)	–7.57 (cm/s)	0.8	0.04	45.7/90		Istok et al. (1994)
	alluvial fan (along a transect in the direction of alluvial transport, coarse sand layer)	–5.49 (cm/s)	0.79	0.42	61/90		Istok et al. (1994)
	alluvial fan (along a transect perpendicular to the direction of alluvial transport, fine sand layer)	–6.55 (cm/s)	0.64	0.4	24.4/90		Istok et al. (1994)
	alluvial fan (along a transect perpendicular to the direction of alluvial transport, coarse sand layer)	–6.44 (cm/s)	1.13	PN[5]	1.5/90		Istok et al. (1994)

Notes:

[1] Y –ln K (hydraulic conductivity, length/time) or Y = ln T (transmissivity, length²/time) in the units of K or T shown. Non-logarithmic statistics are bracketed.

[2] Reported variances are total (they include the nugget effect).

[3] Not reported by investigators.

[4] The reader is referred to Jacobson (1990) for more detailed variogram analyses.

[5] Pure nugget model. In this case, the value reported as integral length reflects the shortest separation distance (m) between measurements.

[6] Represents the variogram sill.

[7] Estimated from tracer residence time data.

[8] The aquifer was split into two separate study areas.

[9] This study presents sampling and analytical methods to achieve less disparity between field and laboratory results.

$$K_{ef} = K_G \left[1 + \left(\frac{1}{2} - \frac{1}{m} \right) \sigma_Y^2 \right]; \quad m = 1, 2, 3 \tag{20}$$

where m denotes the space dimensionality of the problem. First-order analysis implies that Equation (20) is applicable for σ_Y^2 smaller than unity. The structure of Equation (20) has led to the conjecture that the exact expression for the effective conductivity is

$$K_{ef} = K_G \exp\left[\left(\frac{1}{2} - \frac{1}{m} \right) \sigma_Y^2 \right]; \quad m = 1, 2, 3 \tag{21}$$

which was shown in Dagan (1993) to be exact for $m = 1$ and $m = 2$ and leads to Equation (20) at first order in σ_Y^2. Equation (21) holds universally for the $m = 1$ case and for a symmetrical probability distribution function (pdf) of $Y = \ln(K)$ for the $m = 2$ case. A recent study by De Wit (1995) has shown that for $m = 3$, the conjecture (Equation [21]) is correct to second order in σ_Y^2 and that higher-order terms also involve the spatial covariance of Y.

For $m = 3$ and unrestricted variability of K, K_{ef} is bounded by the arithmetic (K_A) and harmonic (K_H) means of the conductivity:

$$K_H \leq K_{ef} \leq K_A \tag{22}$$

where $K_A = <K>$. For a lognormal conductivity, Equation (22) becomes:

$$\exp\left[-\frac{\sigma_Y^2}{2} \right] \leq \frac{K_{ef}}{K_G} \leq \exp\left[\frac{\sigma_Y^2}{2} \right] \tag{23}$$

These bounds can be used to compute the bounds on the flux and on the velocity using the relationship $v_i = q_i/n$, where v_i ($i = 1,2,3$) is the velocity in the ith direction and n is the porosity. Equation (22) is valid regardless of the magnitude of σ_Y^2 and the pdf of K, and it allows for bracketing of the effective hydraulic conductivity without any additional assumptions. When too little site-specific data is available for a meaningful calculation of σ_Y^2, the data from Tables 10.7 and 10.8, and from Gelhar (1993) can be used for rough or preliminary estimates.

10.4.3 Closed-Form Results for Small Variance $(\sigma_Y^2 < 1)$ and Anisotropic Correlation

Closed-form results were derived by Dagan (1989) for axisymmetric anisotropy ($I_1 = I_2 \neq I_3$) and Gelhar and Axness (1983) for $I_1 \neq I_2 \neq I_3$ and some simpler cases. For the former, the effective conductivity is a tensor with principal directions parallel and normal to the plane of stratification, and Equation (19) must be generalized as follows:

$$<q_i> = -K_{ef,ij} <\frac{\partial H}{\partial x_j}>; \quad i, j = 1, 2, 3 \tag{24}$$

with summation over repeated indices. Recalling the definition of the ratio of anisotropy $e = I_3/I_1$, and defining the function λ:

$$\lambda = \frac{e^2}{1-e^2}\left[\frac{1}{e\sqrt{1-e^2}}\tan^{-1}\sqrt{\frac{1}{e^2}-1}-1\right] \tag{25}$$

Dagan (1989) obtained:

$$K_{ef,11} = K_{ef,22} = K_G\left[1+\sigma_Y^2\left(\frac{1}{2}-\frac{\lambda}{2}\right)\right] \tag{26}$$

and

$$K_{ef,33} = K_G\left[1+\sigma_Y^2\left(-\frac{1}{2}+\lambda\right)\right] \tag{27}$$

These results do not depend on the actual form of the spatial covariance of the log conductivity, but they do depend on the integral scales.

For three-dimensional anisotropy ($I_1 \neq I_2 \neq I_3$; see Section 10.3 for the definition of I_i) Gelhar and Axness (1983) obtained:

$$K_{ef,ij} = K_G\left[\left(1+\frac{\sigma_Y^2}{2}\right)\delta_{ij}-F_{ij}\right]; \quad i,j=1,2,3 \tag{28}$$

with

$$F_{ij} = \int_{-\infty}^{\infty}\frac{k_i k_j}{k^2}S_Y(k)d^3k \tag{29}$$

where δ_{ij} is the Kronecker delta, $k^2 = k_1^2 + k_2^2 + k_3^2$, $d^3k = dk_1 dk_2 dk_3$, and boldface letters denote vectors. The log conductivity spectra of Equation (29), S_Y, is related to C_Y as follows:

$$S_Y(k_1,k_2,k_3) = \frac{1}{(2\pi)^3}\int_{-\infty}^{\infty}\exp(-i\boldsymbol{k}\cdot\boldsymbol{r})C_Y(\boldsymbol{r})d^3\boldsymbol{r} \tag{30}$$

For isotropic domains, $F_{ij} = \sigma_Y^2/3$ for $i = j$ and zero otherwise, in agreement with Equation (21). Figure 10.24 depicts the functions $g_{ii} = F_{ii}/\sigma_Y^2$ for the more general case of *statistical anisotropy* and for various ratios of anisotropy (in this figure, λ_i correspond to I_i in our text).

10.4.4 The Self-Consistent Approach

The self-consistent approach (for example, Dagan, 1989) is a very powerful technique for the determination of effective properties. It requires knowledge of the pdf of K, the principal directions of the effective conductivity tensor, and the geometry of the blocks of constant K. The idea is to represent the medium as homogeneous and to perturb it repeatedly by single blocks of arbitrary K value and of arbitrary dimensions. That single block perturbs the hydraulic head field and the mean flow, and the self consistency requires that the averages of the perturbed head and flow fields will be equal to the fields obtained by using effective properties.
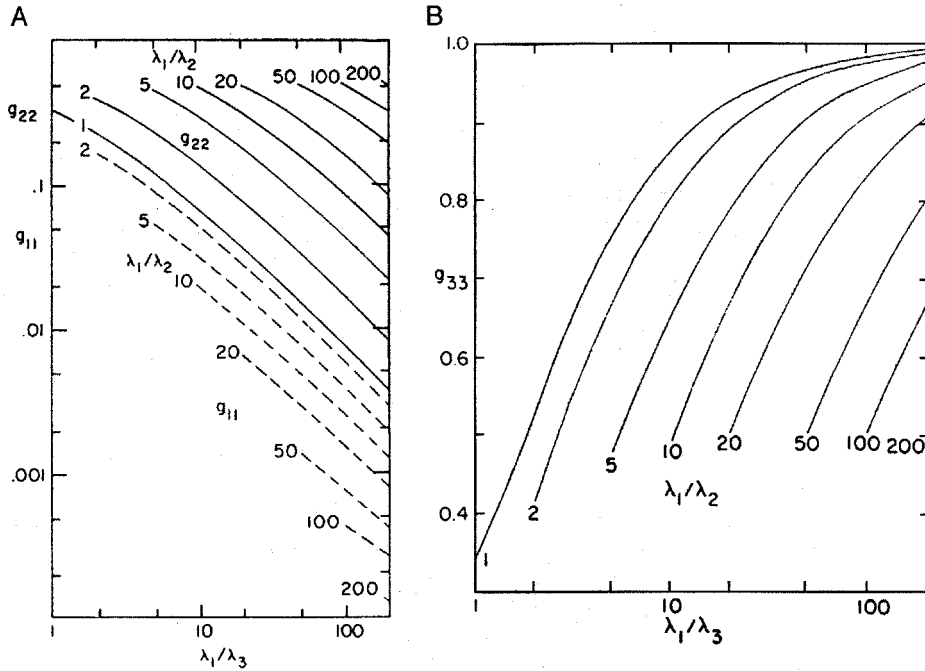
FIGURE 10.24 The g_{ii} functions for the three-dimensional statistically anisotropic media. (From Gelhar L. W., and Axness, C. L. 1983. Three-dimensional stochastic analysis of microdispersion in aquifers. *Water Resources Research*. 19(1), 161-180. With permission.)

In the case of m-dimensional flow ($m > 1$) and isotropic heterogeneity, K_{ef} is also isotropic and is computed as follows:

$$K_{ef} = \frac{1}{m}\left[\int \frac{f(K)dK}{K+(m-1)K_{ef}}\right]^{-1} \tag{31}$$

where $f(K)$ is the pdf of K and integration is over the entire range of variability of K. Equation (31) is a nonlinear integral which can be solved using iterative techniques. If at the N*th* iteration the estimate of the effective conductivity is K_N, then the next estimate, K_{N+1}, is found from:

$$K_{N+1} = \frac{1}{m}\left[\int \frac{f(K)dK}{K+(m-1)K_N}\right]^{-1} \tag{32}$$

This process is repeated until the difference between successive estimates is sufficiently small. The limit of K_N as $N \to \infty$ is K_{ef}.

In the more general case of anisotropic heterogeneity, information is needed regarding the geometry of the blocks. General guidelines are given in Dagan (1989, Part 3). For the quite general case where homogeneous conductivity blocks assume the shape of ellipsoids, one needs to know the anisotropy ratio $e = I_3/I_1$ (the integral scales of the log conductivity; see Section 10.3). Using the definition of λ in Equation (25),
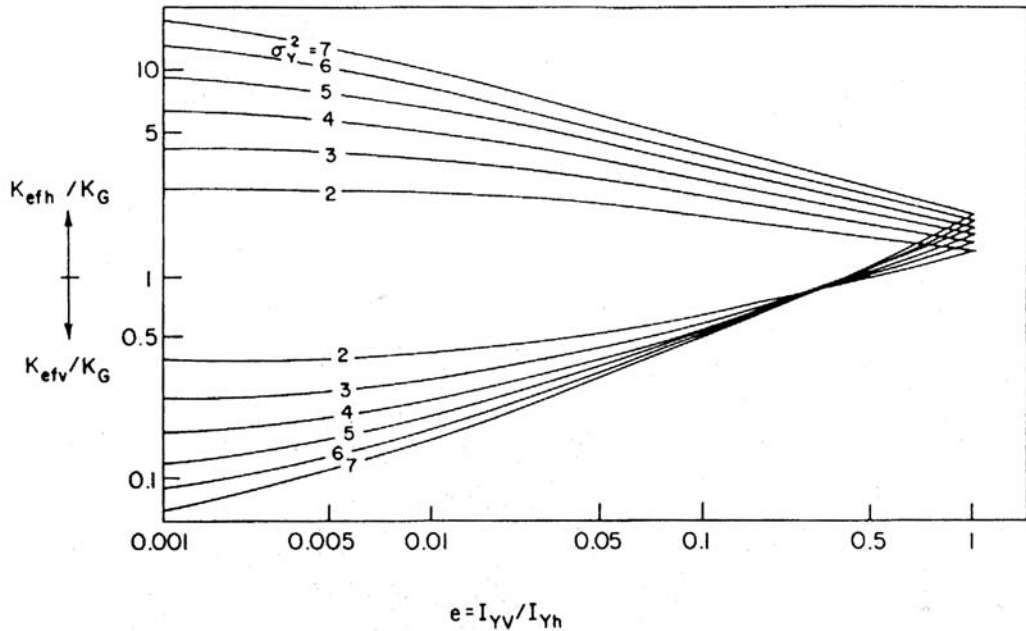
FIGURE 10.25 The dependence of the principal values, horizontal and vertical, of the effective conductivity upon e and upon $\sigma_Y^2 \geq 2$, for a formation of anisotropic axisymmetric heterogeneous structure. The principal directions of C_y are vertical and horizontal, respectively, and the anisotropy ratio e is smaller than unity. The results have been obtained by numerical integration of Equations (33) and (34), i.e., the self-consistent approach. (From Dagan, G. 1989. *Flow and Transport in Porous Formations*. Springer Verlag, Berlin, 465 p. With permission.)

$$K_{ef,11} = \frac{1}{2}\left[\int \frac{f(K)dK}{\left(K - K_{ef,11}\right)\lambda(e) + 2K_{ef}}\right]^{-1} \qquad (33)$$

and an identical expression for $K_{ef,22}$. For the x_3 direction, the self-consistent approach yields:

$$K_{ef,33} = \frac{1}{2}\left[\int \frac{f(K)dK}{K + \left(K_{ef,33} - K\right)\lambda(e)}\right]^{-1} \qquad (34)$$

Figures 10.25 and 10.26 depict some results obtained using the self-consistent approach for various e and σ_Y^2 (in these figures K_{efh} and K_{efv} correspond to K_{ef11} and K_{ef33}, respectively, and I_{Yv} and I_{Yh} correspond to I_3 and I_1, respectively). Figure 10.27 depicts the effective conductivity in isotropic formations as a function of σ_Y^2. Note that the first-order approximation for small σ_Y^2 tends to overestimate the effective conductivity obtained using the self-consistent approach.

A particular case of heterogeneity is that of bimodal (two-phase) formations. Such heterogeneity is encountered in sand-clay or sand-shale formations, or in fractured rocks, where the conductivities of the fractures and of the rock mass differ by orders of magnitude. These cases were investigated by Rubin (1995). Referring to a generic sand-clay mixture, we define V_s and V_c as the sand and clay volumetric fractions, and K_s and K_c as the sand and clay conductivities, respectively. If the medium can be envisioned as homogeneous sand embedded with relatively small cubic inclusions of clay, the self-consistent approach leads to:
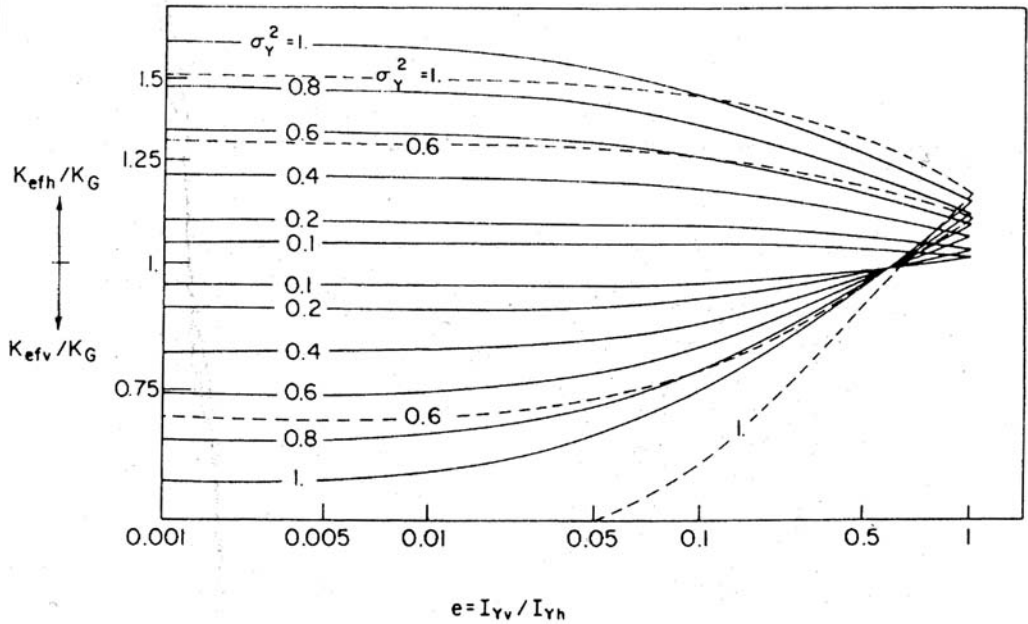
FIGURE 10.26 Same as Figure 10.25 (full line) for $\sigma_Y^2 \leq 1$ and the first-order small σ_Y^2 approximation of Equations (26) and (27) (broken lines). (From Dagan, G. 1989. *Flow and Transport in Porous Formations.* Springer Verlag, Berlin, 465 p. With permission.)
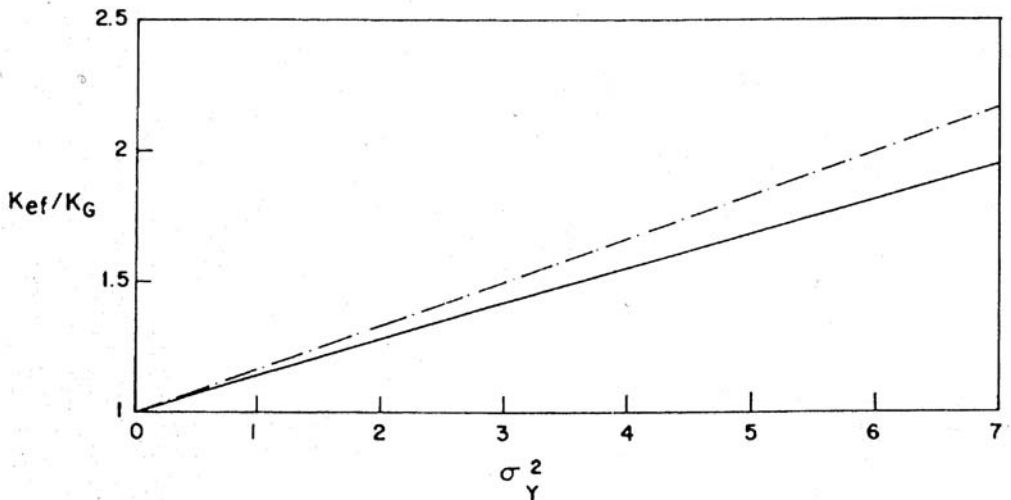
FIGURE 10.27 The dependence of the effective conductivity of a formation of isotropic three-dimensional heterogeneous structure upon σ_Y^2. The numerical solution for the self-consistent approach (Equation [31]; full line) and the first-order small σ_Y^2 approximation $K_{eff}/K_G = 1 + \sigma_Y^2/6$ (broken lines). (From Dagan, G. 1989. *Flow and Transport in Porous Formations.* Springer Verlag, Berlin, 465 p. With permission.)

$$K_{ef} = \frac{1}{3}\left[\frac{V_c}{K_c + 2K_{ef}} + \frac{V_s}{K_s + 2K_{ef}} \right]^{-1} \tag{35}$$

which, in the case of large contrasts between K_s and K_c becomes:

$$\frac{K_{ef}}{K_s} = \frac{1}{2}\left(2 - 3V_c\right) \tag{36}$$

In the case of thin, elongated lenses of clay, the effective conductivity is represented as a tensor, and the self-consistent approach yields the results depicted in Figure 10.28 along with numerical results obtained by Desbarats (1987).
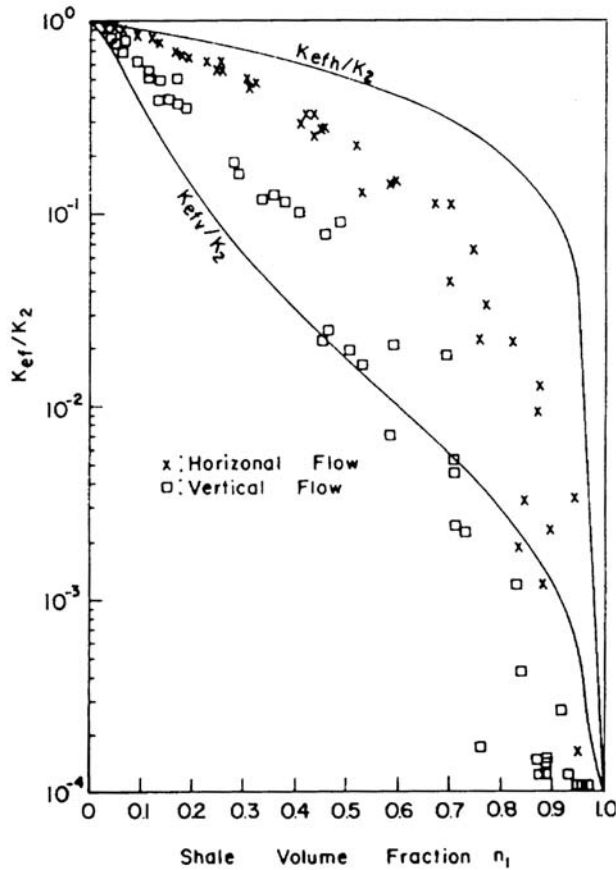
FIGURE 10.28 The effective conductivities (horizontal and vertical) of a two-phase formation made up from a matrix of conductivity K_2 and shale lenses of very low conductivity. The full-line curves are based on the self-consistent approach and the points were obtained by numerical simulations by Desbarats (1987). (From Dagan, G. 1989. *Flow and Transport in Porous Formations.* Springer Verlag, Berlin, 465 p. With permission.)

10.4.5 Effective Conductivity in Nonstationary Formations

The case of a linear trend in the mean log conductivity:

$$\langle Y(\boldsymbol{x}) \rangle = m_0 + \alpha \cdot \boldsymbol{x} \tag{37}$$

was investigated by Rubin and Seong (1994) for two-dimensional, isotropic heterogeneity. Based on a small variability assumption, the effective conductivity was shown to be of the form:
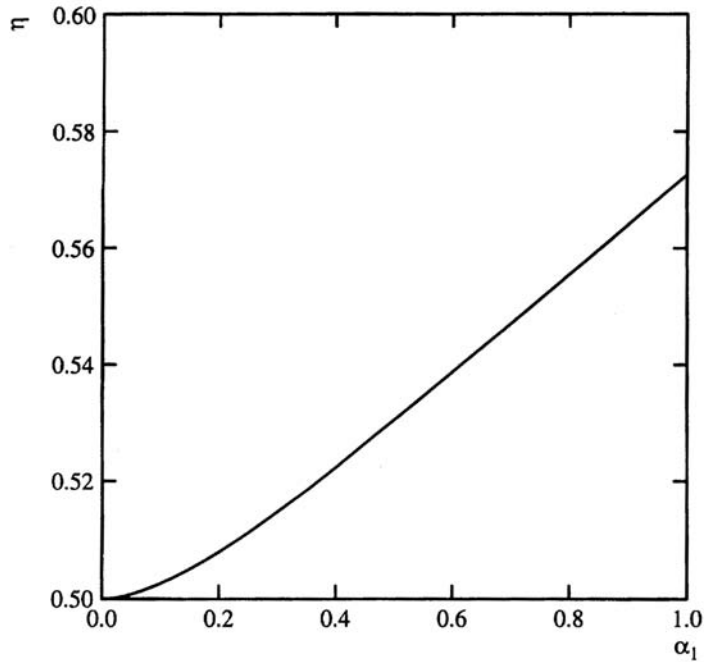
FIGURE 10.29 The coefficient η. (From Rubin, Y. and Seong, K. 1994. Investigation of flow and transport in certain cases of nonstationary conductivity fields. *Water Resources Research*. 30(11), 2901-2912. With permission.)

$$K_{ef}(x) = K_G(x)\left[1 + \sigma_Y^2\left(\frac{1}{2} - A\right)\right] \tag{38}$$

with $K_G(x) = \exp[<Y(x)>]$. A is equal to η (see Figure 10.29) for the case $\boldsymbol{\alpha} = (\alpha_1, 0)$ and to μ (see Figure 10.30) for the case $\boldsymbol{\alpha} = (0, \alpha_2)$. Some results are plotted in Figure 10.31, with $K_{eff} = K_{ef}$.

Based on the earlier conjecture (Equation [21]), the following appears reasonable:

$$K_{ef}(x) = K_G(x)\exp\left[\sigma_Y^2\left(\frac{1}{2} - A\right)\right] \tag{39}$$

Results for two- and three-dimensional, anisotropic heterogeneity are given in Indelman and Rubin (1995, 1996).

10.4.6 Limitations of the Effective Conductivity Approach

The effective hydraulic conductivity approach should generally be used only for statistically stationary domains and for flow which is steady-state and uniform in the average. Because the flow field computed using the effective hydraulic conductivity is the space average of the fluxes of the heterogeneous field, resolution is lost. Thus, the streamlines of the effective field will be smooth and straight, unlike the irregular and tortuous streamlines in the actual flow field. In nonstationary cases such as described in Section 10.4.5, the effective conductivity can only be used for computing the expected value of the fluxes at a given location, but not for computing space averages of the flux.
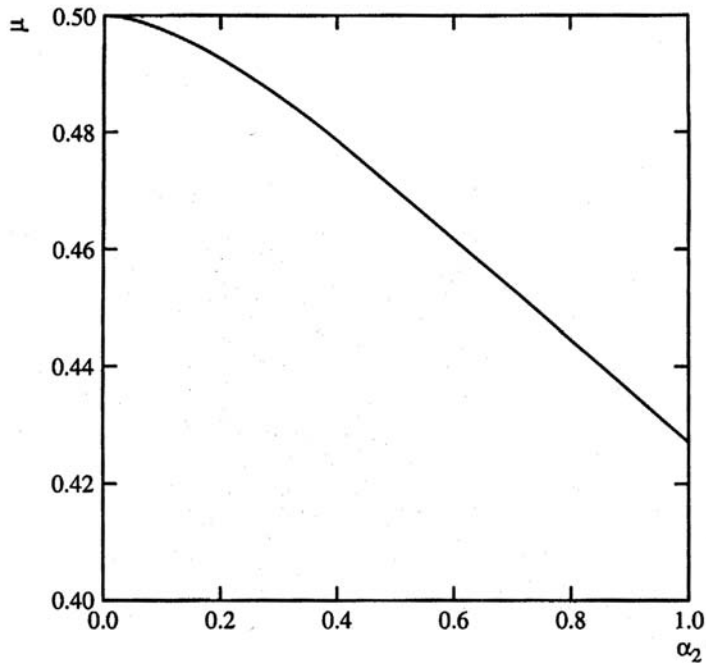
FIGURE 10.30 The coefficient μ. (From Rubin, Y. and Seong, K. 1994. Investigation of flow and transport in certain cases of nonstationary conductivity fields. *Water Resources Research.* 30(11), 2901-2912. With permission.)

10.4.7 Estimation of Natural Recharge

10.4.7.1 Introduction

Natural aquifer recharge is water from precipitation, rivers, lakes, or adjacent aquifers that flows into the groundwater reservoir. In this section, we discuss recharge estimates using mass balance and seepage analyses and piezometric surface geometry. Additional methods such as water budgeting and direct measurement, which can require extensive regional analyses and expensive equipment, are discussed by Ponce (1989) and Lerner et al. (1990). Unsaturated zone modeling is discussed by Stephens (1996), Lerner et al. (1990), and Rushton (1988). Stephens (1996) also addresses the effects of subsurface heterogeneity, the presence of vegetation, and human intervention such as irrigation from agriculture.

Tracers are also used for natural recharge estimates and groundwater dating. Davis and Murphy (1986) and Lerner et al. (1990) give comprehensive reviews, and field and analysis techniques for precipitation recharge estimates from a variety of tracers are widely reported (Ekwurzel et al., 1994; Cook et al., 1994; Landmeyer and Stone, 1995; Solomon et al., 1995).

10.4.7.2 Natural Recharge Estimation by Mass Balance Analysis

Natural recharge (in units of volumetric flux per unit area, or length, L, over time, T) is defined as the inflow to the aquifer resulting from natural infiltration (e.g., precipitation). For a mass balance analysis applied to an arbitrarily defined control volume,

$$R = \frac{\Delta S - \left(V_{\text{in}} - V_{\text{out}}\right) \pm \Sigma}{A_r \, \Delta t} \tag{40}$$

where R is the natural recharge (L/T), A_r is the horizontal area of the control volume (L^2), Δt is the time interval (T), V_{in} and V_{out} (L^3) are the volumes of groundwater flowing into and out of the control volume
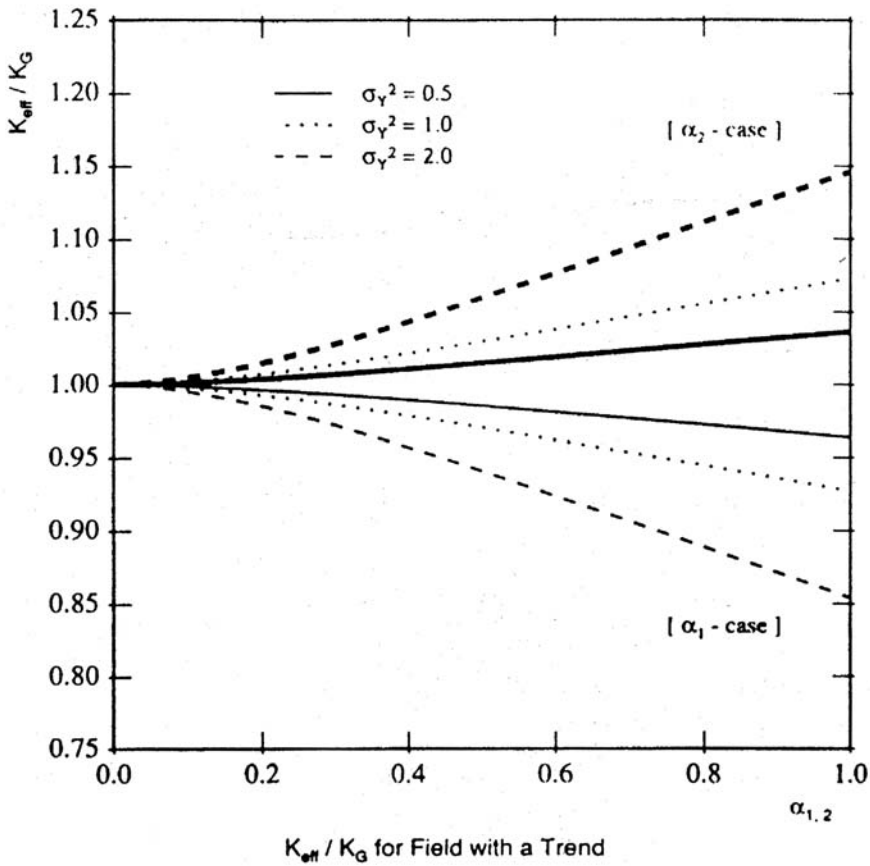
FIGURE 10.31 The ratio $K_{eff}(x)/K_G(x)$ (Equation [39]). (From Rubin, Y. and Seong, K. 1994. Investigation of flow and transport in certain cases of nonstationary conductivity fields. *Water Resources Research*. 30(11), 2901-2912. With permission.)

through the horizontal boundaries during Δt, Σ is the total volume recharged (subtract, $-\Sigma$) or discharged (add, $+\Sigma$) during Δt from other sources, such as pumping wells (L^3), and ΔS is the net change in the volume of water stored in the control volume during the natural recharge event. For control volumes which are large compared to the log conductivity integral scale, the effective hydraulic conductivity and boundary conditions can be used to compute the horizontal fluxes for calculation of V_{in} and V_{out}.

The change in storage, ΔS, is calculated by mapping and contouring water table fluctuations, discretizing the map to an appropriate resolution, determining the change in aquifer volume, and converting to volume of water using

$$\Delta S = S_y \sum_i f_i A_{ri} \tag{41}$$

where S_y is the specific yield (dimensionless) and f_i (L) is the average change in the piezometric surface underlying the horizontal control volume discretized areas A_{ri} (L^2) during Δt. Specific yield, which may be measured in the laboratory (Johnson et al., 1963) or by pumping tests in the field (Fetter, 1994), has typical values 0.22 to 0.27 for coarse sands and gravels, 0.18 to 0.26 for silts and sands, and 0.02 to 0.07 for clays and sandy clays.

An example of the use of Equations (40) and (41) in recharge estimation is provided in Figure 10.32.
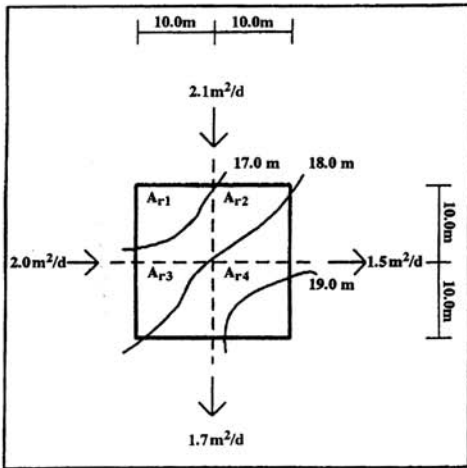
Control volume (S_y=0.2, A_r=400 m²) at time t_1=0 days.
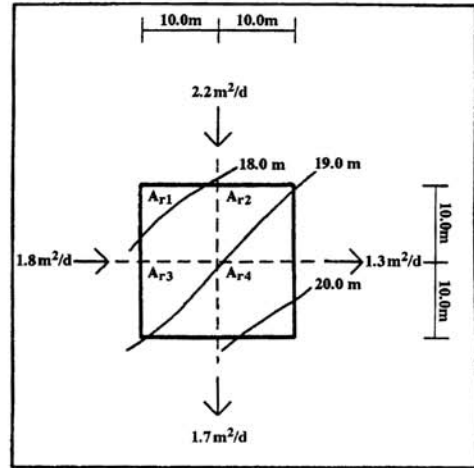
Control volume (S_y=0.2, A_r=400 m²) at time t_2=3 days.

Steps for estimating the natural recharge by mass balance calculation:

1. Calculate V_{in} and V_{out}, the volumes of groundwater flowing into and out of the control volume through the boundaries during the time interval Δt. Since the flux through the boundaries is transient in this example, average over the time period.

$$\Delta t = t_2 - t_1 = 3 \text{ days}$$

$$V_{in} = [\text{(flux in at time } t_2) + \text{(flux in at time } t_1)] / 2 \times \Delta t$$
$$= [(1.8+2.2) \text{ m}^3/\text{d} + (2.0+2.1) \text{ m}^3/\text{d}] / 2 \times 3 \text{ days}$$
$$= 12.15 \text{ m}^3$$

$$V_{out} = [\text{(flux out at time } t_2) + \text{(flux out at time } t_1)] / 2 \times \Delta t$$
$$= [(1.7+1.3) \text{ m}^3/\text{d} + (1.7+1.5) \text{ m}^3/\text{d}] / 2 \times 3 \text{ days}$$
$$= 9.3 \text{ m}^3$$

2. Calculate ΔS, the net change in the volume of water stored in the control volume during the time interval Δt.

$$A_{r1} = A_{r2} = A_{r3} = A_{r4} = 10.0 \text{ m} \times 10.0 \text{ m} = 100.0 \text{ m}^2$$

$$f_1 = 18.0 \text{ m} - 17.0 \text{ m} = 1.0 \text{ m}$$
$$f_2 = 19.0 \text{ m} - 18.0 \text{ m} = 1.0 \text{ m}$$
$$f_3 = 19.0 \text{ m} - 18.0 \text{ m} = 1.0 \text{ m}$$
$$f_4 = 20.0 \text{ m} - 19.0 \text{ m} = 1.0 \text{ m}$$
$$\Delta S = S_y \sum_i f_i A_{ri}$$
$$= (0.2) \times [4 \times (1.0 \text{ m})(100.0 \text{ m}^2)]$$
$$= 80 \text{ m}^3$$

3. Determine Σ, the total volume recharged or discharged during Δt from other sources, such as pumping wells. (In this example, Σ=0.0 m³.)

4. Calculate R, the natural recharge.

$$R = [\Delta S - (V_{in} - V_{out}) \pm \Sigma) / (A_r \Delta t)$$
$$= [80 - (12.15 - 9.3) + 0.0 \text{ m}^3] / (400 \text{ m}^2 \times 3 \text{ days})$$
$$= 0.064 \text{ m/day}$$

FIGURE 10.32 Example calculation of natural recharge by the mass balance method using Equations 40 and 41.

10.4.7.3 Recharge from Rivers and Lakes

Aquifer recharge from rivers may be estimated with a mass balance approach:

$$R = \frac{Q_u - Q_d \pm \Sigma}{A_r} \tag{42}$$

where A_r is the cross-sectional area of the river-aquifer boundary (L^2), Q_u and Q_d are the river flow rates (L^3/T), as measured with stream gages just upstream and downstream of A_r, and Σ accounts for other known volumetric gains (add, $+\Sigma$) or losses (subtract, $-\Sigma$) from the river, such as evaporation. For lake recharge, a network of seepage meters, which measure recharge directly, has been used as an accurate estimation tool (Shaw and Prepas, 1990).

10.4.7.4 Interaquifer Flows

Seepage from over or underlying aquifers can be readily measured using nested piezometers and Darcy's law:

$$R = K \frac{H_1 - H_2}{\Delta L} \tag{43}$$

where H_1 is the hydraulic head of the recharging aquifer (L), H_2 is the hydraulic head of the aquifer being recharged, and K (L/T) and ΔL are, respectively, the hydraulic conductivity and thickness of the inter-vening material. Sharma and Lewis (1994) and Fetter (1994) review field and laboratory hydraulic conductivity testing techniques. Typical values (cm/s) are in the range $10^{-1} - 10$ for gravels, $10^{-3} - 10^{-1}$ for sands, $10^{-6} - 10^{-3}$ for silts and fine sands, and $10^{-9} - 10^{-6}$ for clays.

10.4.7.5 Natural Recharge Estimation Using Piezometric Surface Geometry

In the case of horizontal, steady-state flow with uniformly distributed natural recharge, the linearized, first-order approximation of the flow equation is given by (Rubin and Dagan, 1987a,b; Dagan, 1989)

$$T_G \nabla^2 <H> = -R_{ef} \tag{44}$$

where T_G is the geometric mean of the transmissivity, $<H>$ is the expected value of the hydraulic head, and R_{ef} is the effective recharge. The piezometric surface in the aquifer, or parts of it, can be approximated through the second-order polynomial:

$$<H(x_1, x_2)> = H_0 - J_1 x_1 - J_2 x_2 - \frac{1}{2}\left(c_{11}x_1^2 + 2c_{12}x_1 x_2 + c_{22}x_2^2\right) \tag{45}$$

where J_1, J_2 are the head gradients in the x_1 and x_2 directions, respectively, at $(x_1, x_2) = (0,0)$. Substituting Equation (45) into Equation (44) shows that the quadratic trend coefficients are related to the natural recharge by the relationship:

$$c_{11} + c_{22} = \frac{R_{ef}}{T_G} \tag{46}$$

Since Equation (45) can be viewed as a second-order Taylor expansion around the origin, there can always be defined a vicinity where it is a good approximation, and hence when the head surface has a complex geometry (a polynomial of a higher order than Equation [45] is needed), its analysis can be carried out piecewise. In practical terms, the coefficients c_{11} and c_{22} of the quadratic expression of Equation (45) are estimated based on regression analysis of the field data and then substituted in Equation (46) to yield estimates of the recharge.

For unsteady flow, Equation (44) becomes:

$$T_G \nabla^2 <H> = S \frac{\partial <H>}{\partial t} - R_{ef} \tag{47}$$

where t is time and S is the average storativity. In that case, Equation (46) modifies to the form:

$$c_{11} + c_{22} + S \frac{\partial <H>}{\partial t} = \frac{R_{ef}}{T_G} \tag{48}$$

which requires that the effects of the transients are added to the quadratic coefficients in order to identify the recharge (Dagan and Rubin, 1988). An expansion of the method to account for the effects of pumping wells is included in Dagan and Rubin (1988).

10.5 Stochastic Hydrology and Monte Carlo Simulation Techniques

10.5.1 Introduction

Modeling efforts that require information beyond the mean response of the aquifer involve a more advanced type of analysis than described in Section 10.4. A large body of work exists which describes techniques for treating quantitatively the problems of flow and transport in heterogeneous media under conditions of data scarcity (see Dagan, 1989; Gelhar, 1993).

The concept of stochastic hydrology is summarized in Figure 10.33. The process of aquifer characterization can be seen as that of providing the "correct" input to a transfer function which represents some hydrogeological process. For example, the transfer function may be a numerical simulator of transport, requiring the spatial distributions of the hydraulic conductivity and porosity, and perhaps some chemical reaction parameters, and the output may be the probability distribution function (pdf) of the travel times between a potential contamination source and some regulatory plane of compliance.

The limitations of the characterization process are addressed by replacing the illusive "correct" input by a series of alternative images of the aquifer. Such images should carry in common whatever information is available, and the differences between them should provide a measure of the uncertainty of the input. These images are then processed through the transfer function to yield the corresponding measure of uncertainty in the response function. Analytical transfer functions are available (see Dagan, 1989; and Gelhar, 1993) but are generally limited to small variability and simple aquifer geometry and boundary conditions.

Numerical techniques are more general but computationally demanding. A numerical analysis such as described in Figure 10.33 for processing uncertainty is referred to as a *Monte Carlo simulation*. A partial list of response functions that have been investigated in a Monte Carlo framework is provided in Table 10.9. In a Monte Carlo study, the pdf of the response function is computed. From this pdf, estimates of the expected response with confidence intervals can be obtained. However, the main advantage of the Monte Carlo scheme is that an exhaustive description of the response function is obtained, allowing the investigation and quantification of the occurrence of extreme events.
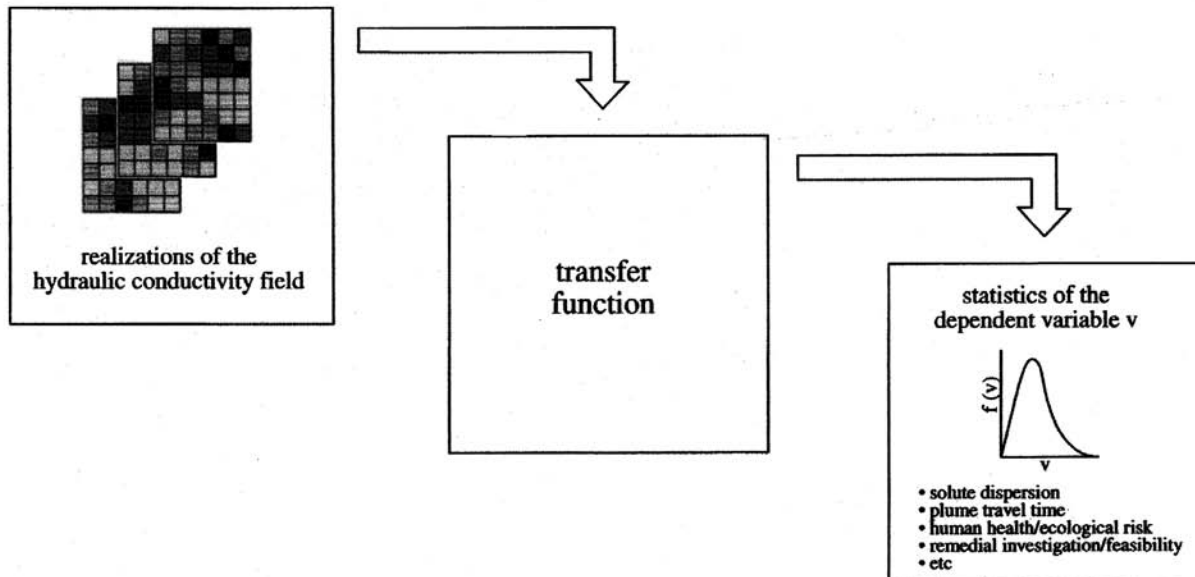
FIGURE 10.33 The Monte Carlo method in aquifer studies.

TABLE 10.9 Environmental Applications of the Monte Carlo Method

	Example References							
	Bellin et al. (1994)	Bellin et al. (1992)	Bosma et al. (1993)	Dakius et al. (1996)	Freeze et al. (1992)	Rubin and Bellin (1994)	Rubin et al. (1994)	Ünlü et al. (1990)
Application or Study								
Passive solute macrodispersion		X				X		
Reactive solute macrodispersion			X					
Plume travel time and breakthrough curves	X	X					X	
Aquifer recharge						X		
Unsaturated flow								X
Risk assessment; investigation and remedial decision analysis				X	X		X	
Type of Monte Carlo Replicate or Distribution								
Heterogeneous flow fields	X	X	X		X	X	X	X
Specific water capacity								X
Pore size								X
Contaminant concentrations				X				

The cornerstone of any Monte Carlo analysis is the numerical simulator of alternative images. To simulate the spatial distribution of a variable $Z(\mathbf{x})$ requires knowledge of its multivariate pdf $f(Z_i)$, $i = 1,\ldots,N$, $N \rightarrow \infty$, with $Z_i = Z(\mathbf{x}_i)$, a task which is generally unattainable given the severe data scarcity commonly encountered in field applications. One exception is the case of multi-Gaussian pdfs, since multi-Gaussian distributions are exhaustively characterized through the first two moments. Many hydrogeological variables have been shown in fact to be Gaussian, and so it is reasonable to assume that they are also multivariate Gaussian. However, it cannot be said that hydrogeological variables are in general Gaussian, and it is not the case that the assumption of Gaussianity is the most conservative in terms of environmental impact assessment. For a review of non-Gaussian models, the reader is referred to Deutsch and Journel (1992).

A review of simulation techniques with an extensive software library is given in Deutsch and Journel (1992). The code HYDRO_GEN, based on the principles for generating multivariate Gaussian fields outlined in Bellin and Rubin (1996), can be obtained free of charge at the web site: http://www. ing.unitn. it/~ bellin/hydro_ge.htm.

10.5.2 Conditional and Unconditional Simulations

The alternative images of the aquifer can be generated with varying degrees of informational constraints. As more information becomes available, more constraints can be imposed, and fewer variations will be observed between alternative images.

At the crudest level, information is available only in the form of statistical moments, which can be obtained from analysis of the available measurements. However, at the initial stages of the characterization effort, the number of measurements may be too small for a meaningful statistical analysis. At this stage the moments can be borrowed from geologically similar yet more extensively sampled aquifers (see Tables 10.7 and 10.8). The images that can be generated at this stage are only required to reproduce the desired moments, and are commonly referred to as unconditional realizations. Due to the minimal constraining level, the generated images will differ greatly. Additionally, the assumptions concerning the multivariate pdf and spatial correlation will be most consequential.

As additional information becomes available, the moments can be evaluated more accurately. Furthermore, the images can be made conditional to the measurements at their locations, meaning that all the generated images will have the measured values as well as the statistical moments in common. Images with this property are referred to as conditional realizations. From a statistical point of view, conditional simulations can be viewed as simulating the conditional pdf $f(Z_i, i = 1,\ldots,N|\{I\})$ over a regular grid, where N is the number of grid nodes and $\{I\}$ is the set of available local information. Bayes' Theorem (Mood et al., 1963) shows that

$$f\left(Z_i, i = 1,\ldots, N\middle|\left\{I\right\}\right) = \frac{f\left(Z_i, i = 1,\ldots, N\left\{I\right\}\right)}{f\left(\left\{I\right\}\right)} \tag{49}$$

where $f(Z_i, i = 1,\ldots,N|\{I\})$ is the joint unconditional pdf of the Z_i and the measurements, and $f(\{i\})$ is the unconditional pdf of the measured values.

In the case where $\{I\}$ represents a group of point measurements $z_p = z(\boldsymbol{x}_p)$, $p = 1,\ldots,P$, conditional simulation amounts to the following three-step procedure:

1. Generate an unconditional realization of Z over the grid, including the measurement points (for example, using the code HYDRO_GEN mentioned in Section 10.5.1);
2. Krige the differences between the measured and generated values (see Chapter 12 of this book for definition of kriging) $z_p - Z(\boldsymbol{x}_p)$, over the same grid; and
3. Add the kriged values to the values generated at the first step.

Once a method has been selected for generating alternative realizations of the random field, the Monte Carlo procedure consists of generating the realizations, solving the transfer function for each of these realizations, and establishing the histograms of the dependent variables of interest. The number of alternative realizations analyzed is determined by the convergence of the statistics of the dependent variable. Higher order statistics or low probability events require a larger number of realizations. Conditioning on measurements can lead to a significant reduction in the required number of realizations.

10.5.3 Upscaling

A central issue in determining the numerical efficiency of a Monte Carlo simulation is the resolution of the randomly generated field. While ideally the field is generated at the highest resolution, for example at the scale of a core sample in three-dimensional simulations, such resolution may pose an unneccessary computational burden, since satisfactory results may be obtained using a coarser numerical grid.

Coarsening of the numerical grid is performed by a procedure called *upscaling*. There is no unique or general upscaling technique; the method is determined by the flow characteristic which is to be preserved when upscaling from a fine to a coarse grid. For example, upscaling can be developed with the goal of preserving the mean flux, or alternatively, the energy dissipation (Indelman and Dagan, 1993). In principle, two different approaches to upscaling can be identified.

In the first approach, the random field is generated at the finest scale of resolution. The domain is then overlain by a coarser grid and the larger block conductivities are obtained by some form of averaging of the fine resolution conductivities falling within the area of each of the larger blocks. Such, for example, is the approach of Durlofsky (1992), which is limited to media displaying periodic variations of the conductivity, and the power law averaging approach of Desbarats (1992), which is more in line with the type of heterogeneity described in Section 10.3. The strength of Desbarats' method is in its simplicity and in its demonstrated success in numerical tests (see Sanchez-Vila et al., 1995, and Desbarats, 1992), but its limitation is that it is empirical.

The second approach consists of upscaling the statistical correlations such that the flow domain can be generated directly at the desired block scale. Such was the technique adopted by Rubin and Gomez-Hernandez (1990) and Sanchez-Vila et al. (1995). The upscaled conductivity in these two methods

preserves the mean flux. In the work of Indelman and Dagan (1993), the upscaled conductivity field preserves the energy dissipation. Numerical testing of the approaches of Desbarats (1992), Rubin and Gomez-Hernandez (1990), and Sanchez-Vila et al. (1995), performed by the last, reveals that these methods yield very similar results and that preservation of the mean fluxes leads also to preservation of energy dissipation. A review of upscaling methods was reported by Wen and Gomez-Hernandez (1996).

10.6 Geochemical Site Characterization

10.6.1 Introduction

An important aspect of site characterization is the identification of portions of an aquifer, basin, or groundwater region which are particularly vulnerable to contaminants from industrial, agricultural, or waste disposal activities. Vulnerability assessment or mapping is a technique for identifying these sub-surface areas based on hydrogeologic factors. Once these regions are identified, precautions can be taken to restrict activities or prevent contamination of the groundwater resources. The techniques are especially useful in analyzing non-point sources of pollution where the entry points to the subsurface environment are not known or cover large areas (e.g., agricultural application of pesticides). The individual regions analyzed are typically on the regional scale (more than a kilometer) but can also be on the field scale (tens to hundreds of meters) if adequate information is available. It is important to note that the indices and factors determined using these assessments are relative values only. Their purpose is to identify sensitive areas, not to indicate that certain groundwater resources are fully resistant to contamination (NRC, 1993).

The simplest and most commonly used vulnerability assessment techniques are overlay and index methods which utilize available soil and chemical property data to assign relative scores to particular regions. This section presents two techniques which have been primarily used to assess the potential for organic contaminants such as pesticides, petroleum products, and solvents to migrate from locations at or near the soil surface to the groundwater table. The first technique, the DRASTIC method, was developed by the U.S. EPA (Environmental Protection Agency) to provide a means, using only readily available site data, to assess systematically the potential for pollutants to migrate to groundwater. The second technique uses chemical-specific information to assign attenuation factors (AF) to areas within the study region. Additional information and descriptions of these and other techniques are provided in NRC (1993) and Knox et al. (1993).

10.6.2 The DRASTIC Method

Aller et al. (1987) developed a vulnerability assessment technique which relies on readily available information for a site, aquifer, basin, or region. Utilizing the hydrogeological setting, inferences are made as to the soil's geochemical nature and the potential for contaminants to migrate from the soil surface to the groundwater table. The term *hydrogeological setting* refers to a composite description of all the major geologic and hydrologic factors. These types of analyses apply to areas on the order of 100 acres or more.

The DRASTIC method is classified as an overlay technique and provides a simple and straightforward means of assessing the susceptibility of certain areas to pollutants. The acronym DRASTIC refers to the seven factors utilized in the rating system — depth to groundwater, recharge rate (net), aquifer media, soil media, topography, impact on vadose zone, and conductivity (hydraulic). Each of these is assigned a value based on a rating system developed by Aller et al. (1987). These factors are adjusted by a weighting factor and summed to calculate the *pollution potential* or DRASTIC index:

$$D_r D_w + R_r R_w + A_r A_w + S_r S_w + T_r T_w + I_r I_w + C_r C_w = \begin{array}{c} \text{pollution potential} \\ \text{or} \\ \text{DRASTIC index} \end{array} \qquad (50)$$

where the subscripts r and w refer to the rating and weighting factors, respectively. Table 10.10 lists these rating and weighting factors. Different weighting factors are listed for pesticides because these are typically less volatile and more persistent in the environment, and the weighting is shifted away from the impact of the vadose zone (I) and hydraulic conductivity (C) to the soil media (S) and topography (T). All other compounds are classified together as non-pesticides. Except for this distinction in weighting factors, no other chemical-specific parameters are utilized for this technique. For this reason, comparisons can only be made between each area's susceptibility to the same compound or to pesticides and non-pesticides in general, not between individual compounds.

To utilize the DRASTIC method, the region is first divided into subregions based on the available hydrogeologic data. Next, using Equation (50) and Table 10.10, the pollution potential or DRASTIC index for each subregion is calculated. These values are then used to generate a map indicating which regions are more susceptible than others. The site characterization information required can be directly determined using the techniques discussed in the sections of this and other chapters or by utilizing available data. Additional sources of data include the Soil Conservation Service (SCS), the U.S. Geological Survey for topographic and geologic maps, and regional and local land-use planning maps.

Interpretation of DRASTIC indices and maps is performed on a relative scale. The higher the calculated index, the more probable that a contaminant will migrate through the vadose zone to the water table. Because these values are relative expressions, indices for a number of areas within a region must be assessed to determine each area's relative susceptibility. An additional comparison can be made by comparing results with those typical for that region of the country. Fifteen groundwater regions of the U.S. have been designated by Heath (1984). Aller et al. (1987) analyzed four to sixteen different types of hydrogeological settings within thirteen of these regions to determine typical DRASTIC indexes for the U.S. Table 10.11 lists the range of indices determined for each region based on a compilation of Aller's results. The lowest indices were typically found for hydrogeological settings which can be classified as mountain slopes, and the highest values were for swamps or marshes and beaches. These values provide an additional relative benchmark from which to evaluate the DRASTIC indices generated for areas within a particular region.

TABLE 10.10 Rating and Weighting Factors for the DRASTIC Method

Depth to Water Table (D)	
Depth (feet)	Rating D_r
0–5	10
5–15	9
15–30	7
30–50	5
50–75	3
75–100	2
> 100	1
Type	Weighting D_w
Non-Pesticide	5
Pesticide	5

TABLE 10.10 (continued)

Net Recharge (R)

Recharge (inches)	Rating R_r
0–2	1
2–4	3
4–7	6
7–10	8
> 10	9

Type	Weighting R_w
Non-Pesticide	4
Pesticide	4

Aquifer Media (A)

Aquifer Types	Rating Range A_r	Typical Rating A_r
Massive Shale	1–3	2
Metamorphic/Igneous	2–5	3
Weathered Metamorphic/Igneous	3–5	4
Glacial Till	4–6	5
Bedded Sandstone, Limestone, and Shale Sequences	5–9	6
Massive Sandstone	4–9	6
Massive Limestone	4–9	6
Sand and Gravel	4–9	8
Basalt	2–10	9
Karst Limestone	9–10	10

Type	Weighting A_w
Non-Pesticide	3
Pesticide	3

Soil Media (S)

Media Type	Rating S_r
Thin or Absent	10
Gravel	10
Sand	9
Peat	8
Shrinking and/or Aggregated Clay	7
Sandy Loam	6
Loam	5
Silty Loam	4
Clay Loam	3
Muck	2
Nonshrinking and Nonaggregated Clay	1

Type	Weighting S_w
Non-Pesticide	2
Pesticide	5

TABLE 10.10 (continued)

Topography (T)	
Slope (%)	Rating T_r
0–2	10
2–6	9
6–12	5
12–18	3
> 18	1

Type	Weighting T_w
Non-Pesticide	1
Pesticide	3

Impact of the Vadose Zone (I)		
Media Type	Rating Range I_r	Typical Rating I_r
Confining Layer	1	1
Silt/Clay	2–6	3
Shale	2–5	3
Limestone	2–7	6
Sandstone	4–8	6
Bedded Limestone, Sandstone, and Shale	4–8	6
Sand and Gravel with significant Silt and Clay	4–8	6
Metamorphic/Igneous	2–8	4
Sand and Gravel	6–9	8
Basalt	2–10	9
Karst Limestone	8–10	10

Type	Weighting I_w
Non-Pesticide	5
Pesticide	4

Hydraulic Conductivity (C)	
Hydraulic Cond. (gpd/ft^2)	Rating C_r
1–100	1
100–300	2
300–700	4
700–1000	6
1000–2000	8
> 2000	10

Type	Weighting C_w
Non-Pesticide	3
Pesticide	2

(Adapted from Aller, L., Bennett, T., Lehr, J., Petty, R., and Hackett, G. 1987. *DRASTIC: A Standardized System for Evaluating Ground Water Pollution Potential Using Hydrogeological Settings.* U.S. Environmental Protection Agency, RSKERL, Ada, OK. EPA/600/2-87/035.)

TABLE 10.11 Range of DRASTIC Indexes for Groundwater Regions in the U.S.

Region Number	Region Name	Range of DRASTIC Indexes	Range of Pesticide DRASTIC Indexes
1	Western Mountain Ranges	65–196	91–221
2	Alluvial Basins	74–202	99–224
3	Columbia Lava Plateau	86–179	92–208
4	Colorado Plateau and Wyoming Basin	87–176	108–213
5	High Plains	80–196	88–229
6	Nonglaciated Central Region	71–196	96–221
7	Glaciated Central Region	88–202	111–225
8	Piedmont and Blue Ridge	70–176	102–213
9	Northeast and Superior Uplands	75–190	102–198
10	Atlantic and Gulf Coastal Plain	82–202	113–233
11	Southeast Coastal Plain	190–224	224–251
12	Alluvial Valleys	NA	NA
13	Hawaiian Islands	165–201	174–230
14	Alaska	92–141	118–166
15	Puerto Rico and Virgin Islands	NA	NA

NA—Not analyzed by Aller et al. (1987).

(Data from Aller, L., Bennett, T., Lehr, J., Petty, R., and Hackett, G. 1987. *DRASTIC: A Standardized System for Evaluating Ground Water Pollution Potential Using Hydrogeological Settings.* U.S. Environmental Protection Agency, RSKERL, Ada, OK. EPA/600/2-87/035.)

10.6.3 Attenuation Factors

Rao et al. (1985) developed an alternate vulnerability assessment technique which accounts for the chemical properties of the compound(s) of interest. The technique utilizes an attenuation factor (AF) as a quantitative index for ranking contaminants based on their potential for migrating from the soil surface to the water table. This factor represents the fraction of the contaminant remaining after it has traveled through the vadose zone to the water table. The AF accounts for the primary processes which control the transport of contaminants — sorption, advection, and transformation, and it ranges in value from 0 to 1. The equations for calculating the AF are listed in Figure 10.34. This approach has been used by Khan and Liang (1989) and Loague et al. (1990) to assess the vulnerability of aquifers and basins to different pesticides.

The processes of sorption and volatilization of the compound are accounted for in the AF through the retardation factor (RF). The RF value ranges from 1 to infinity, where RF = 1 represents a nonreactive, nonvolatile tracer compound. The product $f_{oc}K_{oc}$ in the expression for RF is the contaminant distribution coefficient which represents its equilibrium partitioning between the soil and the water. The premise for this relationship is that a majority of the soil's sorption capacity for organic compounds such as solvents, petroleum products, and pesticides can be attributed to the organic carbon within the soil. If the f_{oc} for the soil is low (< 0.005) then this relationship might not be appropriate. Table 10.12 lists the K_{oc} for some common contaminants of environmental interest.

Values for K_H are also readily available in the literature (e.g., Lyman et al., 1982; U.S. EPA, 1986; Mackay et al., 1992; Knox et al., 1993) and several are listed in Table 10.12. K_H constants are reported in the literature in a variety of units; the values used here are in a dimensionless form (for relating air and water concentrations), and constants from the literature should be adjusted appropriately. For nonvolatile compounds, the last term in the expression for RF can be discarded.

The expression for AF also accounts for transformations, which can include chemical or biological degradation as well as radioactive decay. Using a first-order relationship, the half-life ($t_{1/2}$) is related to a first-order degradation rate coefficient ($t_{1/2} = 0.693/k$). Literature data for $t_{1/2}$ is limited but ranges of values have been reported by Howard et al. (1991) and Mackay et al. (1992). Values for several common contaminants are included in Table 10.12. Local conditions such as soil pH, temperature, water content, and microbial populations can substantially alter these rates. If available, site-specific values based on

The attenuation factor (*AF*) is given by:

$$AF = e^{-\frac{0.693\, d\, RF\, \theta_{FC}}{q\, t_{1/2}}}$$

d = distance from the soil surface to the water table [L]
θ_{FC} = soil-water content at field capacity (volume fraction)
q = net annual groundwater recharge [L/t]
$t_{1/2}$ = transformation half-life [t] = $0.693/k$
k = first-order degradation rate coefficient [1/t]

The retardation factor is given by:

$$RF = 1 + \frac{\rho_b f_{oc} K_{oc}}{\theta_{FC}} + \frac{n_a K_H}{\theta_{FC}}$$

ρ_b = soil bulk density [M/L^3]
f_{oc} = mass fraction of organic carbon in the soil
K_{oc} = organic carbon and water partitioning coefficient [L^3/M]
n_a = soil air-filled porosity $(n_a = n - \theta_{FC})$
n = soil porosity [$n = 1 - (\rho_b/\rho_p)$]
ρ_p = soil particle density [M/L^3]
K_H = dimensionless Henry's constant

FIGURE 10.34 Equations for the attenuation factor (AF) and retardation factor (RF).

TABLE 10.12 Chemical Characteristics for Selected Organic Compounds

Compound	K_{oc}^{a} (mL/g)	K_H^{c}	$t_{1/2}^{d}$ (days)
Benzene	83	0.24	5–16
Toluene	300	0.28	4–22
Benzo(a)pyrene	5.5×10^6	6×10^{-5a}	57–530
Trichloroethylene (TCE)	126	0.42	180–365
Atrazine	160[b]	2.5×10^{-7b}	70[b]
DDT	2.4×10^5	0.0017	730–5700
Ethylene dibromide (EDB)	44	0.028	28–180

[a] U.S. EPA (1986)
[b] Jury et al. (1984)
[c] Lyman et al. (1982)
[d] Howard et al. (1991)

laboratory tests with the site soils and the chemical of interest are preferred. This technique does not account for transformation products which can have longer half-lives and also be of environmental concern.

The remaining parameters are soil-specific and can be determined based on available site measurements or, for large areas, estimated by utilizing databases for soil properties. These estimates can be based on the soil order classification of each subregion of the study area. SCS (1988) has identified ten soil orders — Alfisols, Aridisols, Entisols, Histosols, Inceptisols, Mollisols, Oxisols, Spodosols, Ultisols, and Vertisols. The SCS maintains a computerized database for a large number of soil properties (including those listed in Figure 10.34) for many classifications and multiple subclassifications of soils for different regions throughout the U.S. Soil properties can be estimated based on values in the database for that region. For example, Loague et al. (1990) estimated soil properties based on the soil order classification represented by each area in their study region (the Pearl Harbor basin in Oahu). Their selected soil properties were based on a database of soils for all the islands in the state of Hawaii. If no other information is available, typical or representative values should be used. For example, ρ_p is usually 2.65 g/cm^3, and n is in the range of 0.24 to 0.7 for unconsolidated deposits and 0.05 to 0.5 for rock formations (Freeze and Cherry, 1979). For the fraction of organic carbon (f_{oc}), the ranges for eight of the soil orders are as follows (based on a review of data in SCS, 1988):

Soil Order	f_{oc} (%)
Alfisols	<3.0
Aridisols	<1.0
Entisols	<1.5
Inceptisols	<12.0
Mollisols	<2.0
Oxisols	<4.0
Spodosols	<5.0
Vertisols	<2.0

These are merely typical values for representative soils; actual site values could vary substantially.

Table 10.13 provides sample results for several compounds for two hypothetical soil types with different depths to the water table. An additional scenario is included with a higher net recharge rate (equivalent to an irrigated site). The rankings indicate the relative likelihood of a compound reaching the water table (with 1 being the most likely). The soil properties are listed and $\rho_p = 2.65$ g/cm^3 is used for all scenarios. The chemical properties are taken from Table 10.12, and the highest values are used for $t_{1/2}$ for a conservative comparison. For the scenarios shown, the top rankings shift under different hydrogeological conditions and the AF increases with the higher recharge rate. These factors are highly dependent on $t_{1/2}$; as $t_{1/2}$ approaches infinity, AF approaches unity regardless of the hydrogeological conditions. These factors should be used for comparative ranking purposes only.

Once the AF has been calculated for each area within the study region, maps of the AF values can be overlain on site maps. Assigning qualitative values to ranges of AF can provide a more convenient means for interpreting attenuation maps. Table 10.14 lists the designations used by Khan and Liang (1989); using these designations, they mapped the vulnerability of basins on the island of Oahu to various pesticides. Based on their analysis, land-use planners and regulators could identify which areas were of greatest concern and the regions where groundwater contamination was most likely.

Using either of these two vulnerability assessment techniques can provide additional information regarding the characteristics of a site or subregions within an aquifer or basin. Those subsurface areas which are more susceptible to the migration of contaminants can be identified using readily available site or regional information. Additional information on these vulnerability assessment techniques can be found in Aller et al. (1987), Loague et al. (1990), Knox et al. (1993), and NRC (1993).

TABLE 10.13 Sample Attenuation Factors (AF)

	Soil #1			Soil #2			Soil #1 w/irrigation		
	q = 350 mm/yr d = 1 m f_{oc} = 0.01 ρ_b = 1.1 g/cm^3 θ_{FC} = 0.4			q = 350 mm/yr d = 5 m f_{oc} = 0.0075 ρ_b = 1.4 g/cm^3 θ_{FC} = 0.45			q = 1200 mm/yr d = 1 m f_{oc} = 0.01 ρ_b = 1.1 g/cm^3 θ_{FC} = 0.4		
	ln(AF)	AF	Rank	ln(AF)	AF	Rank	ln(AF)	AF	Rank
Non-Pesticides:									
Benzene	−61.3	$10^{-26.6}$	3	−300	~0	3	na	na	na
Trichlorethylene (TCE)	−3.69	0.025	2	−17.6	$10^{-7.7}$	1	na	na	na
Pesticides:									
DDT	−335	~0	4	−1598	~0	4	−97.6	10^{-42}	2
Ethylene dibromide (EDB)	−3.57	0.028	1	−18.3	10^{-8}	2	−1.04	0.35	1

na — not applicable: only the pesticides are included in the irrigation scenario.

TABLE 10.14 Qualitative Designations for Attenuation Factors (AF)

Attenuation Factor (AF)	Probability of Groundwater Contamination
$0–10^{-4}$	Very unlikely
10^{-4}–10^{-2}	Unlikely
10^{-2}–0.1	Moderately likely
0.1–0.25	Likely
0.25–1.0	Very likely

(As defined by Khan, M. A. and Liang, T. 1989. Mapping pesticide contamination potential. *Environmental Management.* 13(2), 233-242.)

Acknowledgment

References

Aller, L., Bennett, T., Lehr, J., Petty, R., and Hackett, G. 1987. *DRASTIC: A Standardized System for Evaluating Ground Water Pollution Potential Using Hydrogeological Settings.* U.S. Environmental Protection Agency, RSKERL, Ada, OK. EPA/600/2-87/035.

Annan, A. P. and Cosway, S. 1992. Ground penetrating radar survey design. *Proc. of the Symposium on the Applications of Geophysics to Engineering and Environmental Problems (SEGEEP).* Chicago.

Archie, G. E. 1942. The electric resistivity log as an aid in determining some reservoir characteristics. *Trans. AIME.* 146, 54-62.

Banton, O. 1993. Field- and laboratory-determined hydraulic conductivities considering anisotropy and core surface area. *Soil Science Society of America Journal.* 57(1), 10-15.

Bellin A. and Rubin, Y. 1996. HYDRO_GEN: A spatially distributed random field generator for correlated properties. *Stochastic Hydrology and Hydraulics.* 10(4), 253-278.

Bellin, A., Rubin, Y., and Rinaldo, A. 1994. Eulerian-Lagrangian approach for modeling of flow and transport in heterogeneous geological formations. *Water Resources Research.* 30(11), 2913-2924.

Bellin, A., Salandin, P., and Rinaldo, A. 1992. Simulation of dispersion in heterogeneous porous formations: statistics, first-order theories, convergence of computations. *Water Resources Research.* 28(9), 2211-2227.

Beres, M., Jr. and Haeni, F. P. 1991. Application of ground-penetrating radar methods in hydrogeologic studies. *Ground Water.* 29(3), 375-386.

Best, M. E. 1992. *Geological Association of Canada Short Course Notes Volume 10.* Wolfville, Nova Scotia, May 28-29.

Bjerg, P. L., Hinsby, K., Christensen, T. H., and Gravesen, P. 1992. Spatial variability of hydraulic conductivity of an unconfined sandy aquifer determined by a mini slug test. *Journal of Hydrology.* 136, 107-122.

Bosma, W. J. P., Bellin, A., van der Zee, S., and Rinaldo, A. 1993. Linear equilibrium adsorbing solute transport in physically and chemically heterogeneous porous formations, 2. Numerical results. *Water Resources Research.* 29(12), 4031-4043.

Brewster, M. L. and Annan, A. P. 1994. Ground-penetrating monitoring of a controlled DNAPL release: 200 MHZ radar. *Geophysics.* 59(8): 1211-1221.

Burger, H. R. 1992. *Exploration Geophysics of the Shallow Subsurface.* Prentice Hall, Englewood Cliffs, NJ.

Buselli, G., Barber, C., Davis, G. B., and Salama, R. B. 1990. Detection of groundwater contamination near waste disposal sites with transient electromagnetic and electrical methods, in *Geotechnical and Environmental Geophysics Vol. 2: Environmental and Groundwater, S.E.G. Investigations in Geophysics 5.* Ed. Stanley Ward. 27-39.

Butler, D. K. 1991. Tutorial — engineering and environmental applications of microgravity. *Proc. of the Symposium on the Application of Geophysics to Engineering and Environmental Problems.* Knoxville, TN: 139-177.

Clifton, P. M. and Neuman, S. P. 1982. Effects of kriging and inverse modeling on conditional simulation of the Avra Valley aquifer in southern Arizona. *Water Resources Research.* 18(4), 1215-1234.

Cook, et al. 1994. Unsaturated zone tritium and chlorine 36 profiles from souther Australia: their use as a tracers of soil water movement. *Water Resources Research.* 30(6), 1709-1719.

Dagan, G. and Rubin, Y. 1988. Stochastic identification of recharge, transmissivity and storativity in aquifer's unsteady flow: a quasi-steady approach. *Water Resources Research.* 24(10), 1698-1710.

Dagan, G. 1993. Higher-order correction of effective permeability of heterogeneous isotropic formations of lognormal conductivity distribution. *Transport in Porous Media.* 12, 279-290.

Dagan, G. 1989. *Flow and Transport in Porous Formations.* Springer Verlag, Berlin, 465 p.

Daily, W., and Ramirez, A. 1995. Electrical resistance tomography during in-situ trichloroethylene remediation at the Savannah River site. *Journal of Applied Geophysics.* 33: 239-249.

Dakius, M. E., Toll, J. E., Small, M. J., and Brand, K. P. 1996. Risk-based environmental remediation: Bayesian Monte Carlo analysis and the expected value of sample information. *Risk Analysis.* 16, 67-79.

Daniels, D., Fritz, S., and Leap, D. 1991. Estimating recharge rates through unsaturated glacial till by tritium dating. *Ground Water.* 29(1), 26-24.

Davis, J. and Annan, A. 1989. Ground-penetrating radar for high-resolution mapping of soil and rock stratigraphy. *Geophysical Prospecting.* 37: 531-551.

Davis, S. and Murphy, E. 1986. Dating ground water and the evaluation of repositories for radioactive waste. U.S. Regulatory Commission Report No. NUREG/CR-4912.

Desbarats, A. 1987. Numerical estimation of effective permeability in sand-shale formations. *Water Resources Research.* 23, 273-286.

Desbarats, A. 1992. Spatial averaging of hydraulic conductivity in three-dimensional heterogeneous porous media. *Mathematical Geology.* 24(3), 249-267.

Deutsch, C. V. and Journel, A. G. 1992. GSLIB: A geostatistical software library and user's guide. Oxford University Press, London.

De Wit, A. 1995. Correlation structure dependence of the effective permeability of heterogeneous porous media. *Phys. Fluids.* 7(11), 1553-2562.

Durlofsky, L. 1992. Representation of grid block permeability in coarse scale models of randomly heterogeneous porous media. *Water Resources Research.* 28(7), 1791-1800.

Ekwurzel, B. et al. 1994. Dating of shallow groundwater: comparison of the transient tracers ^{3}H/^{3}He, chlorofluorocarbons, and ^{85}Kr. *Water Resources Research.* 30(6), 1693-1708.

Fetter, C. W. 1994. *Applied Hydrogeology.* Macmillan College Publishing Co., New York.

Fisher, E., McMechan, G., and Annan, A. P. 1992. Acquisition and processing of wide-aperture ground-penetrating radar data. *Geophysics.* 57(3): 495-504.

Freeze, R. A. and Cherry, J. A. 1979. *Groundwater.* Prentice-Hall, Inc., Englewood Cliffs, NJ.

Freeze, R.A., James, B., Massmann, J., Sperling, T., and Smith, L. 1992. Hydrogeological decision analysis, 4., The concept of data worth and its use in the development of site investigation strategies. *Ground Water.* 30(4), 574-588.

Frischknecht, F. C., Labson, V. F., Speis, B. R., and Anderson, W. L. 1991. Profiling methods using small sources, in *Electromagnetic Methods in Applied Geophysics. Vol. 2, S.E.G. Investigations in Geophysics 3.* Ed. M. Nabighian. 105-270.

Gelhar, L. W. and Axness, C. L. 1983. Three-dimensional stochastic analysis of microdispersion in aquifers. *Water Resources Research.* 19(1), 161-180.

Gelhar L. W., Welty, C., and Rehfeldt, L. 1992. A critical review of data on field-scale dispersion in aquifers. *Water Resources Research.* 28(7), 1955-1974.

Gelhar, L. W. 1993. *Stochastic Subsurface Hydrology.* Prentice Hall, Englewood Cliffs, NJ, 390.

Geller, J. T. and Myer, L. R. 1995. Ultrasonic imaging of organic liquid contaminants in unconsolidated porous media. *Journal of Contaminant Hydrology.* 19, 85-104.

Goldstein, N. E. 1994. Expedited site characterization geophysics: geophysical methods and tools for site characterization. *Lawrence Berkeley Laboratory LBL-35384.*

Graham, W. D. and Neff, C. R. 1994. Optimal estimation of spatially variable recharge and transmissivity under steady-state groundwater flow, 2, case study. *Journal of Hydrology.* 157, 267-285.

Greaves, R. J., Lesmes, D. P., Le, J. M., and Toksoz, M. N. 1996. Velocity variations and water content estimated from multi-offset, ground-penetrating radar. *Geophysics.* 61(3), 683-695.

Greenhouse, J., Brewster, M., Schneider, G., Redman, J. D., Annan, P., Olhoeft, G., Lucius, J., Sander, K., and Mazzella, A. 1993. Geophysics and solvents: the Borden experiment. *The Leading Edge.* 12(4), 261-267.

Greenhouse, J. P. and Harris, R. D. 1983. Migration of contaminants in groundwater at a landfill: a case study. *Journal of Hydrology.* 63, 177-197.

Healy, R. W. and Mills, P. C. 1991. Variability of an unsaturated sand unit underlying a radioactive-waste trench. *Soil Science Society of America Journal.* 55(4), 899-907.

Heath, R. C. 1984. *Ground-water Regions in the United States.* U. S. Geological Survey, Water Supply Paper 2242.

Hinze, W. J. 1990. The role of gravity and magnetic methods in engineering and environmental studies, in *Geotechnical and Environmental Geophysics Vol. 1: Review and Tutorial, SEG Investigations in Geophysics No. 5.* Ed. Stanley Ward. 75-126.

Hoekstra, P. and Blohm, M. 1990. Case histories of time-domain electromagnetic soundings in environmental geophysics, in *Geotechnical and Environmental Geophysics Vol. 2: Environmental and Groundwater, S.E.G. Investigations in Geophysics 5.* Ed. Stanley Ward. 1-17.

Hoekstra, P., Lahti, R., Hild, J., Bates, R.C. and Phillips, D. 1992. Case histories of shallow time domain electromagnetics in environmental site assessment. *Groundwater Monitoring Review.* 12(4), 110-117.

Howard, P. H., Boethling, R. S., Jarvis, W. F., Meylan, W. M., and Michalenko, E. M. 1991. *Handbook of Environmental Degradation Rates.* Lewis Publishers, Boca Raton, FL.

Hubbard, S., Rubin, Y., and Majer, E. 1997a. Ground penetrating radar-assisted saturation and permeability estimation in bimodal systems. *Water Resources Research.* 33(5).

Hubbard, S. S., Peterson, J. E., Jr., Major, E., Zawislanski, P., Williams, K., Roberts, J. and Wabber, F. 1997b,. Estimation of permeable pathways and water content using tomographic radar data, *The Leading Edge.* 16(11), 1623-1628.

Huntley, D. 1986. Relations between electrical resistivity in granular aquifers. *Ground Water*. 24(4), 466-474.

Hyndman, D. W., Harris, J. M., and Gorelick, S. M., 1994. Capled seismic and tracer test inversion for aquifer property characterization. *Water Resources Research*. 30(7), 1965-1977.

Hyndman, D. W. and Gorelick, S. M. 1996. Estimating lithologic and transport properties in three dimensions using seismic and tracer data, the Kesterson aquifer. *Water Resources Research*. 32(9), 2659-2670.

Indelman, P. and Dagan, G. 1993. Upscaling of permeability of anisotropic heterogeneous formations, 1. The general framework. *Water Resources Research*. 29(4), 917-923.

Indelman, P. and Rubin, Y. 1995. Flow in heterogeneous media displaying a linear trend in the logconductivity. *Water Resources Research*. 31(5), 1257-1265.

Indelman, P. and Rubin, Y. 1996. Average flow in heterogeneous media of trending conductivity. *Journal of Hydrology*. 183, 57-68.

Istok, J. D., Blout, D. O., Barker, L., Johnejack, K. R., and Hammermeister, D. P. 1994. Spatial variability in alluvium properties at a low-level nuclear waste site. *Soil Science Society of America Journal*. 58(4), 1040-1051.

Istok, J. D., Rautman, C. A., Flint, L. E., and Flint, A. L. 1994. Spatial variability in hydrologic properties of a volcanic tuff. *Ground Water*. 32(5), 751-760.

Jackson, P. D., Taylor-Smith, D., and Stanford, P. N. 1978. Resistivity-porosity-particle shape relationships for marine sands. *Geophysics*. 43, 1250-1268.

Jacobson, E. A. 1990. Investigation of the spatial correlation of saturated hydraulic conductivities from a vertical wall of a trench, paper presented at the Canadian/American Conference on Hydrogeological Parameter Identification and Estimation for Aquifer and Reservoir Characterization, Alta. Res. Counc., Assoc. of Groundwater Scientists and Engineers, Calgary, Alta., Canada, September 18-20.

Johnson, A. I., Prill, R. C., and Morris, D. A. 1963. Specific yield — column drainage and centrifuge moisture content. *U.S. Geological Survey Water-Supply Paper 1662-A*.

Jury, W. A., Spencer, W. F., and Farmer, W. J. 1984. Behavior assessment model for trace organics in soil: III. Application of screening model. *J. Environ. Qual.* 13(4), 573-579.

Jussel, P., Stauffer, F., and Dracos, T. 1994. Transport modeling in heterogeneous aquifer: 1. Statistical description and numerical generation of gravel deposits. *Water Resources Research*. 30(6), 1803-1817.

Kachanoski, R. G., Gregorich, E. G., and Van Wesenbeeck, I. J. 1988. Estimating spatial variations of soil water content using noncontacting electromagnetic inductive methods. *Canadian Journal of Soil Science*. 68, 715-722.

Kachanoski, R. G., De Jong, E., and Van Wesenbeeck, I. J. 1990. Field scale patterns of soil water storage from non-contacting measurements of bulk electrical conductivity. *Canadian Journal of Soil Science*. 70, 537-541.

Kelly, W. E. 1977. Geoelectric sounding for estimating aquifer hydraulic conductivity. *Ground Water*. 15(6), 420-425.

Keys, S. W. 1989. *Borehole Geophysics Applied to Ground-Water Investigations*. National Water Well Association, Dublin, OH.

Khan, M. A. and Liang, T. 1989. Mapping pesticide contamination potential. *Environmental Management*. 13(2), 233-242.

Kick, J. F. 1985. Depth to bedrock using gravimetry. *Geophysics: The Leading Edge of Exploration, Society of Exploration Geophysicists*. 4(4), 38-42.

Klimentos, T. 1991. The effects of porosity-permeability-clay content on the velocity of compressional waves. *Geophysics*. 56, 1930-1939.

Knoll, M., Knight, R., and Brown, E. 1995. Can accurate estimates of permeability be obtained from measurement of Dielectric properties? *SAGEEP Annual Meeting Extended Abstracts*. Orlando, FL, Apr. 23-26.

Knox, R. C., Sabatani, D. A., and Canter, L. W. 1993. *Subsurface Transport and Fate Processes*. Lewis Publishers, Boca Raton, FL.

Landmeyer, J. and Stone, P. 1995. Radiocarbon and $\delta^{13}C$ values related to ground-water recharge and mixing. *Ground Water*. 33(2), 227-234.

Lankston, R. W. 1990. High-resolution refraction seismic data acquisition and interpretation, in *Geotechnical and Environmental Geophysics Vol. 1: Environmental and Groundwater, S.E.G. Investigations in Geophysics 5*. Ed. Stanley Ward. 45-73.

Lerner, D., Issar, A., and Simmers, I. 1990. *Groundwater Recharge, A Guide to Understanding and Estimating Natural Recharge*. R. Van Acken GmbH, West Germany.

Loague, K., Green, R. E., Giambelluca, T. W., Liang, T. C., and Yost, R. S. 1990. Impact of uncertainty in soil, climatic, and chemical information in a pesticide leaching assessment. *Journal of Contaminant Hydrology*. 5, 171-194.

Lyman, W. J., Reehl, W. F., and Rosenblatt, D. H. 1982. *Handbook of Chemical Property Estimation Methods*. McGraw-Hill Book Company, New York.

Mackay, D., Shiu, W. Y., and Ma, K. C. 1992. *Illustrated Handbook of Physical-Chemical Properties and Environmental Fate for Organic Chemicals*, Volume I-IV. Lewis Publishers, Boca Raton, FL.

Mac Lean, H. D. 1992. *Commercially Available Broadband Electromagnetic Systems for U.S. Department of Energy Waste-Site Characterization*. Chem-Nuclear Geotech, Inc., U.S. DOE, Grand Junction Projects Office, Grand Junction, CO, GJPO-GP-2.

Marion, D., Nur, A., Yin, H., and Han, D. 1992. Compressional velocity and porosity in sand-clay mixtures. *Geophysics*. 57, 554-563.

Mazac, O., Kelly, W. E., and Landa, I. 1985. A hydrogeophysical model for relations between electrical and hydraulic properties of aquifers. *Journal of Hydrology*. 79, 1-19.

Mazac, O., Cislerova, M., Kelly, W. E., Landa, I., and Venhodova, D. 1990. Determination of hydraulic conductivities by surface geoelectrical methods, in *Geotechnical and Environmental Geophysics: Environmental and Groundwater, S.E.G. Investigations in Geophysics 5*. Ed. Stanley Ward. 125-131.

McLaughlin, D., Reid, L. B., Li, Shu-Guang, and Hyman, J. 1993. A stochastic method for characterizing ground-water contamination. *Ground Water*. 31(2), 237-249.

McNeill, J. D. 1990. Use of electromagnetic methods for groundwater studies, in *Geotechnical and Environmental Geophysics Vol. 1: Review and Tutorial, SEG Investigations in Geophysics No. 5*. Ed. Stanley Ward. 191-218.

Mohanty, B. P., Kanwar, R. S., and Horton, R. 1991. A robust-resistant approach to interpret spatial behavior of saturated hydraulic conductivity of a glacial till soil under no-tillage system. *Water Resources Research*. 27(11), 2979-2992.

Moltyaner, G. L. 1986. Stochastic versus deterministic — a case study. *Hydrogeologie*. 2, 183-196.

Mood, A. M., Graybill, F. A., and Boes, D. C. 1963. *Introduction to the Theory of Statistics*. McGraw-Hill, New York., 564 p.

NRC (National Research Council). 1993. *Water Vulnerability Assessment: Contamination Potential Under Conditions of Uncertainty*. National Academy Press, Washington, D.C.

Neuman, S. P. 1990. Universal scaling of hydraulic conductivity and dispersivities in geologic media. *Water Resources Research*. 26(8), 1749-1758.

Peterson, J. E., Paulsson, B. N., and McEvilly, T. V. 1985. Applications of algebraic reconstruction techniques to crosshole seismic data. *Geophysics*. 50, 1566-1580.

Pfeifer, M. C. and Anderson, H. T. 1995. CD-resistivity array to monitor fluid flow at the INEL-infiltration test. Proceedings of the Symposium on the Application of Geophysics to Engineering and Environmental Problems. Orlando, FL. 709-718, April 23-26.

Ponce, V. 1989. *Engineering Hydrology*. Prentice Hall, Englewood Cliffs, NJ.

Press, F. 1966. Seismic velocities, in *Handbook of Physical Constants*, rev. ed., Clark, S. P., Jr., Ed. Geologic Society of America Memoir. 97: 97-173.

Pullan, S. E. and Hunter, J. A. 1990. Delineation of buried bedrock valleys using the optimum offset shallow seismic reflection technique, in *Geotechnical and Environmental Geophysics: Vol. 3, Geotechnical, S.E.G. Investigations in Geophysics 5*. Ed. Stanley Ward. 75-87.

Ramirez, A., Daily, W., LaBrecque, D., Owen, E., and Chesnut, D. 1993. Monitoring an underground steam injection process using electrical resistance tomography. *Water Resources Research*. 29(1), 73-87.

Rao, P. S. C., Hornsby, A. G., and Jessup, R. E. 1985. Indices for ranking the potential for pesticide contamination of groundwater. *Soil and Crop Science Society of Florida Proceedings*. 44, 1-8. (Forty-Fourth Annual Meeting, Jacksonville Beach, FL, October 23-25, 1984).

Rehfeldt K. R., Boggs, J. M., and Gelhar, L. W. 1992. Field study in a heterogeneous aquifer 3. Geostatistical analysis of hydraulic conductivity. *Water Resources Research*. 28(12), 3309-3324.

Robinson, E. S. and Coruh, C. 1988. *Basic Exploration Geophysics*, John Wiley and Sons, New York.

Rubin, Y. and Bellin, A. 1994. The effects of recharge on flow nonuniformity and macrodispersion. *Water Resources Research*. 30(4), 939-948.

Rubin, Y., Cushey, M. A., and Bellin, A. 1994. Modeling of transport in groundwater for environmental risk assessment. *Stochastic Hydrology and Hydraulics*. 8, 57-77.

Rubin, Y. and Dagan, G. 1987a. Stochastic identification of transmissivity and effective recharge in steady groundwater flow, 1, Theory. *Water Resources Research*. 23(7), 1185-1192.

Rubin, Y. and Dagan, G. 1987b. Stochastic identification of transmissivity and effective recharge in steady groundwater flow, 2, Case study. *Water Resources Research*. 23(7), 1193-1200.

Rubin, Y. and Gomez-Hernandez, J. J. 1990. A stochastic approach to the problem of upscaling of conductivity in disordered porous media: theory and unconditional numerical simulations. *Water Resources Research*. 26(4), 691-701.

Rubin, Y., Mavko, G., and Harris, J. 1992. Mapping permeability in heterogeneous aquifers using hydrological and seismic data. *Water Resources Research*. 28(7), 1192-1800.

Rubin, Y. and Seong, K. 1994. Investigation of flow and transport in certain cases of nonstationary conductivity fields. *Water Resources Research*. 30(11), 2901-2912.

Rubin, Y. 1995. Flow and transport in bimodal heterogeneous formations. *Water Resources Research*. 31(10), 2461-2468.

Rushton, K. R. 1988. Numerical and conceptual models for recharge estimation in arid and semi-arid zones. *Estimation of Natural Groundwater Recharge*. Ed. I. Simmers. D. Reidel Publishing Co., Dordrecht, Holland.

Sanchez-Vila, X., Girardi, J. P., and Carrera, J. 1995. A synthesis of approaches to upscaling of hydraulic conductivities. *Water Resources Research*. 31(4), 867-882.

Sharma, H. and Lewis, S. 1994. *Waste Containment Systems, Waste Stabilization, and Landfills — Design and Evaluation*. John Wiley and Sons, New York.

Shaw, R. D. and Prepas, E. E. 1990. Groundwater-lake interactions: I. Accuracy of seepage meter estimates of lake seepage. *Journal of Hydrology*. 119(1-4), 105-120.

Sheets, K. R. and Hendricks, J. M. H. 1995. Noninvasive soil water content measurement using electromagnetic induction. *Water Resources Research*. 31(10), 2401-2409.

Shuter, E. and Teasdale, W. E. 1981. Application of drilling, coring, and sampling techniques to test holes and wells, in *Techniques of Water Resources Investigations TWI 02-F1*.

Smith, D. G. and Jol, H. M. 1992. Ground-penetrating radar investigation of a Lake Bonneville Delta, Provo level, Bringham City, Utah. *Geology*. 20, 1083-1086.

Soil Conservation Service (SCS). 1988. *Soil Taxonomy*. Robert E. Krieger Publishing Company, Malaber, FL.

Solomon, D. K., et al. 1995. Site characterization using ^{3}H/^{3}He ground-water ages, Cape Cod, MA. *Ground Water*. 33(6), 988-996.

Steeples, D. W. and Miller, R. D. 1990. Seismic reflection methods applied to engineering, environmental, and groundwater problems, in *Geotechnical and Environmental Geophysics, S.E.G. Investigations in Geophysics 5, Vol. 1.* Ed. Stanley Ward. 1-30.

Stephens, D. B. 1996. *Vadose Zone Hydrology.* CRC Press, Boca Raton, FL.

Telford, W. M., Geldart, L. P., and Sheriff, R. E. 1990. *Applied Geophysics.* Second Edition, Cambridge University Press.

Topp, G. C., Davis, J. L., and Annan, A. P. 1980. Electromagnetic determination of soil-water content: Measurements in co-axial transmission lines. *Water Resources Research.* 16, 574-582.

Ulriksen, P. F. 1982. *Application of Impulse Radar to Civil Engineering.* Ph.D. Thesis: Lund University, Sweden.

Ünlü, K., Nielsen, D. R, and Biggar, J. W. 1990. Stochastic analysis of unsaturated flow: one dimensional Monte Carlo simulations and comparisons with spectral perturbation analysis and field observations. *Water Resources Research.* 26(9), 2207-2218.

U.S. Environmental Protection Agency (U.S. EPA). 1986. *Superfund Public Health Evaluation Manual.* Washington, D.C., EPA/540/1-86/060.

VanNostrand, R. G. and Cook, K. L., 1966. Interpretation of Resistivity Data, U.S. Geological Survey Professional Paper, PO499, 310pp.

Ward, S., 1990. Geotechnical and environmental geophysics, in *Geotechnical and Environmental Geophysics: Environmental and Groundwater*, S.E.G. Investigations in Geophysics 5.

Wen, X. H. and Gomez-Hernandez, J. J. 1996. Upscaling hydraulic conductivities in heterogeneous media — an overview. *Journal of Hydrology.* 186(1-2), R9-R32.

Wierenga, P. J., Hills, R. G., and Hudson, D. B. 1991. The Las Cruces trench site: characterization, experimental results, and one-dimensional flow predictions. *Water Resources Research.* 27(10), 2695-2705.

Wilt, M., Morrison, H. F., Becker, A., Tseng, H., Lee, K., Torres-Verdin, C., and Alumbaugh, D. 1995. Crosshole electromagnetic tomography: a new technology for oil field characterization. *The Leading Edge.* 14(3), 173-177.

Wyatt, D. E., Waddell, M. G., and Sexton, G. B. 1996. Geophysics and shallow faults in unconsolidated sediments. *Ground Water.* 34(2), 236-334.

Yilmaz, O. 1997. *Seismic Data Processing: Investigations in Geophysics*, Vol 2. Society of Exploration Geophysics, Vol. 2, Tulsa, OK.

Wharton, R. P., Hazen, G. A., and Best, D. L. 1980. Electromagnetic propagation logging: Advances in technique and interpretation. Soc. Pet. Eng. Pap. 9267.

Zodhy A. A., Eaton, G. P., and Mayhey, D. R. 1974. Applications of surface geophysics to groundwater investigations. Techniques of Water Resources Investigations of the U.S. Geological Survey, Book 2, Chapter D1.

Glossary

Conditional Monte Carlo Simulation   Modeling uncertainty through generation of alternative images of the flow domain using the space random function model and, in addition, requiring that the generated images fit field data at specified locations.

Crosshole Geophysical Methods   Acquisition geometry used by seismic, radar, electrical resistivity, and electromagnetic induction methods whereby energy from sources located in one borehole are recorded by sensors in another.

Dielectric Constant   A measure of the separation (polarization) of opposite electrical charges within a material that has been subjected to an external electrical field.

Dynamic Process Monitoring   Collection of data at the same location as a function of time. Processing of these data as "time difference" cross sections can reveal information about processes such as water infiltration, stream flooding, or contaminant plume migration.

Effective Conductivity   The coefficient multiplying the expected value of the head gradient to yield the expected value of the flux. Must not depend on the parameters defining the flow problem or the boundary conditions. In a stationary domain, the expected value of the flux is also equal to its space average.

Electrical Profiling   An electrical survey made at several surface locations using a constant electrode separation distance. Electrical profiles, also called constant spread profiles or resistivity profiles, provide information about lateral changes in apparent resistivity.

Electrical Resistivity   A measure of the ability of electrical current to flow through materials, measured in Ohm-m. Electrical resistivity is an intrinsic property of a material and is the inverse of electrical conductivity.

Electrical Sounding   An electrical survey made at a single surface location by moving the electrodes progressively farther apart. Electrical soundings, also called expanding spread profiles, provide information about apparent electrical resistivity as a function of depth.

Monte Carlo Simulation   Modeling uncertainty through generation of alternative images of the flow domain using the space random function model.

Seismic Velocity   The propagation velocity of a seismic disturbance through a material. As seismic velocity is a function of the elastic parameters of a material, seismic velocity has been used to estimate lithology and porosity.

Space Random Function   A probabilistic model for describing the variations of attributes such as hydraulic conductivity in heterogeneous media.

Spatial Correlation   A statistical measure of similarity between attributes in a heterogeneous domain, usually as a function of separation distance between two locations. Spatial correlation is modeled through the spatial covariance or the semivariogram.

Stationarity   Invariance of the space random function model to translation in space.

Statistical Anisotropy   Implies that the spatial correlations are direction dependent.

Statistical Isotropy   Invariance of the spatial correlation to rotation of the separation distance.

Upscaling   Representing the heterogeneous domain for modeling purposes through elements larger than the scales of heterogeneity but not large enough to warrant the use of effective properties. The effects of the sub-element heterogeneity are represented using coefficients which depend on the element size.

11

Geophysical and Tracer Characterization Methods

David W. Hyndman
Michigan State University

11.1 Introduction

This chapter presents aquifer characterization methods that use geophysical and tracer data. Detailed aquifer properties are needed to predict the fate of groundwater solutes and to design efficient remediation strategies for contaminated aquifers, yet the hydrogeologic data used to estimate these properties are generally local and sparse. Low-resolution aquifer property estimates may fail to describe the dominant flow and transport paths through an aquifer, resulting in significant errors in groundwater flow and solute transport predictions. One approach for improving the resolution of heterogeneous aquifer properties combines densely sampled geophysical information with sparse hydrogeologic data in numerical inversions. This chapter provides: a synopsis of simulation-optimization methods; a description of surface seismic, ground-penetrating radar (GPR), and cross-well seismic tomography methods; and a discussion of approaches to combine geophysical and hydrogeologic data.

11.2 Simulation/Optimization Methods for Aquifer Property Estimation

A powerful approach for estimating aquifer properties involves coupling nonlinear optimization methods with numerical groundwater flow and solute transport simulations. Such numerical simulations overcome the difficulties associated with incorporating highly heterogeneous aquifer properties into analytic solutions to the equations that govern groundwater flow and solute transport. The groundwater flow equation is used to solve for hydraulic head (h) given estimates of hydraulic conductivity (K) and specific storage (S_s) (for transient simulations)

$$S_s \frac{\partial h}{\partial t} = \nabla \cdot \left(K \cdot \nabla h \right) - W \tag{1}$$

where t is time; ∇ is the gradient operator ($\partial/\partial x,\ \partial/\partial y,\ \partial/\partial z$), x, y, and z are Cartesian coordinates, and W is the source fluid flux. Darcy's law can then be used to approximate the average linear velocity of the fluid (V) given the calculated hydraulic head values as well as estimates of hydraulic conductivity and effective porosity (θ).

$$V = -\frac{K}{\theta} \nabla h \tag{2}$$

The advection-dispersion equation is used to approximate the spatial and temporal distribution of solute concentrations (C) using the calculated velocities and given estimates of the hydrodynamic dispersion tensor (D)

$$R\frac{\partial C}{\partial t} = \nabla \left(D \cdot \nabla C \right) - V \nabla C + \frac{W(C - C')}{\theta} - B \tag{3}$$

where $D = D(\alpha, v)$, α is the dispersivity, R is the retardation factor, C' is the solute concentration in the source fluid, and B is the reaction term (i.e., biodegradation). Numerical approximations of these equations, however, can easily incorporate realistic aquifer heterogeneity in three dimensions, while optimization methods provide an efficient method of adjusting aquifer properties to best match all available data.

Despite the power of optimization methods, parameters for hydrogeologic models are commonly determined using trial and error calibration, which is inefficient and highly dependent on the knowledge and experience of the model operator. A more efficient and mathematically rigorous approach can be developed using weighted nonlinear least squares *regression*, which can incorporate measurement uncertainty and operator judgment through the assigned weights. Regression approaches provide a consistent statistical framework for estimating aquifer properties. The goal (or objective function) of least squares regression for aquifer property estimation is generally to minimize the weighted sum-of-squared residuals between measured and simulated values of hydraulic heads, tracer concentrations, and other measured states of the system. Other objective functions include minimizing the sum of the absolute value of residuals, or the squared residuals between some index of the data, such as spatial or temporal concentration *moments* or concentration arrival time quantiles. A well-posed problem will rapidly converge toward the *global minimum* objective value (i.e., no combination of parameters will provide a lower objective value) from a variety of starting parameter sets. Optimization methods work best when the available data (hydraulic head, tracer concentrations, advective front location, water temperature, velocity, etc.) are highly sensitive to the estimated *aquifer properties*, such as hydraulic conductivity, dispersivity, and effective porosity. These concepts are discussed in more detail along with a discussion of other parameter estimation methods in Peck et al. (1988).

11.2.1 Tracer Tests

The main limiting factor for inferring the detailed heterogeneous structure of aquifers is the sparse nature of hydrogeologic data. Readily available hydrogeologic data include local hydraulic conductivity values from well cores, *pump tests and slug tests*, as well as hydraulic head and tracer concentration measurements at observation wells. Concentration and hydraulic head measurements provide spatially averaged information about aquifer properties, while the limited core and pump test data provide local information about some irregularly shaped region near observation wells. Unfortunately, the size and geometry of the region characterized by single well pump tests or slug tests is unknown because water preferentially flows through high conductivity regions. The hydraulic response during such tests thus provides a hydraulic conductivity estimate for the high conductivity region near the wellscreen. Early time pump test data provide information about properties very close to the well, while later measurements provide information about more distant regions. A relatively new hydraulic testing method is to measure flow rates at multiple locations in a well due to natural or imposed flow conditions using a borehole flowmeter. This approach can provide high resolution information about heterogeneous aquifer properties as discussed by Molz (1994).

Tracer tests are valuable for estimating aquifer properties, such as effective porosity (θ) and dispersivity (α), yet hydraulic head should be measured at all possible locations during such tests to estimate the hydraulic conductivity field. Tracer concentration histories from a densely sampled well network provide a direct indication of transport velocities, thus providing information about the ratio of hydraulic conductivity to effective porosity for a known hydraulic gradient. Hydraulic head measurements are sensitive to the regional variations in hydraulic conductivity and provide information about the gradients. These measurements are inexpensive to obtain, and densely sampled hydraulic head data may be used in conjunction with seismic data to develop detailed hydraulic conductivity estimates, while head measurements alone provide limited information about the spatial patterns of hydraulic conductivity. In addition, head data can be used as observations for flow simulations, which are less computationally intensive than tracer simulations. When geophysical information is also available, it should be jointly analyzed with tracer and hydraulic data to provide high resolution aquifer property estimates, as will be described in Section 11.6 *Combining Seismic and Tracer Data*.

11.2.1.1 Natural Gradient Methods

The two categories of tracer tests are natural gradient tests and forced gradient tests. Natural gradient methods are often preferred because they do not disturb the natural flow system and the measured concentrations are more sensitive to the dispersivity. Disadvantages of natural gradient tests include their time-consuming nature and the difficulty in designing a downgradient sampling system that adequately measures the tracer plume through time. Classic examples of natural gradient tracer experiments were conducted at the Borden Aquifer in Ontario (relatively homogeneous glacial deposit), the Cape Cod Aquifer in Massachusetts (heterogeneous glacial outwash), and the MAcroDispersion Experiment Site (MADE) in Mississippi (highly heterogeneous alluvial deposit). These tests illustrated that heterogeneous aquifer properties strongly influence groundwater flow and solute transport.

The dispersivities at the Borden site were analyzed using *spatial moments* (*M*) of the measured concentrations (Freyberg, 1986)

$$M_{ijk}(t) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \theta C(x,y,z,t) x^i y^j z^k \, dx\,dy\,dz \qquad (4)$$

where θ is effective porosity, C is the concentration of a specific tracer at time t and location (x, y, z), and the indices $(i, j, \text{and } k)$ define the components of each moment:

M_{000} is the zeroth spatial moment, which is the total mass;
$M_{100}, M_{010}, M_{001}$, are the x, y, and z components of the first spatial moment;
$M_{200}, M_{020}, M_{002}$, are the x, y, and z components of the second spatial moment.

The concentration covariance tensor (σ) can be defined about the center of mass as described by Freyberg (1986):

$$\sigma_{xx} = \left(M_{200}/M_{000}\right) - x_c^2;$$

$$\sigma_{yy} = \left(M_{020}/M_{000}\right) - y_c^2;$$

$$\sigma_{zz} = \left(M_{002}/M_{000}\right) - z_c^2;$$

$$\sigma_{xy} = \sigma_{yx} = \left(M_{110}/M_{000}\right) - x_c y_c;$$

$$\sigma_{yz} = \sigma_{zy} = \left(M_{011}/M_{000}\right) - y_c z_c;$$

$$\sigma_{xz} = \sigma_{zx} = \left(M_{101}/M_{000}\right) - x_c z_c;$$

where the center of mass is defined as:

$$x_c = M_{100}/M_{000}$$

$$y_c = M_{010}/M_{000}$$

$$z_c = M_{001}/M_{000}$$

The dispersivity tensor (A, with components α_l, α_t, …) can then be approximated using

$$A = \frac{1}{2|V|}\frac{d}{dt}\left(\sigma\right) \tag{5}$$

for cases with negligible molecular diffusion, which is a common assumption for groundwater transport studies due to the larger magnitude of hydrodynamic dispersion. This analysis of the Borden tracer moments showed a significant temporal increase in both the longitudinal and transverse components of the covariance tensor (related to dispersivity with Equation [5]), which may indicate the effects of the plume encountering more heterogeneity at larger scales (Figure 11.1). In addition, the center of mass sinks through time, especially early in the experiment, and the original plume appears to split into two distinct plumes at late time despite the relatively homogeneous material at this site. Sudicky (1986) developed detailed estimates of log hydraulic conductivity along two transects at this site using permeameter measurements on cores. These were used to estimate macrodispersivity coefficients, using the theory of Gelhar and Axness (1983), which were consistent with those obtained from the tracer test analysis.

The tracer tests at the Cape Cod site were used to analyze the heterogeneous hydraulic and chemical properties of the shallow aquifer (Hess et al., 1992; LeBlanc et al., 1991; Garabedian et al., 1991). Each tracer plume was irregular in shape and appeared to sink as it moved downgradient (Figure 11.2). Two of the likely causes for this effect are recharge-induced downward flow, and the slight density contrast between the tracer plume and the regional groundwater. The concentration moment analysis indicated an order of magnitude larger longitudinal dispersion for a reactive tracer (lithium) as compared to a nonreactive tracer (bromide) (Gelhar, 1993), which indicates the need to properly describe both the physical and chemical transport processes of aquifers.

At the MADE site in Mississippi, only about half of the injected tracer could be accounted for at late times as the tracer plume advected down an unidentified channel deposit in advance of the expanding
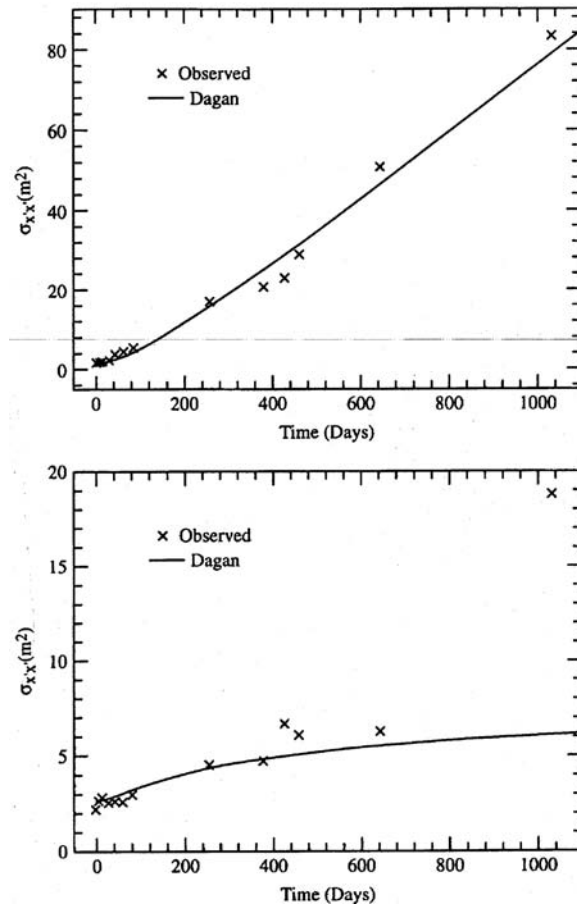
FIGURE 11.1 Plots of estimated longitudinal and transverse components of the spatial covariance tensor illustrating the temporal dependence of these statistics. These values can be used to approximate the dispersivities in these directions. (From Freyberg, D. L. 1986. A natural gradient experiment on solute transport in a sand aquifer 2. Spatial moments and the advection and dispersion of nonreactive tracers. *Water Resour. Res.* 22(13):2031-2046. With permission.)

sampling network (Adams and Gelhar, 1992; Boggs et al., 1992). This result indicates the need for a detailed tracer test design that accounts for the likely heterogeneous aquifer conditions. The tracer data clearly illustrated the effect of large-scale heterogeneities in hydraulic conductivity. This was considered a trend in the spatial covariance analysis (Rehfeldt et al., 1992), with the primary hydraulic conductivity estimates obtained using a borehole flowmeter. In addition, this experiment demonstrated the need for high-resolution noninvasive geophysical techniques to identify such large-scale heterogeneities.

11.2.1.2 Forced Gradient Methods

The long duration of natural gradient methods is often inconsistent with the need for rapid site assessment, thus forced gradient tests are often used in practice. The first stage of a forced gradient test is pumping for a period of time to approximately achieve steady-state hydraulic heads in the region of the test. In divergent flow tracer tests, water is pumped into a well while concentrations are measured at several different observation wells at radius (r). This type of test can be used to estimate effective porosity θ in a homogeneous hydraulic conductivity field using (Sanchez-Villa et al., 1992)
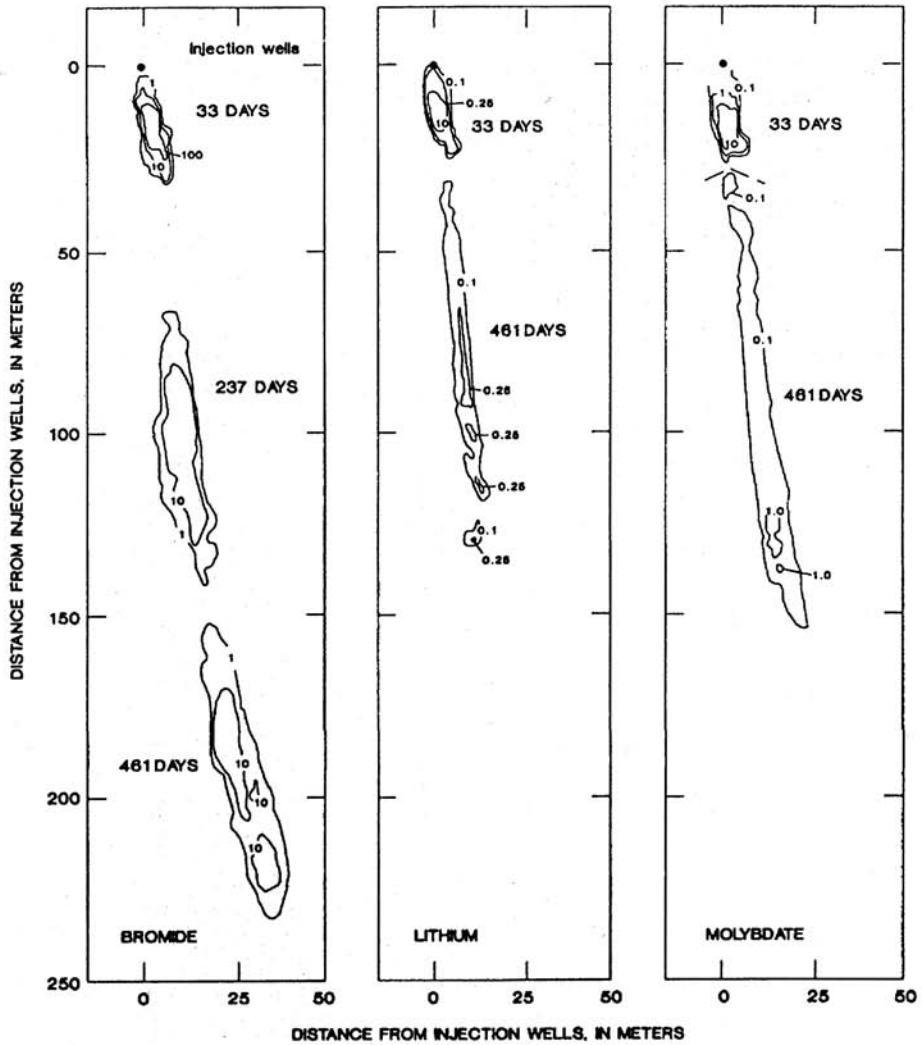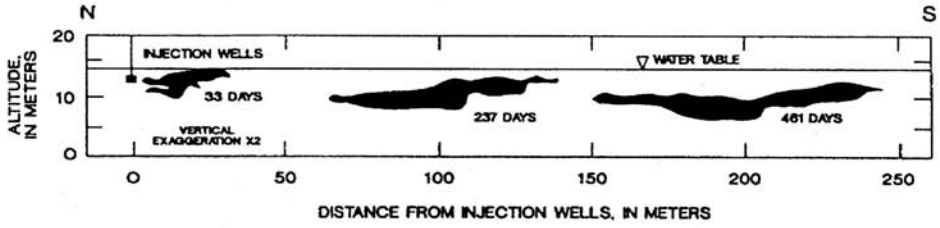
FIGURE 11.2 Vertical and horizontal plots of tracer concentration for the Cape Cod natural gradient tracer test. The variable plume geometry demonstrates the influence of physical heterogeneities on solute transport, while the difference in shapes of the three tracers illustrates the influence of geochemistry on solute transport. (From LeBlanc, D. R. et al. 1991. Large-scale natural gradient tracer test in sand and gravel, Cape Cod, Massachusetts. 1. Experimental design and observed tracer movement. *Water Resour. Res.* 27(5):895-910. With permission.)

$$\theta = \frac{Qt_r}{\pi r^2 b} \tag{6}$$

where Q is the injection flow rate over the entire saturated thickness of the aquifer (b), and t_r is the concentration travel time from the injection to observation well (i.e., can be approximated as the peak concentration arrival time or center of mass for a pulse injection of a conservative tracer).

Convergent flow tracer tests can be used to explore anisotropy in hydraulic conductivity and dispersivity. In a convergent flow tracer test, water is pumped out of one well which is monitored for tracer concentrations derived from small injected tracer pulses from other wells. Each test has to be conducted at different times, or multiple tracers with similar transport characteristics must be injected into the different wells. This type of analysis showed significant directional dependence in tracer concentration histories in an anisotropic fractured gneiss and the Culebra dolomite near the Waste Isolation Pilot Plant (Sanchez-Villa et al., 1992).

At the Kesterson Aquifer in California's central valley, a forced gradient tracer test was conducted using both an injection and a withdrawal well to assess the heterogeneous properties of a shallow aquifer between these wells (Benson et al., 1991). Benson (1988) performed a series of one-dimensional transport simulations to approximate the longitudinal dispersivities of this site. This site was also used to test several approaches that combine seismic, tracer, and hydraulic data for aquifer property estimation, as described later in this chapter.

11.3 Application of Geophysical Methods to Environmental Problems

Spatially averaged hydrogeologic data can be supplemented with higher resolution geophysical information to develop detailed estimates of aquifer structure and properties. Geophysical techniques, such as seismic and radar methods, have successfully imaged a variety of subsurface features including heterogeneous lithologies in oil reservoirs, aquifers, and mines; the geometry of subsurface cavities and buried objects; variations in fluid properties and saturation; and the structure of glaciers. The following sections discuss: (1) an overview of surface seismic methods; (2) two promising geophysical methods that have limited discussion in current textbooks (ground-penetrating radar and cross-well seismic tomography); and (3) approaches to combining geophysical and tracer data.

11.3.1 Surface Seismic Reflection and Refraction Surveys

Seismic *reflection* and *refraction* surveys have been widely used to characterize subsurface environments for decades. Such surveys provide information about the geometry and nature of subsurface lithologies for both petroleum exploration and environmental remediation work. Reflection and refraction seismic surveys generate sound waves at a near-surface source (e.g., hammer on a steel plate, explosion) and measure the resulting seismic signal at multiple receivers called geophones (Figure 11.3). The measured signals are then processed for amplitudes and travel times for refracted and reflected seismic energy.

A portion of the incident seismic energy is reflected when it encounters a contrast in seismic impedance, which equals density (ρ) times seismic velocity (V), and the amount of reflected energy is proportional to the degree of contrast. The time (t) for reflected seismic energy to reach the surface from a single flat interface is

$$t = \sqrt{\frac{4d^2}{V^2} + \frac{x^2}{V^2}} \tag{7}$$
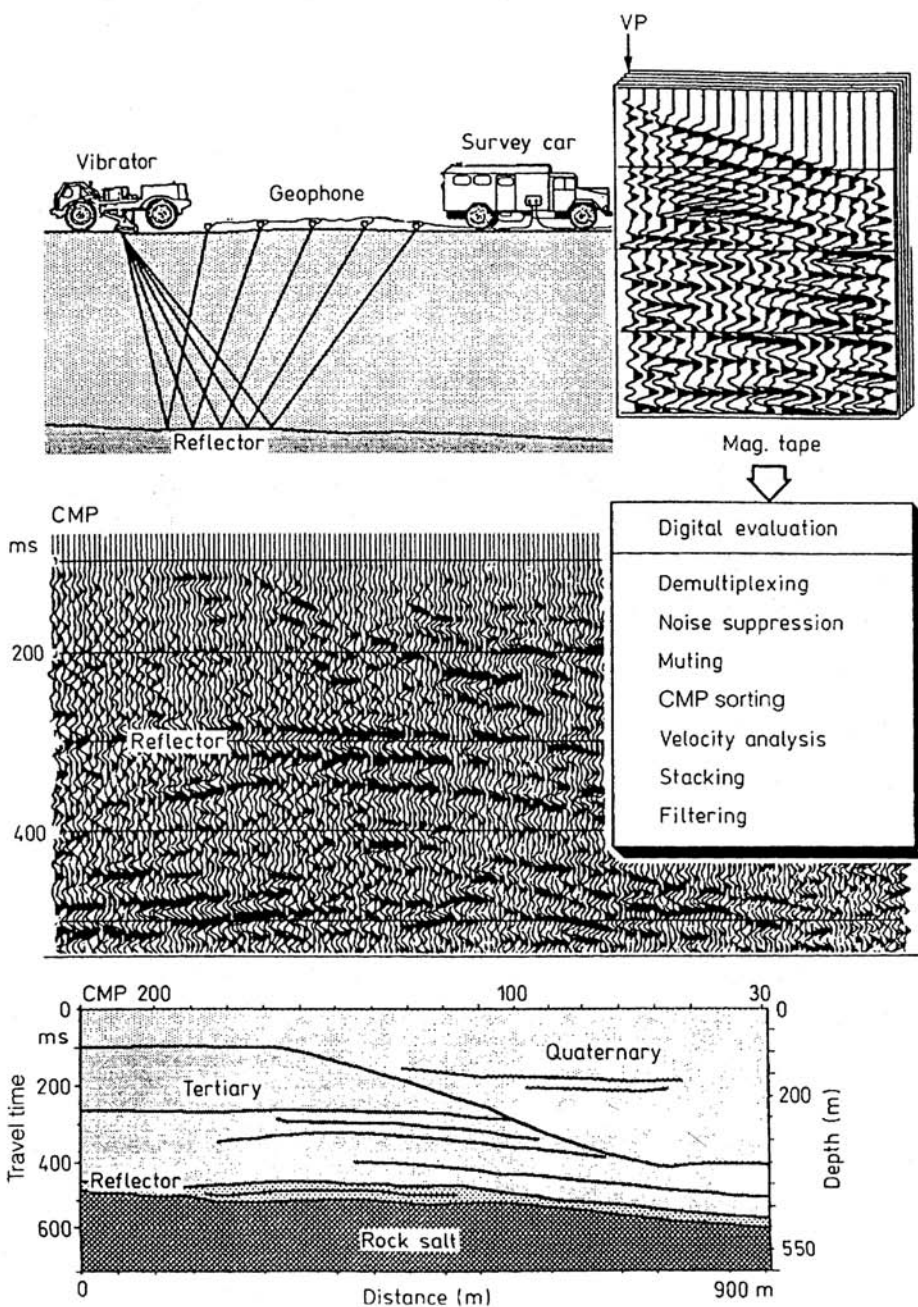
FIGURE 11.3 A diagram of seismic reflection data collection, processing, and interpretation. (From Vogelsang, D. 1995. *Environmental Geophysics: A Practical Guide.* Springer Verlag, Berlin, 173 p. With permission.)

where *d* is the depth to the interface. The first term under the square root represents the time for a normal incidence reflection (source and receiver at same location), while the second term represents parabolic moveout due to offset *x* between source and receiver. More complex formulations for multiple dipping and curved interfaces can be found in introductory geophysics texts (e.g., Telford et al., 1990). Surface seismic reflection data can be processed by adjusting the velocity for a *common midpoint gather* to flatten the observed hyperbolic shape of the reflector, which is called a *normal moveout correction*.
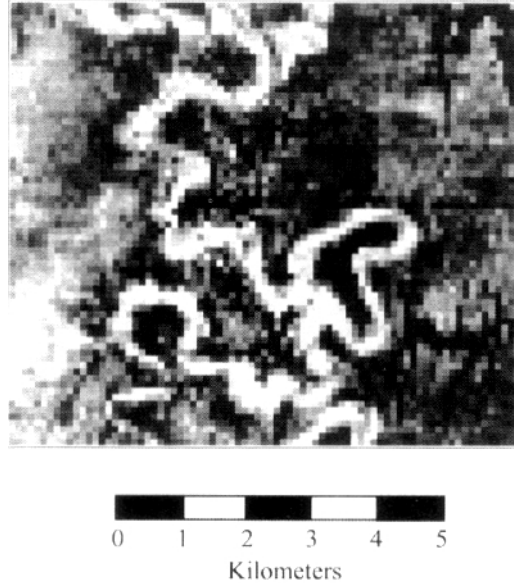
FIGURE 11.4 A horizontal slice through a three-dimensional processed seismic reflection survey from the Gulf of Thailand that illustrates a buried meandering river channel. (Data courtesy of the Stanford Exploration Project.)

Through this correction, the measured times are corrected from their offset positions back to a normal incidence position (no offset). The data are then stacked across each corrected common midpoint gather to provide an estimate of the seismic velocity structure at the midpoint of the plane between source and receiver locations (see Figure 11.3). This method does not account for interfaces that dip out of the plane, so the energy reflected from out of the image plane will appear at incorrect locations. As a result, the best approach is to collect and process surface seismic data in three dimensions. An excellent example of the type of information that can be obtained by three-dimensional analysis of surface seismic data is the clear evidence of an old meandering channel that can be observed in the processed data from the Gulf of Thailand (Figure 11.4). The fundamentals of seismic reflection methods are described in more detail in Steeples and Miller (1990).

Seismic refraction surveys are collected with far offset receivers to measure energy that travels along an interface with a higher velocity at depth (Figure 11.5a). Refracted arrivals can be discerned from directly reflected arrivals by the slope of the distance versus time curve (Figure 11.5b). Direct reflected arrivals for a two-layer geologic setting will fall on a line that intersects the origin, while a line connecting refracted arrivals will have a positive zero-offset intercept. To ensure refracted energy is the first to arrive at the receivers, a refraction survey needs to be several times longer than the depth of investigation and velocity must increase with depth. The slope of the time versus distance line at the appropriate offset gives the reciprocal of the velocity of the deeper layer (Figure 11.5). In complex heterogeneous systems, detailed velocity analysis and ray tracing may be necessary to infer velocities from measured seismic data (Claerbout, 1992). More detailed descriptions of surface seismic surveys are covered in texts such as Telford et al. (1990).

11.4 Ground-Penetrating Radar

Ground-penetrating radar (GPR) techniques generate high frequency electromagnetic waves (typically 10 to 1000 MHz) at a source antenna directed toward the ground. There is a trade-off between the high resolution obtained with high-frequency methods and the corresponding rapid signal *attenuation*. Preliminary surveys are often conducted with low-frequency radar antennas (50 to 100 MHz), with detailed higher frequency (200 to 900 MHz) surveys conducted in selected regions where detailed information
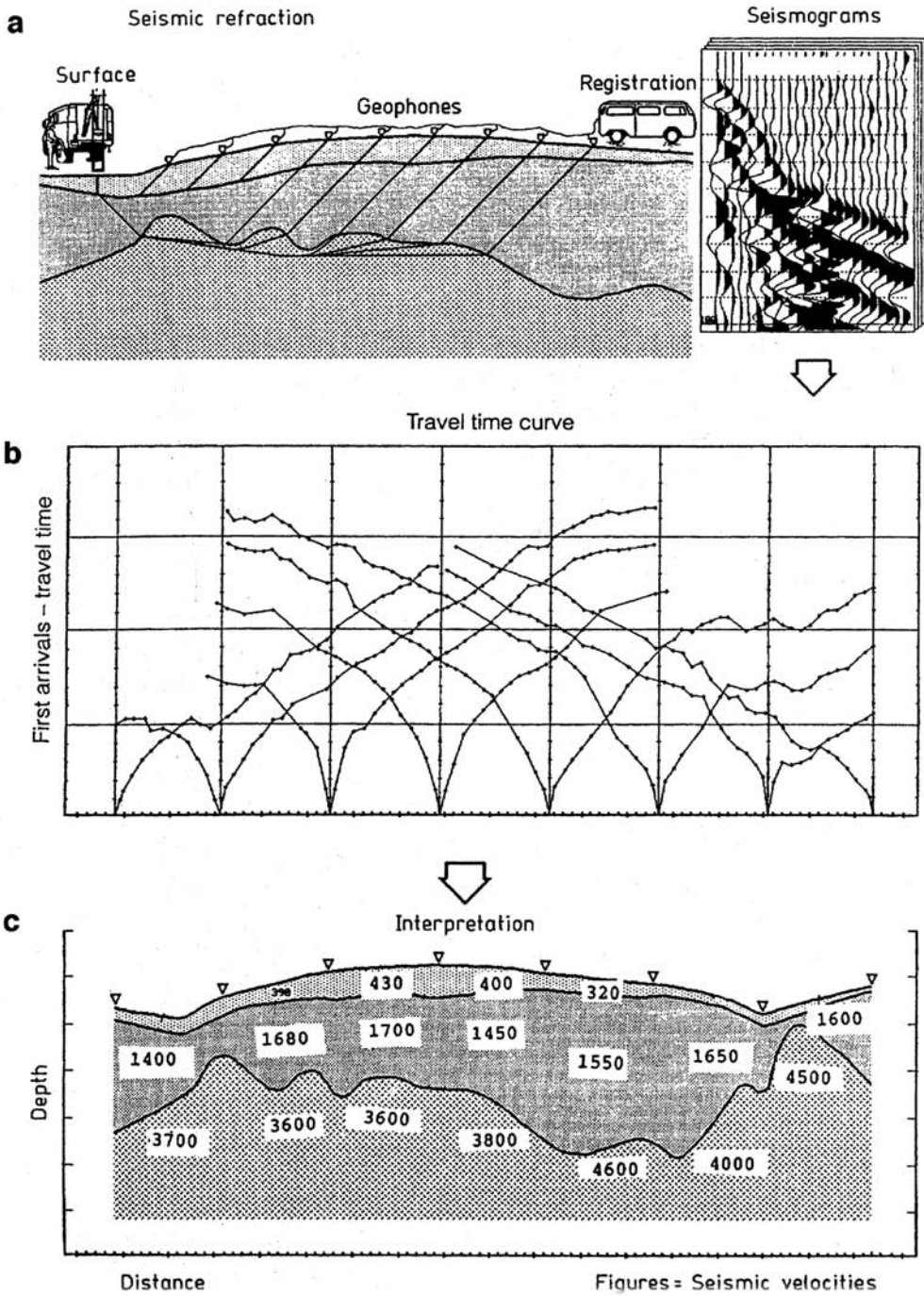
FIGURE 11.5 A diagram of a seismic refraction survey with the corresponding distance versus time plot for direct, reflected, and refracted waves. (From Vogelsang, D. 1995. *Environmental Geophysics: A Practical Guide.* Springer Verlag, Berlin, 173 p. With permission.)

about shallow properties is needed. The propagated direct wave (in cross-well GPR) or the reflected wave (in surface GPR) is measured at a receiver antenna. Surface GPR data are generally collected with a single spacing between antennas to provide nearly normal incidence reflections, although more information can be obtained using multiple offsets around each central point of interest, called common midpoint gathers, as described in Section 11.3.1. In both cases, multiple traces should be collected at each measurement point and stacked (or averaged) to improve the signal-to-noise ratio. Normal incidence reflection surveys have the benefit of rapid collection, since antennas can be towed across the surface at a constant rate while a signal is generated at a specified delay. This allows large regions to be rapidly surveyed and variable towing rates can be compensated using a timing mark at specified locations in a survey grid. A global positioning system (GPS) is a valuable addition to large radar surveys to continuously monitor the location of the radar system and map topography. Topographic data is necessary to correct the interpreted GPR depths to elevations above a reference, and lateral location information allows for rapid surveys of irregularly gridded regions.

Collecting traces in common midpoint gathers provides a suite of data that can be processed using surface seismic methods, yet the survey is more difficult to obtain because radar systems with multiple antennas are not readily available. The measured wavefields are processed for the first arrival traveltime and wave amplitudes. These wave attributes can then be processed using normal moveout corrections and mapped into depth versus amplitude plots which can be compared to lithological and moisture content variations in cores across the survey region. Tools that are commonly used to process seismic data such as frequency filtering, time-dependent gain, and migration can be used to improve interpretation of GPR data as described by Knoll et al. (1991) for the Cape Cod Aquifer. Rea and Knight (1998) estimated the near horizontal *variogram* (correlation vs. offset plot) of processed radar data from a gravel pit in southwestern British Columbia and compared this to the modeled variogram of a digitized image of a cliff face at the site. They found good agreement between both the direction of maximum correlation and the estimated correlation length determined from the two types of data. They then used this approach to infer the paleoflow direction of a buried channel and the correlation length in that direction from a three-dimensional GPR survey (Figure 11.6) over this site.

The subsurface propagation of electromagnetic energy depends on the electrical properties of the ground, commonly described using the relative dielectric permativity ε_r in the electromagnetic wave equation,
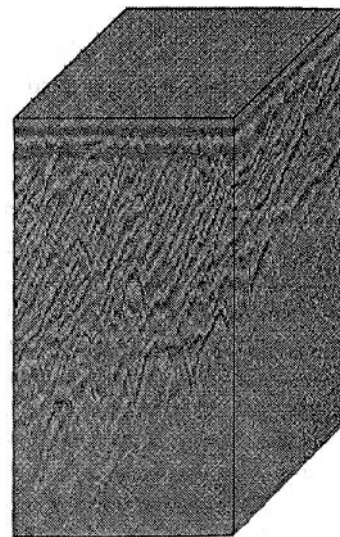
FIGURE 11.6 A three-dimensional GPR survey conducted with 100 MHz antennas over dipping foreset beds in southwestern British Columbia. (From Rea, J. and Knight, R. 1998. Geostatistical analysis of ground penetrating radar data: a means of describing spatial variations in the subsurface. *Water Resour. Res.* 34(3).

$$\nabla^2 E + i\mu\omega\sigma E + \omega^2 \varepsilon^* \mu E = 0 \tag{8}$$

where E is the electrical field vector {Volts/m}; $i = \sqrt{-1}$; μ is the magnetic permeability {Henry/m}; ω is the angular frequency {radian/s} $= 2\pi\upsilon$ (υ = frequency in Hz); s is the electrical conductivity {Siemen/m}; ε^* is the complex dielectric permativity of the media {Farad/m}.

The electromagnetic wave equation is rarely used because of its complexity and uncertainty in all the parameters (i.e., μ, σ, ε^*). Instead, the imaginary portion of the dielectric permativity is generally ignored, unless the goal is to describe electromagnetic wave attenuation, while the real portion of the relative dielectric permativity is used to describe radar propagation velocities using

$$V = \frac{c}{\sqrt{\varepsilon_r}} \tag{9}$$

where V is the electromagnetic wave velocity {m/ns}; c is the free space light velocity (0.3 m/ns); ε_r is the relative dielectric permativity $= (\varepsilon/\varepsilon_0)$ {dimensionless ratio}; ε is the real portion of the media dielectric permativity; ε_0 is the real portion of the free space dielectric permativity (8.854×10^{-12} Farad/m). This approach has been successful in low electrical loss materials such as dry sand, which are the environments that allow reasonable radar penetration depths.

The relative dielectric permativity represents the ability of an electric field to polarize a material and is dependent on properties of both the solid and fluid. This property is inversely proportional to frequency, especially in saturated sediments (Telford et al., 1990). Some typical values are plotted in Figure 11.7 for high-frequency (>100 MHz) measurements.

The large contrast between the dielectric permativity of water and that of dry sediment can be used to infer the saturation of sediments from radar measurements. In this sense, it may be possible to dynamically image wetting fronts, which would provide information about the hysteretic nature of unsaturated flow. The capillary fringe often appears as a dominant reflector in radar records because of the large vertical gradient in saturation at this depth. The reflector depth should only be interpreted as the water table in coarse-grained sediments, since the capillary fringe will be significantly shallower than the water table in fine-grained sediments. A detailed GPR survey of the Cape Cod Aquifer imaged the capillary fringe (labeled as the water table although the capillary fringe is likely inches above the water table) as well as lithologic variability in the saturated zone (Figure 11.8) (Knoll et al., 1991). The GPR penetration depth is reduced in the saturated zone because a large portion of the incident energy is reflected from the capillary fringe due to the contrast in dielectric constant. At other sites, low-frequency radar (25 to 200 MHz) has provided deeper penetration and has been used to image lithologic contrasts in the saturated zone, although the trade-off for increased depth is decreased spatial resolution.

GPR has significant potential to image lithologic variations in the unsaturated zone due to the different water retention properties of lithologies. Thus, lithologic heterogeneity may be detected by comparing the results of two radar surveys made under different moisture conditions. Although the dielectric permativity of clay is different than sand with the same moisture content (Figure 11.7), the change is more pronounced when both materials are saturated and subsequently drained, because under these conditions clay has a higher moisture content due to strong capillary forces in the small pore spaces. Greaves et al. (1996) interpreted variations in moisture content from the high-quality multichannel radar data collected at the Chalk River site, Ontario, which were processed by Fisher et al. (1992) (Figure 11.9). The clarity of this processed data shows the ability of multichannel GPR surveys to infer heterogeneities in lithology and water content.

Cross-well radar surveys provide the opportunity to image subsurface features with higher resolution and more control over the path of energy between source and receiver than surface radar surveys. In the cross-well radar configuration, the transmitted radar signal is measured in a receiver well, thus cross-well tomography can estimate both the attenuation and velocity of electromagnetic energy throughout the plane between wells (Olhoeft, 1988). Radar signals are strongly attenuated by clays and saline fluids,

Dielectric Permativity
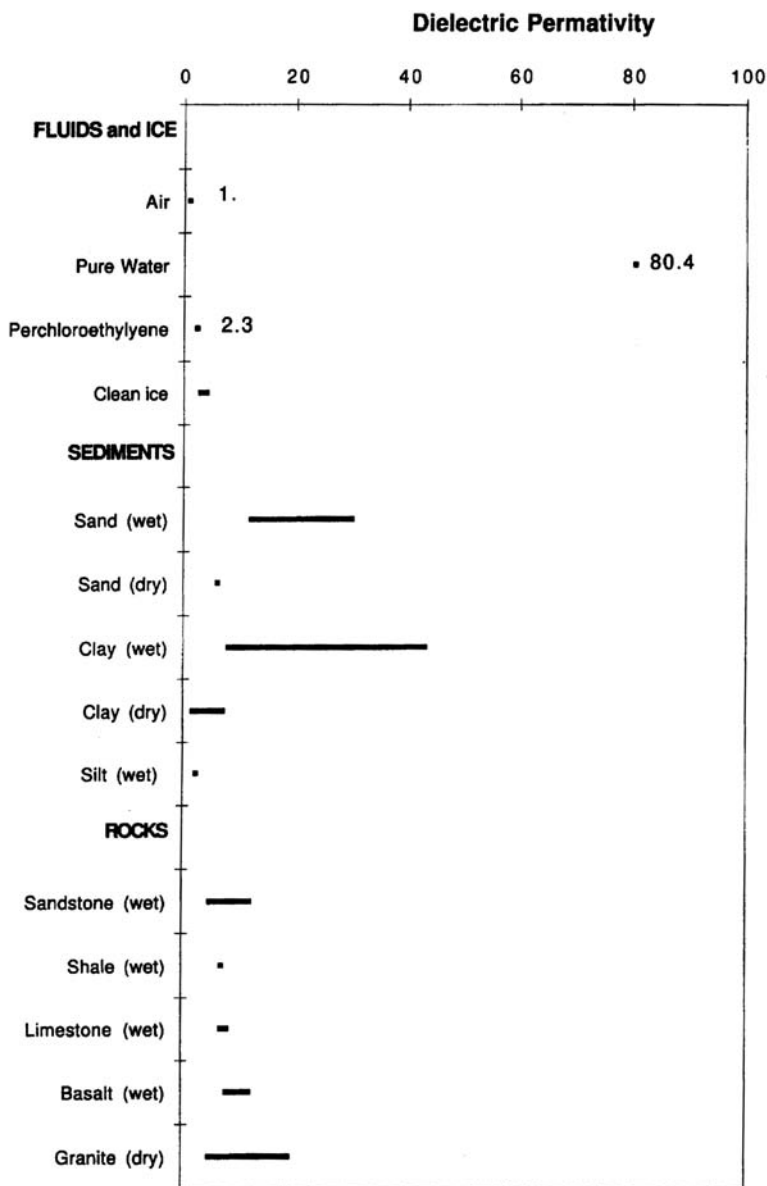
FIGURE 11.7 Plot of relative dielectric permativity values of various materials. (Data from Placzek and Haeni, 1995; Beres and Haeni, 1991; Arcone et al., 1995; Telford et al., 1990.)

thus limiting their use in these environments. This attenuation, however, can be used to identify such zones using short offset cross-well surveys in regions where clay lenses are embedded in a sand matrix. Surface and cross-well radar can be used to dynamically image the location of contaminant plumes (for nonpolar contaminants with permativities very different from that of clean water), as was done in the Borden Aquifer (Greenhouse et al., 1993; Brewster et al., 1992). GPR reflections were observed as the dense nonaqueous phase liquid (DNAPL) pooled on successively deeper low-permeability lithologies. The late time GPR profiles illustrated the average increase in GPR velocity due to the introduction of the low dielectric constant DNAPL into the test region. The dynamic imaging of a tracer that influences the radar propagation velocity or attenuation can provide high resolution estimates of DNAPL saturation,
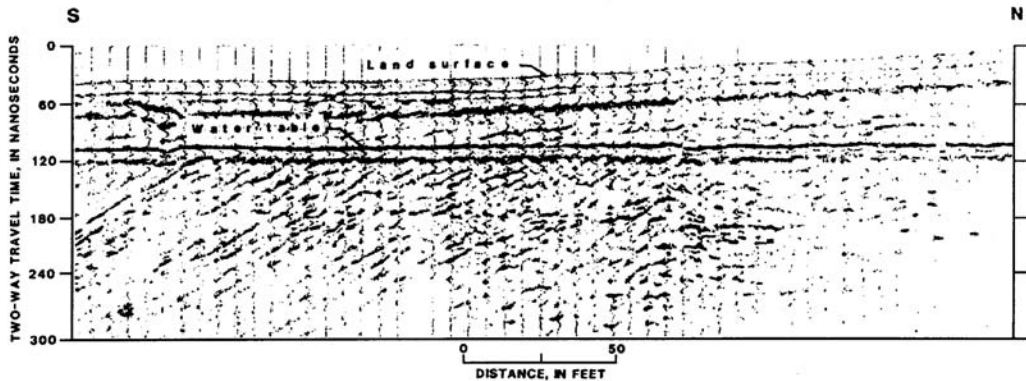
FIGURE 11.8 A single line of processed radar data from a three-dimensional survey over the Cape Cod Aquifer. The land surface and capillary fringe (labeled as water table) are noted on the plot. (From Knoll, M. D., Haeni, F. P., and Knight, R. J. 1991. Characterization of a sand and gravel aquifer using ground-penetrating radar, Cape Cod, Massachusetts. Ed. G. E. Mallard and D. E. Aronson, U. S. Geological Survey toxic substances hydrology program — Proceedings of the technical meeting, March 11-15, 1991, Monterey, CA. With permission.)

which can be used to estimate the hydraulic conductivity structure of the aquifer and enhance our knowledge about transport processes in heterogeneous aquifer environments.

11.5 Cross-Well Seismic Tomography

Cross-well seismic tomography is a relatively new technique that shows great promise for imaging heterogeneities between wells. The philosophy of seismic tomography is similar to that of the computed axial tomography (CAT) scan, except that seismic waves are used to image the subsurface rather than using X-rays to image the human body. The resolution of seismic tomography is much lower than resolutions obtained from medical imaging techniques because much lower frequencies are used for seismic imaging (1 to 20 KHz), and the coverage is limited to a plane between wells, whereas complete three-dimensional coverage is available in a CAT scan.

Cross-well seismic tests involve generating sound waves at specified depth intervals in one well, while hydrophones or geophones measure the propagated wavefield at multiple depths in nearby wells. A variety of sources can be used including: (1) a weight drop, (2) a hydraulic pulse generator, (3) an air pulse, or (4) a piezoelectric bender bar. Piezoelectric sources provide a high-quality, frequency-controlled signal by applying an electrical impulse to material that expands due to this charge. The received wavefield is analyzed for key events, such as the compressional- and shear-wave arrival times at each receiver (Telford et al., 1990) or reflections (Lazaratos et al., 1993).

Cross-well seismic tomography fills a gap between sparse well logs and low resolution surface seismic estimates (Harris et al., 1996). This geophysical method involves estimating seismic properties that best reproduce aspects of the measured wavefield through wave propagation simulations. The best possible vertical resolution is a quarter wavelength (λ), which can be roughly calculated using the dominant frequency (f) of the source as $\lambda = f/(4 \times V)$, where V is the average seismic velocity. This indicates that cross-well seismic tomography has the potential to image lithologic heterogeneities at scales from 10 cm to several meters, for typical sediment velocities (1000 to 2500 m/s) (Telford et al., 1990; Kearey and Brooks, 1991) and source frequencies (1 to 10 KHz).

One seismic tomography approach is to approximate wave propagation through a homogeneous media with embedded scattering objects using the elastic wave equation
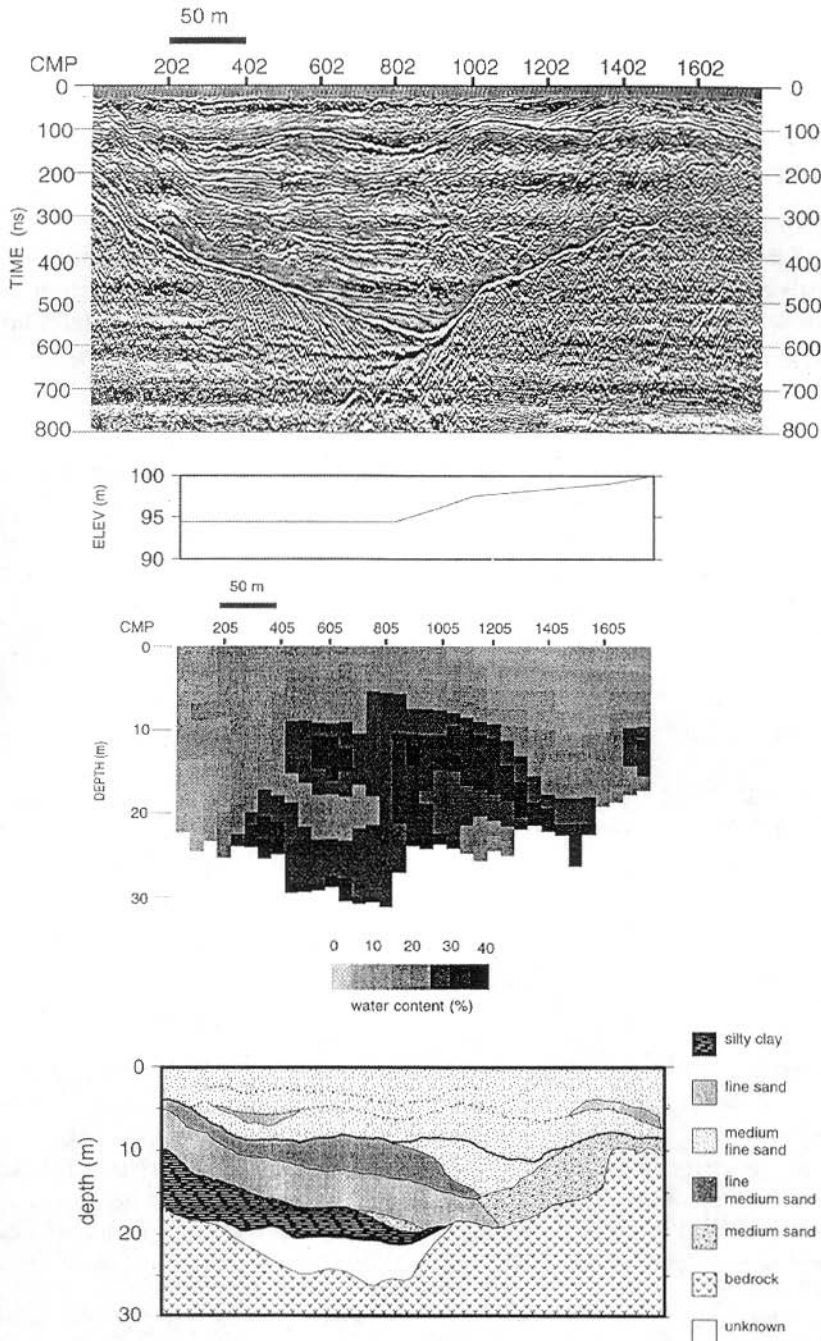
FIGURE 11.9 Plot of processed multiple offset GPR data collected at the Chalk River test site along with interpretations of water content and lithology for the surveyed region. (From Greaves, R. J., Lesmes, D. P., Le, J. M., and Toksoz, M. N. 1996. Velocity variations and water content estimated from multi-offset, ground penetrating radar. *Geophysics.* 61(3):683-695. With permission. Original presentation in Fisher, E., McMechan, G., and Annan, A. P. 1992. Acquisition and processing of wide-aperture ground penetrating radar data. *Geophysics.* 57:495-504.)

$$\rho \frac{\partial^2 \Delta}{\partial t^2} = \left(\lambda + 2\mu\right)\nabla^2 \Delta \tag{10}$$

where the compressional wave velocity $\alpha^2 = \dfrac{(\lambda + 2\mu)}{\rho}$, Δ is the dilatation, or change in volume per unit volume due to strain, λ is Lame's constant, μ is the shear modulus, and ρ is the bulk density. Diffraction tomography provides one method of approximating wave propagation through a background model with embedded point objects, thus heterogeneous media can be split up into different levels of heterogeneities (Harris and Wang, 1996). Reiter and Rodi (1996) presented a nonlinear waveform tomography method that relaxes some of the assumptions necessary for diffraction tomography, however this approach is highly computationally intensive.

11.5.1 Traveltime Tomography

Rather than simulating the dynamics of wave propagation, this approach assumes that waves are of infinite frequency and thus can be represented using rays, which is fairly reasonable for high-frequency cross-well seismic tomography. In this case the goal of the estimation problem is to minimize the absolute or squared difference between measured and simulated traveltimes for a particular type of seismic wave, such as a compressional wave. The wave equation can be simplified to the traveltime equation using the ray theory approximation

$$t_i = \int_{C_i} S\left(x, z\right) d\ell \qquad i = 1, \ldots, M \text{ rays} \tag{11}$$

where the traveltime (t_i) for source-receiver pair (i) is approximated as the integral of slowness (S) = (1/seismic velocity) along the simulated ray path (C_i), and ($d\ell$) is the increment of path length along the ith ray. Rays can be traced using a numerical approximation of Snell's law (which was originally derived for optics)

$$\frac{\sin\phi_1}{V_1} = \frac{\sin\phi_2}{V_2} = \frac{\sin\phi_i}{V_i} = \frac{\sin\phi_{i+1}}{V_{i+1}} \tag{12}$$

where ϕ_i is the incidence angle between the ray in zone i and the normal to an interface separating zone i from zone $i + 1$, and V_i is the seismic velocity in zone i. The top portion of Figure 11.5 illustrates the refraction (or bending) of rays as they travel through progressively higher velocity regions with depth. The traveltime equation is nonlinear because both the ray path and the slowness field are unknown, thus we cannot directly invert this equation to estimate the slowness field. For each iteration, the traveltime equation can be linearized by iteratively tracing ray paths through a known background slowness field (S_0) to calculate slowness perturbations (ΔS).

$$S = S_0 + \Delta S \tag{13}$$

The slowness estimate from the previous iteration is used for the known background field (S_0), except on the first iteration when an arbitrary starting model is used. A critical decision for tomographic inversions is the appropriate parameterization of the slowness model ($S(x, z)$). The most common parameterization for cross-well seismic tomography has been rectangular constant-velocity pixels (S_j) because this simplifies the estimation problem. The solution to the traveltime equation reduces to a linear set of equations that relate traveltime residuals (Δt_i) to slowness perturbations in each of j pixels (ΔS_j).

$$\Delta t_i = \sum_{j=1}^{N} \Delta S_j \ell_{ij} \qquad i = 1, \ldots, M \text{ rays} \tag{14}$$

where the matrix (ℓ_{ij}) has elements of ray path length in each pixel when the slowness field is defined using rectangular pixels.

The slowness perturbations (ΔS_j) are calculated by inverting Equation (14) using an algorithm such as the simultaneous iterative reconstruction technique (McMechan et al., 1987). This algorithm simultaneously back-projects the traveltime residuals along the estimated ray paths and averages the result to determine the slowness perturbation to assign to each cell.

The slowness field can be parameterized in different ways depending on the goal of the inversion. The most common parameterization for seismic tomography is a field of rectangular pixels, which are assigned slowness values using Equations (11) through (14). An alternative approach was presented by Michelena and Harris (1991) who parameterized the slowness inversion using "natural pixels" that account for the data collection geometry in cross-well tomography. These natural pixels are rectangular elements aligned in the direction of wave propagation. This approach can account for the sensitivity of the measured signal at a receiver to some average of the properties along a rectangular or ellipsoidal beam between wells.

For cases where the goal of the geophysical survey is identification of lithologic zones (e.g., sand, gravel, and clay), better results may be obtained by limiting the number of possible slowness values to the number of major lithologies identified in cores. Hyndman and Harris (1996) parameterized a cross-well inversion using a small number of velocity values to estimate the geometry of the primary lithologies of the shallow Kesterson Aquifer in California's San Joaquin Valley. This approach, called the multiple population inversion (MPI), coinverts traveltimes between all available well pairs for the spatial distribution of a small number of seismic slowness populations. This can provide more information about the dominant scale of subsurface heterogeneity than common parameterizations that provide smooth slowness fields. The lithologies and hydraulic parameters for the estimated populations can then be inferred from core data and hydraulic testing, or a co-inversion can be used to simultaneously estimate all parameters of interest as discussed in the next section of this chapter.

The MPI method iteratively assigns pixels to slowness populations based on the histogram of slowness perturbations, which are calculated by inverting Equation (14). By constraining the number of slowness values, this method is less susceptible to ray-based inversion artifacts and can resolve finer-scale sedimentary structures than methods that smooth the slowness field. Hyndman and Harris (1996) showed that the MPI converges to an equal or smaller average traveltime residual than obtained with an approach based on nonzonal parameterization for both a two-dimensional synthetic aquifer, and for a three-dimensional aquifer case from the Central Valley of California.

This approach reduced the artifacts commonly observed in ray-based inversions without smoothing the estimated field, which is the common method of reducing such artifacts. Since tomography is an iterative estimation approach, the zonal slowness model can be used as the background model for an unconstrained velocity inversion. At the Kesterson site, this provided an updated seismic slowness estimate that preserved much of the structure of the zonal slowness estimates with finer-scale resolution (Figure 11.10). These Kesterson tomograms are used in later sections of this chapter for hydraulic and transport property estimation in three dimensions.

11.5.2 Attenuation Tomography

Traveltime tomography only uses one attribute of the measured wavefield, while attenuation tomography attempts to reproduce the wave amplitudes along the paths predicted by traveltime tomography. The philosophy of this inversion is that the bulk of the seismic energy travels near the ray paths predicted for traveltime tomography, thus the amplitude attenuation can also be described along such paths. Thus the attenuation coefficients can be iteratively updated to best reproduce some aspect of the amplitudes measured at the receivers. For example, Quan and Harris (1996) demonstrated an approach that used
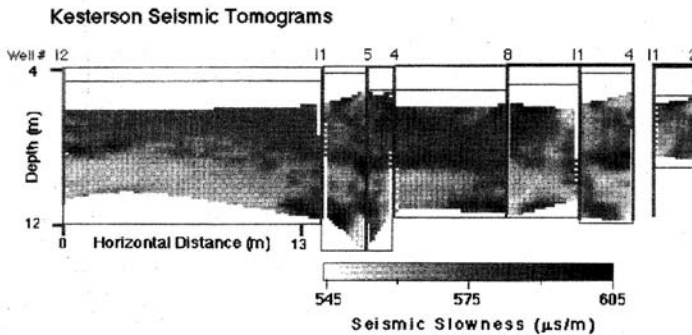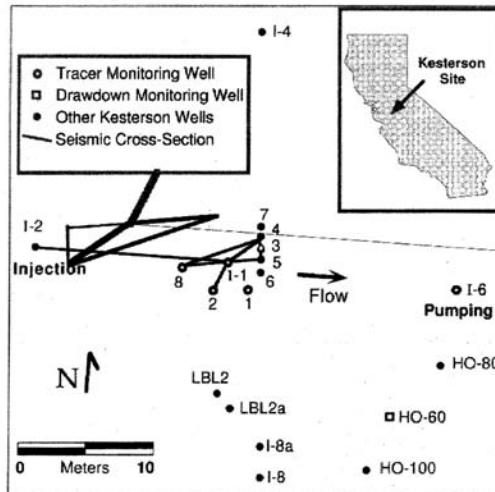
FIGURE 11.10 Location map for the Kesterson Aquifer site along with seismic velocity tomograms estimated by Hyndman and Harris. The low slowness zones were identified as low hydraulic conductivity regions. (From Hyndman, D. W. and Harris, J. M. 1996. Traveltime inversion for the geometry of aquifer lithologies. *Geophysics.* 61(6):1728-1737. With permission.)

the observed frequency downshift in a window around the first arrival to estimate attenuation coefficients for an oil reservoir in Texas. If high-quality signals are measured, significant information can be gained about the lithologic heterogeneities through analysis of both attenuation and traveltime tomography results. Another approach to attenuation tomography is to minimize the difference between the simulated and measured amplitudes of the first arrival peak.

11.6 Combining Seismic and Tracer Data

A great deal of interest has recently focused on combining geophysical and hydrogeologic data for aquifer property estimation. Methods that estimate aquifer properties can benefit from different sampling scales for each type of data. Lithologic cores provide point values. Well-test data provide local information around a well. Cross-well seismic tomography provides dense spatial information for a cross section between wells. Ground-penetrating radar (GPR) provides three-dimensional information about variability in shallow lithologies and fluid saturation. Measured hydraulic heads and tracer concentrations are indirectly related to hydraulic conductivity through the groundwater flow and advection dispersion equations. The scale measured by each type of data is different but overlapping, which provides a suite of information about spatial variability that could not be fully obtained using any single type of data

alone. These independent data types complement each other because they contain information about a range of important aquifer scales.

11.6.1 Geostatistical Methods

Geostatistical methods can be used to combine geophysical and hydrogeologic data. These methods interpolate between available data points using weights that represent modeled spatial correlation structures and the uncertainty in the different measurement methods. Weights are assigned through a covariance matrix that incorporates the inferred spatial correlation of each variable and the interrelationships between variables. An excellent summary of basic geostatistical concepts is provided by Isaaks and Srivastiva (1989) and a suite of geostatistical computer codes is available in Deutsch and Journel (1992).

Geophysical surveys provide an excellent source of soft information for geostatistical estimation methods. Several geostatistical methods have been developed that use surface seismic and radar datasets to provide information about the heterogeneous structure of aquifers and reservoirs. The simplest approach is to *cokrig* the different datasets based on the inferred correlation structures of each variable and cross-correlation between datasets. The benefit of adding high resolution seismic data to well-log data has been demonstrated in a number of studies. For example, Araktingi and Bashore (1992) cokriged three-dimensional surface seismic velocity estimates with porosity measurements and found that even poor-quality seismic estimates provided information about reservoir properties. However, traditional cokriging cannot incorporate nonlinear and potentially non-unique relations between a parameter of interest and the accompanying soft geophysical estimates.

Sequential Gaussian co-simulation (Deutsch and Journel, 1992) provides an additional simple geostatistical method of developing hydraulic conductivity realizations based on hydraulic conductivity measurements as hard data and the slowness estimates as soft data. Sequential Gaussian simulation honors the values at all seismic tomogram locations, as well as the sample probability distribution and variograms of the tomographic slowness estimates. For the Kesterson Aquifer in California's San Joaquin Valley (Hyndman, 1995) found a correlation coefficient of +0.74 between slowness and natural log hydraulic measurements, which establishes the level of control the seismic data has in the geostatistical simulation (Figure 11.11). If this coefficient were +1.0 the seismic data would have a perfect positive linear correlation to log conductivity, with large slowness values indicating large ln(K) values. If this coefficient were 0.0, the seismic data would have no linear correlation to log conductivity and would thus have no influence on the ln(K) estimates. One limitation of this approach is the reliance on a linear relation between parameters.

Figure 11.11 demonstrates the value of adding seismic information to the conductivity estimation problem in cases where significant correlation is observed between slowness and log conductivity. The log conductivity field in Figure 11.11a is a randomly chosen realization generated using sequential Gaussian simulation of log hydraulic conductivity data alone, which results in the corresponding simulated concentration histories. Figure 11.11b illustrates a log conductivity realization generated using sequential Gaussian co-simulation for the same random path that was used to generate Figure 11.11a. For this realization, a slowness realization conditioned on the estimated seismic tomography planes (Figure 11.10) was used as soft data and hydraulic conductivity measurements were hard data. Thus to obtain the conductivity estimates first involved generating a slowness realization using sequential Gaussian simulation, then this field was updated using sequential Gaussian co-simulation to constrain the regions that are near conductivity measurements.

In the Kesterson case, Hyndman (1995) illustrated that the slowness estimates improved the concentration histories at all wells, except for at well 8, relative to the simulations through the estimate based on conductivity data alone (Figure 11.11a). This improvement would likely be more pronounced if fewer conductivity measurements were available. In addition, the simulation with conductivity data alone required estimates of the correlation structure, which in this case were derived from the seismic estimates. Thus the simulations with conductivity alone would have been poorer if no seismic data were available to help infer the correlation structure of subsurface lithologies.
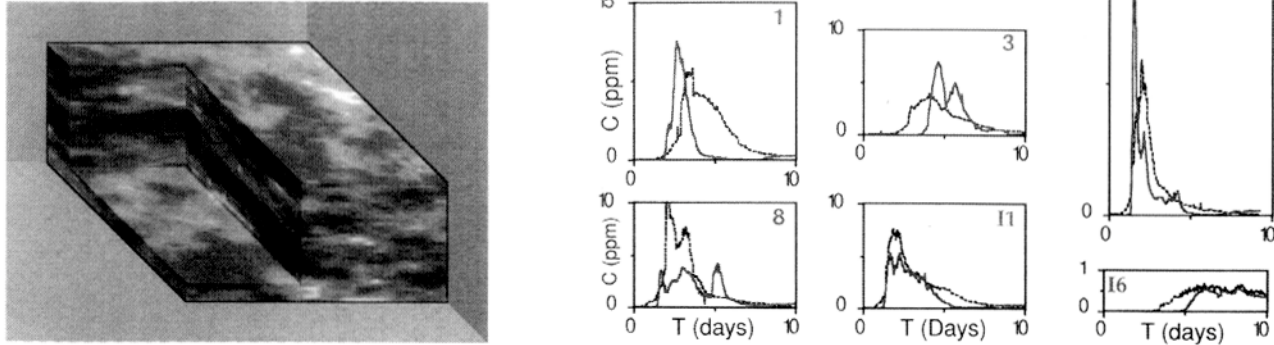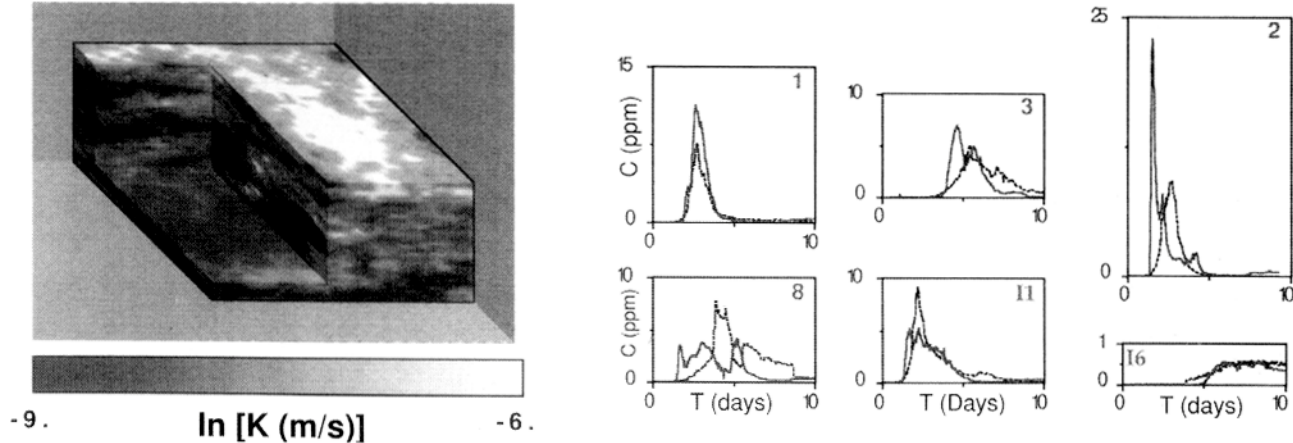
FIGURE 11.11 Comparison of observed (solid) and simulated (dashed) tracer concentration histories for the shallow Kesterson Aquifer, simulated through conductivity realizations (a) generated using conductivity measurements alone, and (b) generated using sequential co-simulation with hard conductivity data and a soft seismic slowness field based on estimates from seismic tomography (see Figure 11.10).

Bayesian geostatistics provides a more flexible alternative to traditional cokriging for incorporating "soft" geophysical data. Copty and Rubin (1995) developed a conditional log hydraulic conductivity distribution based on seismic and lithologic data for a synthetic case study. They first estimated interval seismic velocities using *normal moveout corrections* (described in Section 11.3.1 of this chapter). Then they converted these interval velocities into a prior lithology indicator probability field, which represents the likelihood of each lithology across the survey region, using calibration curves that represented the relation between seismic velocity from well logs and lithology from core samples. This prior distribution was then conditioned on lithology logs from wells to provide an updated probability field. The log hydraulic conductivity pdf was then developed by weighting the conductivity probability distribution function (pdf) for each lithology by the indicator lithology pdf. The results of this study indicate that the best hydraulic conductivity estimates are obtained when both seismic and lithologic information are used. Figure 11.12 compares the true log conductivity model with the expected value of the conditional log conductivity pdf based on all available data. Hubbard et al. (1997) developed a similar *stochastic* technique that estimates binary permeability and saturation fields from GPR data based on petrophysical relationships from Knoll et al. (1995). In their numerical example, the GPR data dramatically improved the estimates relative to those based only on borehole data even with significant measurement error.

Other geostatistical techniques develop a covariance matrix between measurements and properties of interest through approximations of an appropriate governing equation. For example, hydraulic head can be related to hydraulic conductivity using a first-order approximation of the steady-state groundwater flow equation (Hoeksema and Kitanidis, 1984), and solute tracer arrival times can be related to conductivity through a first-order approximation of the advection dispersion equation (Harvey and Gorelick, 1995). The conductivity estimates obtained from such methods can be combined with high resolution seismic velocity estimates from inversion methods like those discussed in Section 11.3 of this chapter. For example, Rubin et al. (1992) and Copty et al. (1993) demonstrated the value of seismic velocity estimates given an assumed relation between seismic velocity and permeability (Marion et al., 1992). They showed that seismic information significantly enhanced estimates of hydraulic properties even in synthetic cases where random error was added to the seismic velocity values, when information about the relation between properties is available. For some sites, it may be possible to infer a relation between seismic and hydraulic properties using core data, which then allows the use of such stochastic estimation methods to estimate the distribution of hydraulic properties across a region. Knowledge of the relationship between seismic velocity and hydrogeologic properties (lithology or permeability), however, is generally unknown at the field scale.

Conditional indicator simulation can be used to combine a wide variety of datasets including geologic knowledge, seismic estimates, geophysical logs, and local lithologic and hydraulic conductivity measurements. This approach involves modeling the spatial correlation structure (i.e., variogram) for several sets of property ranges separated by indicator values, and developing realizations of a property that honor the correlation structure and are conditioned to all measured values. McKenna and Poeter (1995) classified the *hydrofacies* of a site using indicator seismic velocity thresholds. These thresholds were determined using discriminant analysis in a four-dimensional space representing borehole logs (sonic, neutron, electrical conductivity, and natural gamma). This approach allowed for a nonlinear and non-unique relation between hydraulic conductivity and seismic velocity. The concept behind this analysis is to group locations with similar lithologies as indicated by the combination of log responses at each location. Inverse flow modeling was then used to estimate the likely range of hydraulic conductivity for each hydrofacies and eliminate realizations that did not honor the hydraulic head data. This approach was used to delineate the hydrofacies of a fluvial deposit near Golden, Colorado, where the combined analysis of seismic, hydraulic, and geologic data reduced the uncertainty in the hydraulic conductivity estimates. Poeter and McKenna (1995) stressed the importance of incorporating hydraulic data to constrain the conductivity estimates. Using geologic and geophysical data (hydrofacies classification) from wells alone, they found that some of the geostatistical realizations were implausible.
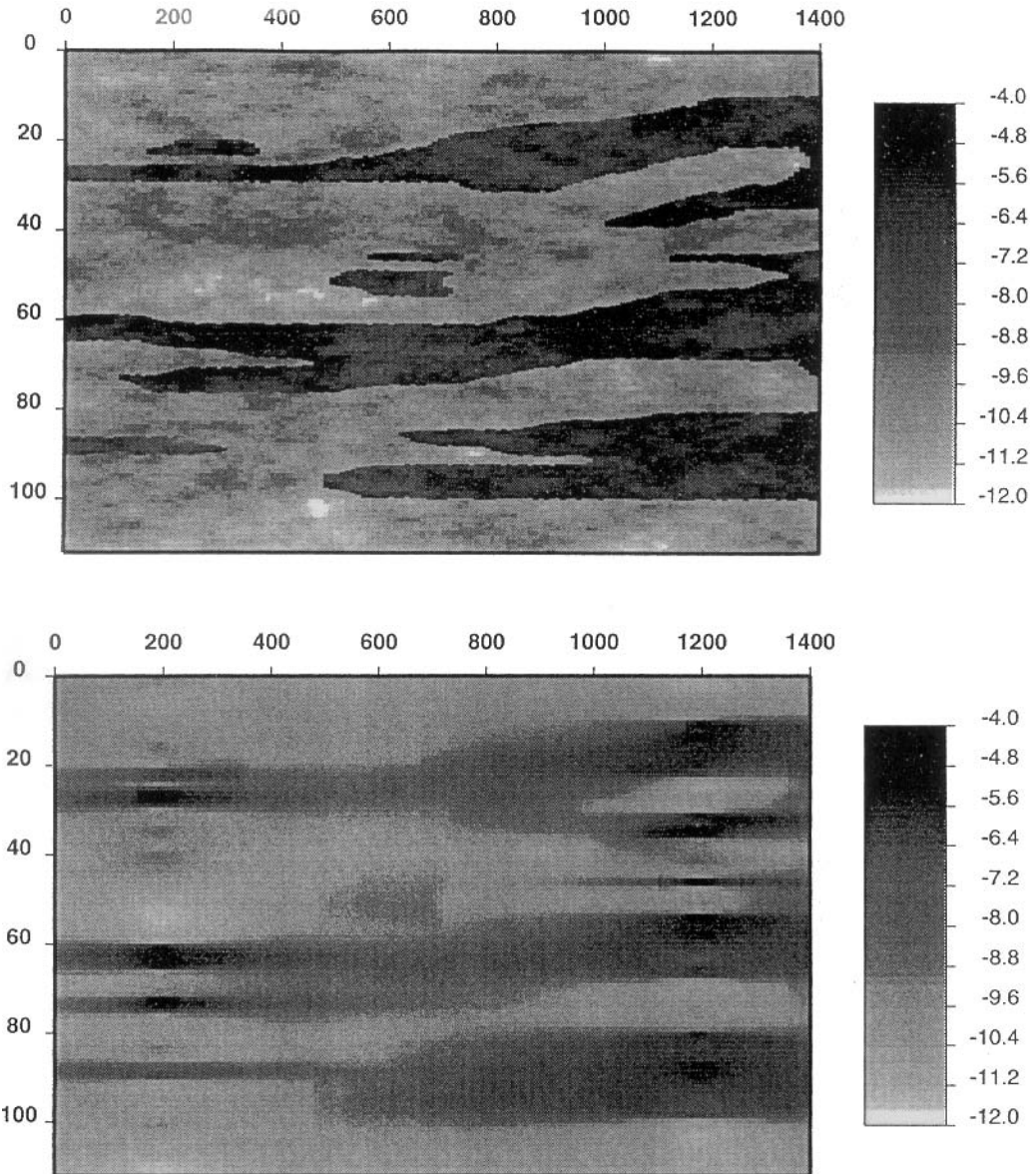
FIGURE 11.12 Schematic log permeability section below Fremont, California, along with the estimated conditional log hydraulic conductivity distribution based on seismic and lithologic data. (From Copty, N. and Rubin, Y. 1995. A stochastic approach to the characterization of lithofacies from surface seismic and well data. *Water Resour. Res.* 31(7):1673-1686. With permission.)

11.6.2 Zonal Inversion Methods

Another approach to combining geophysical and hydrogeologic data involves inferring the large-scale zonation of lithologies and effective hydrogeologic properties that are consistent with seismic, tracer, and hydraulic data. The split inversion method (SIM) was introduced by Hyndman et al. (1994) to co-invert independently collected datasets (e.g., cross-well seismic traveltimes and tracer concentration histories) for the zonation of lithologies, effective hydraulic conductivity values for each zone, and an effective dispersivity for the region. The motivation for the approach is that the combined analysis of seismic

energy and tracer concentrations, which have independently sampled portions of the same environment, should provide better estimates of the geometry of subsurface lithologies and effective zonal properties than obtained with either dataset alone. The SIM does not rely on knowledge of the relationship between seismic velocity and hydraulic conductivity, although it assumes some relationship exists for large-scale lithologic zones.

The ability to represent nonlinear and non-unique relations between these properties was illustrated by the accurate estimates obtained by Hyndman et al. (1994) for two synthetic aquifers with the same seismic profile (a heterogeneous high seismic velocity matrix with embedded low seismic velocity zones) but different hydraulic conductivity profiles. The low velocity zones in one aquifer represented low conductivity zones, while these zones in another aquifer represented both high and low conductivity zones. Thus seismic information alone did not provide enough information to discern the lithologic zonation or hydrogeologic properties, yet the combined use of seismic and tracer data allowed for accurate estimates of both the zonation of lithologies and the hydrogeologic properties of the zones.

The SIM was also applied to the Kesterson Aquifer by Hyndman and Gorelick (1996). The first stage of this approach was to estimate the seismic slowness for the available cross sections between wells as discussed at the end of Section 11.5.1. The correlation structure of the estimated seismic tomograms was then modeled to provide horizontal and vertical slowness variograms for the region. These were used to develop multiple three-dimensional conditional seismic slowness realizations (e.g., Figure 11.13) for a region surrounding the tomograms using sequential Gaussian simulation (Deutsch and Journal, 1992). The SIM was then used to split each slowness realization into three lithologic classes (e.g., Figure 11.13) and to estimate both the effective hydraulic conductivity for each lithology and a single dispersivity value for the entire model domain. The objective function of this inversion was to minimize the squared residual between measured and simulated tracer arrival time quantiles, and drawdown during the tracer test. This study again illustrated the value of seismic tomography for estimating the hydraulic conductivity structure of a heterogeneous aquifer. The seismic slowness estimates provided information about the geometry of lithologic zones and the horizontal and vertical correlation structure of the aquifer lithologies. The drawdown measured during the tracer test provided information about the average regional hydraulic conductivity at the site, while the tracer concentrations provided information about the continuity of hydraulic conductivity across the region of the test.

11.6.3 Estimating the Relation Between Geophysical and Hydrogeologic Properties

One difficulty with incorporating geophysical data into hydrogeologic estimation methods is the uncertain and potentially non-unique relation between estimated geophysical properties (e.g., seismic velocity) and hydrogeologic properties (e.g., hydraulic conductivity). There is no reason to expect a fundamental relation between these properties since seismic velocities do not directly depend on the ability of sediment to transmit fluids. For particular frequencies of seismic energy, however, some researchers have found that fluid flow is quite important to the stiffness of the media, which is related to seismic velocity and attenuation (Bourbie et al., 1987). The effect of fluid moving through pores to accommodate the stress imposed by a sound wave was explored by Dvorkin et al. (1995) as an alternative to the Biot theory.

A primary goal in the field of rock physics (or petrophysics) is identifying relations between rock/sediment properties and measured geophysical signals to enable better predictions of properties in field settings. Petrophysics research has demonstrated empirical relations between seismic velocity and both porosity and clay content (Han et al., 1986). Marion et al. (1992) developed a conceptual mixing model for the relationship between seismic velocity and porosity that involved sand-clay mixtures. The porosity had a minimum when the clay content was large enough to fill in the voids between sand grains, which resulted in a maximum seismic velocity. The porosity increased from this minimum toward both end member lithologies (sand and clay), with a corresponding reduction in the seismic velocity since the velocity of sound is lower through water than it is through sand or clay. Marion's model described one possible process leading to a relation between porosity and velocity. However, laboratory slowness
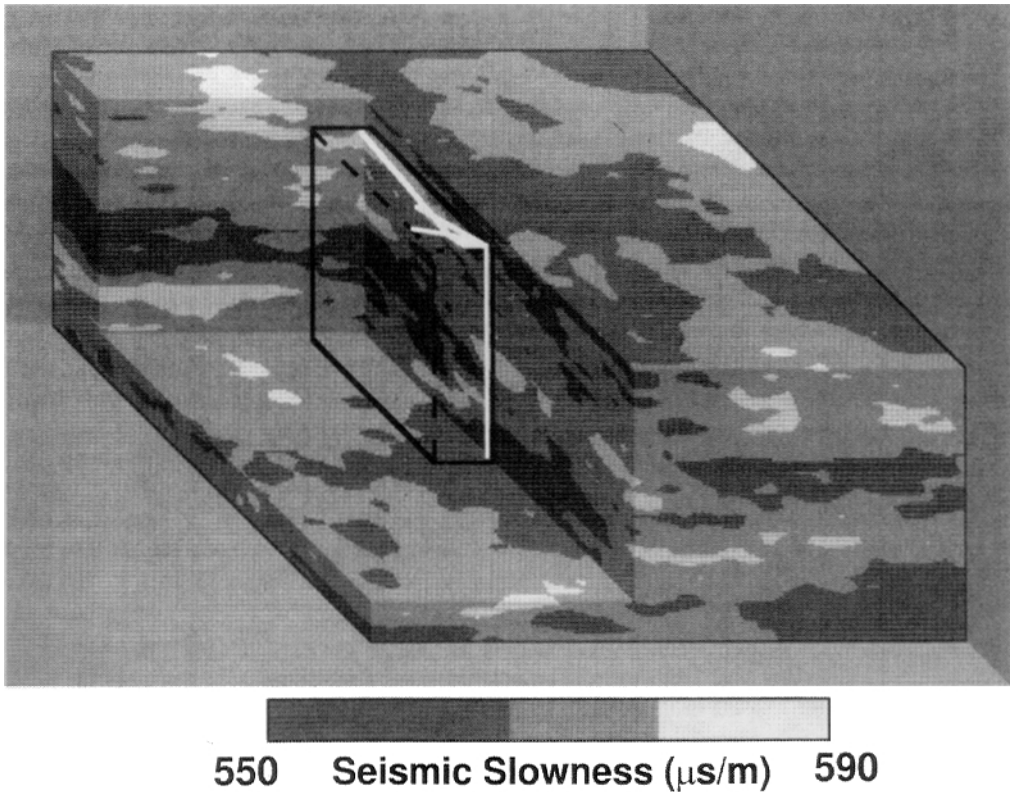
FIGURE 11.14 Zonation of hydraulic conductivity estimates obtained using the split inversion method. For this slowness realization (Figure 11.12), the estimated slowness splits are 569.4 and 587.8 µs/m, and the hydraulic conductivity values are: 1.4×10^{-4} m/s for the black regions; 3.6×10^{-4} m/s for the gray regions; and 5.0×10^{-4} m/s for the white regions.

is a highly site- and scale-specific endeavor since no general relation is expected. The slowness (*S* in *µs/m*) appears to be positively correlated to log hydraulic conductivity (*K* in *m/s*) *for the shallow Kesterson Aquifer* as described with the following relation.

$$\ln(K) = 31.5 + 0.041S \tag{15a}$$

$$\ln(K) = c + b(S - S_{\text{median}}) \tag{15b}$$

where the modified intercept c = (a + bS$_{\text{median}}$).

The coefficients in Equation (15a) are the average slope and intercept for five slowness realizations with standard deviations of 5.8 for the intercept and 0.01 for the slope. The median slowness for each realization was removed prior to estimation using Equation (15b), to reduce the dependence of the average natural log conductivity on the estimated slope. The slope is then a weighting factor that explains the dependence of log hydraulic conductivity on the relative magnitude of seismic slowness (high versus low). In addition, the longitudinal dispersivity was estimated to be 8.3 cm with a standard deviation of 1.4 cm. The fit to the tracer and hydraulic data was significantly enhanced by the addition of the seismic data relative to those obtained using sequential simulation to interpolate between hydraulic conductivity measurements alone. This approach determined the likely conductivity structure of the shallow Kesterson Aquifer and explored the range of likely relations between slowness and conductivity at the field scale.

All the approaches described above indicated that geophysical information has the potential to increase the resolution of hydraulic property estimates. Changes in subsurface lithology are likely to cause a change in both geophysical and hydraulic properties at some scale, thus combining these different datasets provides better estimates than can be obtained with only one dataset. The primary difficulty in direct incorporation of these different datasets is the uncertain link between the measured properties. Different relations are expected for different depositional environments, and these relations may be nonlinear or non-unique if they are measurable at all.

Other geophysical measurements could also be inverted to provide additional information about the hydraulic and lithologic properties of aquifers. For example, seismic attenuation is thought to be correlated to hydraulic conductivity. In all cases, high-quality data are necessary to perform detailed estimates of aquifer properties. This includes accurate surveys of well geometry for all cross-well surveys, since even small location errors may be larger than the variations in the property of interest between the wells. Hydraulic and tracer test data should be sampled with high spatial and temporal resolution to measure the fine scale variations in flow and transport that can be critical to solute transport predictions.

For Further Information

Peck et al. (1988) provide a detailed overview of simulation optimization methods and other parameter estimation methods. Telford et al. (1990) provide an excellent overview of introductory geophysical methods, but have a limited discussion of tomography or GPR. Ward (1990) compiled a large set of papers relating to environmental geophysics, including case studies and theoretical investigations. McLaughlin and Townley (1996) provide an excellent review of inverse methods for estimating groundwater flow and solute transport properties.

References

Adams, E. E. and Gelhar, L. 1992. Field study of dispersion in a heterogeneous aquifer, 2, Spatial moments analysis. *Water Resour. Res.* 28(12):3293-3307.

Araktingi, U. G. and Bashore, W. M. 1992. Effects of properties in seismic data on reservoir characterization and consequent fluid-flow predictions when integrated with well logs, *Society of Petroleum Engineers* (SPE 24752, presented at 67th Annual SPE conference in Washington, D.C.).

Arcone, S. A., Lawson, D. E., and Delaney, A. J. 1995. Short-pulse radar wavelet recovery and resolution of dielectric contrasts within englacial and basal ice of Matanuska Glacier, Alaska, USA. *Journal of Glaciology.* 41(137):68-86.

Benson, S. M. 1988. Characterization of the hydrogeologic and transport properties of the shallow aquifer under Kesterson Reservoir. Ph.D. thesis. University of California Berkeley.

Benson, S. M., White, A. F., Halfman, S., Flexser, S., and Alavi, M. 1991. Groundwater contamination at the Kesterson Reservoir, California: 1. Hydrogeologic setting and conservative tracer transport. *Water Resour. Res.* 27(6):1071-1084.

Beres, Milan and Haeni, F. P. 1991. Application of ground penetrating radar methods in hydrologic studies. *Ground Water.* 29(3):375-386.

Boggs, M. J., Yound, S. C., Beard, L. M., Gelhar, L. W., Rehfeldt, K. R., and Adams, E. E. 1992. Field study of dispersion in a heterogeneous aquifer 1. Overview and site description. *Water Resour. Res.* 28(12):3281-3291.

Bourbie, T., Coussy, O., and Zinszner, B. 1987. *Acoustics of porous media.* Gulf Publishing Company, Houston.

Brewster, M. L., Annan, A. P., and Redman, J. D. 1992. GPR monitoring of DNAPL migration in a sandy aquifer. *Geological Survey of Finland, Special Paper* 16:185-190.

Claerbout, J. 1992. *Earth Sounding Analysis: Processing Versus Inversion.* Blackwell Scientific, Boston, MA, 304 p.

Copty, N. and Rubin, Y. 1995. A stochastic approach to the characterization of lithofacies from surface seismic and well data. *Water Resour. Res.* 31(7):1673-1686.

Copty, N., Rubin, Y., and Mavko, G. 1993. Geophysical-hydrological identification of field permeabilities through Bayesian updating. *Water Resour. Res.* 29(8):2813-2825.

Deutsch, C. V. and Journel, A. G. 1992. *Geostatistical Software Library and User's Guide.* Oxford University Press, Oxford.

Dvorkin J., Mavco, G., and Nur, A. 1995. Squirt flow in fully saturated rocks. *Geophysics.* 60(1):97-107.

Fisher, E., McMechan, G., and Annan, A. P. 1992. Acquisition and processing of wide-aperture ground penetrating radar data. *Geophysics.* 57:495-504.

Freyberg, D. L. 1986. A natural gradient experiment on solute transport in a sand aquifer 2. Spatial moments and the advection and dispersion of nonreactive tracers. *Water Resour. Res.* 22(13):2031-2046.

Garabedian, S. P., LeBlanc, D. R., Gelhar, L. W., and Celia, M. A. 1991. Large-scale natural gradient tracer test in sand and gravel, Cape Cod, Massachusetts. 2. Analysis of spatial moments for a nonreactive tracer. *Water Resour. Res.* 27(5):911-924.

Gelhar, L. W. 1993. *Stochastic Subsurface Hydrology.* Prentice Hall, Englewood Cliffs, NJ, 390 p.

Gelhar, L. W. and Axness, C. L. 1983. Three dimensional stochastic analysis on macrodispersion in aquifers. *Water Resour. Res.* 19(1):161-180.

Greaves, R. J., Lesmes, D. P., Le, J. M., and Toksoz, M. N. 1996. Velocity variations and water content estimated from multi-offset, ground penetrating radar. *Geophysics.* 61(3):683-695.

Greenhouse, J., Brewster, M., Schneider, G., Redman, G., Annan, P, Olhoeft, G., Lucius, J., Sander, K., and Mazzella, A. 1993. Geophysics and solvents: the Borden experiment. *The Leading Edge.* 12(4):261-267.

Han, D., Nur, A., and Morgan, D. 1986. Effect of porosity and clay content on wave velocity in sandstones. *Geophysics.* 51(11):2093-2107.

Harris, J. M., Nolen-Hoeksema, R. C., Langan, R. T., Schaack, M. V., Lazaratos, S. K., and Rector III, J. W. 1996. High resolution cross-well imaging of a west Texas carbonate reservoir: Part 1 — Project summary and interpretation. *Geophysics.* 60(3):667-681.

Harris, J. M. and Wang, G. Y. 1996. Diffraction tomography for inhomogeneities in layered background medium. *Geophysics.* 61(2):570-583.

Harvey, C. F. and Gorelick, S. M. 1995. Mapping hydraulic conductivity: Sequential conditioning with measurements of solute arrival time, hydraulic head, and local conductivity. *Water Resour. Res.* 31(7):1615-1626.

Hess, K. M., Wolf, S. H., and Celia, M. A. 1992. Large-scale natural gradient tracer test in sand and gravel, Cape Cod, Massachusetts. 3. Hydraulic conductivity variability and calculated macrodispersivities. *Water Resour. Res.* 28(8):2011-2027.

Hoeksema, R. J. and Kitanidis, P. K. 1984. An application of the geostatistical approach to the inverse problem in two-dimensional groundwater modeling. *Water Resour. Res.* 20(7):1003-1020.

Hubbard, S. S., Rubin, Y., and Majer, E. 1997. Ground-penetrating-radar-assisted saturation and permeability estimation in bimodal systems. *Water Resour. Res.* 33(5):971-990.

Hyndman, D. W. 1995. Inferring aquifer properties using seismic travel times and tracer concentrations. PhD dissertation, Stanford University.

Hyndman, D. W. and Gorelick, S. M. 1996. Mapping lithologic and transport properties in three dimensions using seismic and tracer data: the Kesterson aquifer. *Water Resour. Res.* 32(9):2659-2670.

Hyndman, D. W. and Harris, J. M. 1996. Traveltime inversion for the geometry of aquifer lithologies. *Geophysics.* 61(6):1728-1737.

Hyndman, D. W., Harris, J. M., and Gorelick, S. M. 1994. Coupled seismic and tracer test inversion for aquifer property characterization. *Water Resour. Res.* 30(7):1965-1977.

Hyndman, D. W., Gorelick, S. M., and Harris, J. M. Inferring the relationship between seismic slowness and hydraulic conductivity in heterogeneous aquifers. Submitted to *Water Resour. Res.*

Isaaks, E. H. and Srivastiva, R. M. 1989. *An Introduction to Applied Geostatistics*. Oxford Press, New York, 561 p.

Kearey, P. and Brooks, M. 1991. *An Introduction to Geophysical Exploration*, 2nd Ed., Blackwell Scientific Publications. Oxford.

Knoll, M. D., Haeni, F. P., and Knight, R. J. 1991. Characterization of a sand and gravel aquifer using ground-penetrating radar, Cape Cod, Massachusetts. Ed. G. E. Mallard and D. E. Aronson, U. S. Geological Survey toxic substances hydrology program — Proceedings of the technical meeting, March 11-15, 1991, Monterey, CA.

Knoll, M., Knight, R., and Brown, E. 1995. Can accurate estimates of permeability be obtained from measurements of dielectric properties? *SAGEEP Annual Meeting Extended Abstracts*, Orlando, FL.

Lazaratos, S. K., Rector III, J. W., Harris, J. M., and Van Schaack, M., 1993 High-resolution, cross-well reflection imaging: potential and technical difficulties. *Geophysics.* 58(9):1270-1280.

LeBlanc, D. R., Garabedian, S. P., Hess, K. M., Gelhar, L. W., Quadri, R. D., Stollenwerk, K. G., and Wood, W. W. 1991. Large-scale natural gradient tracer test in sand and gravel, Cape Cod, Massachusetts. 1. Experimental design and observed tracer movement. *Water Resour. Res.* 27(5):895-910.

Marion, D., Nur, A., Yin, H., and Han, D. 1992. Compressional velocity and porosity in sand-clay mixtures. *Geophysic.* 57(4):554-563.

McKenna, S. A. and Poeter, E. P. 1995. Field example of data fusion in site characterization, *Water Resour. Res.* 31(12):3229-3240.

McLaughlin, D. and Townley, L. R. 1996. A reassessment of the groundwater inverse problem. *Water Resour. Res.* 32(5):1131-1161.

McMechan, G. A., Harris, J. M., and Anderson, L. M., Cross-hole tomography for strongly variable media with applications to scale model data. *Bol. Seis. Soc. Am.,* 77(6):1945-1960.

Michelena, R. and Harris, J. 1991. Tomographic traveltime inversion using natural pixels, *Geophysics.* 56(5):635-644.

Molz, F. 1994. Borehole flowmeters: field application and data analysis. *Journal of Hydrology.* 163(3-4): 347-371.

Olhoeft, G. R. 1988. Interpretation of hole-to-hole radar measurements. *Proceedings of the Third Technical Symposium on Tunnel Detection — Golden, Colorado.* 616-626. (also in USGS *Open File Report 90-0414*).

Peck, A., Gorelick, S., Marsily, G. D., Foster, S., and Kovalevsky, V. 1988 *Consequences of Spatial Variability in Aquifer Properties and Data Limitations for Groundwater Modeling Practice.* IAHS Publication 175, 271 p.

Placzek, G. and Haeni, F. P. 1995. Surface-geophysical techniques used to detect existing and infilled scour holes near bridge piers. U.S. Geological Survey Water Resources Investigations Report 95-4009, Denver, CO.

Poeter, E. P. and McKenna, S. A. 1995. Reducing uncertainty associated with ground-water flow and transport predictions. *Ground Water.* 33(6):899-904.

Quan, Y. and Harris, J. 1998. Seismic attenuation tomography using the frequency shift method. *Geophysics.* In press.

Rea, J. and Knight, R. 1998. Geostatistical analysis of ground penetrating radar data: a means of describing spatial variations in the subsurface. *Water Resour. Res.* 34(3).

Rehfeldt, K. R., Boggs, M. J., and Gelhar, L. 1992. Field study of dispersion in a heterogeneous aquifer, 3, Geostatistical analysis of hydraulic conductivity. *Water Resour. Res.* 28(12):3281-3291.

Reiter, D. T. and Rodi, W. 1996. Nonlinear waveform tomography applied to crosshole seismic data. *Geophysics.* 61(3):902-913.

Rubin, Y., Mavko, G., and Harris, J. 1992. Mapping permeability in heterogeneous aquifers using hydrologic and seismic data. *Water Resour. Res.* 28(7):1809-1816.

Sanchez-Villa, X., Carrera, J., and Colominas, I. 1992. Directional effects on convergent flow tracer tests, in *Tracer Hydrology*, H. Hotzl and A. Werner Eds. A. A. Balkema, Rotterdam, 407-414.

Steeples, D. W. and Miller, R. D. 1990. Seismic reflection methods applied to engineering, environmental, and groundwater problems, in *Geotechnical and Environmental Geophysics*, Ward, S. (Ed). Vol. I, Investigations in Geophysics, SEG Series. SEG, Tulsa, OK.

Sudicky, E. A. 1986. A natural gradient experiment on solute transport in a sand aquifer: spatial variability of hydraulic conductivity and its role in the dispersion process. *Water Resour. Res.* 22(13):2069-2082.

Telford, W. M., Geldart, L. P., and Sheriff, R. E. 1990. *Applied Geophysics*. 2nd Ed. Cambridge University Press, Cambridge.

Vogelsang, D. 1995. *Environmental Geophysics: A Practical Guide*. Springer Verlag, Berlin, 173 p.

Ward, S. (Ed). 1990. *Geotechnical and Environmental Geophysics*, Vols. I, II, and III. Investigations in Geophysics, SEG Series. SEG, Tulsa, OK.

Glossary

Aquifer Properties   Characteristics of an aquifer that control the transport of fluid and solutes.

Attenuation   Reduction in signal amplitude due to energy loss (because of reflection, dispersion, etc.).

Cokriging   Interpolating between a series of measured values based on multiple types of data and the modeled spatial correlation structures of each data type.

Common Midpoint Gather   A collection of geophysical data with source and receiver at different locations along a line, each equidistant from a central midpoint. If subsurface reflectors are flat and layered with homogeneous seismic properties in each layer, each trace sampled from the common midpoint gather will be reflected from the same point, which will be directly below the chosen common midpoint.

Converging   Iteratively approaching a value or solution with successively smaller changes between iterations.

Global Minimum   The smallest value of an objective function for all parameter combinations.

Hydrofacies   Lithologies categorized by their flow and transport properties.

Inversion   Iterative adjustment of model parameters to minimize the difference between simulated and observed values (e.g., hydraulic head or tracer concentrations).

Non-uniqueness   For optimization problems, this condition indicates that multiple sets of parameters provide similar minimum objective values, thus no global minimum exists and the optimal parameter set cannot be determined using the pπrovided data alone.

Normal Moveout Correction (NMO)   Adjusting seismic or radar velocity estimates to flatten the parabolic appearance of reflectors due to offset between sources and receivers.

Objective Function   A mathematical function that is minimized or maximized (i.e., the residual between measured and simulated head and concentration values for parameter estimation, or total cost for managing aquifer pumping).

Offset   Horizontal distance between a source and a receiver.

Optimization   Finding parameters that achieve the optimal objective value (minimum or maximum).

Pump Tests or Slug Tests   Methods of estimating hydraulic conductivity near a well by pumping water into or out of the well and measuring the hydraulic response. A slug test involves rapid removal or addition of a small volume of water (or a solid slug that displaces water), while a pump test involves continuous pumping of water for a period of time.

Reflection   Energy that bounces back from a surface due to a change in physical properties, such as seismic impedance in the case of sound waves.

Refraction   Energy that bends according to changes in the velocity of sound waves in adjacent media. The angle can be described according to Snell's law, which is included in the text of this chapter.

Regression   Iteratively adjusting the parameters of a model to minimize (or in some cases maximize) the objective value.

Simulation-Optimization Methods   These methods estimate optimal parameter values by applying an optimization package, such as one based on nonlinear least-squares regression, to physically based models of groundwater flow and solute transport to obtain the most desirable system response.

Slowness   1/seismic velocity. This terminology is used because the traveltime equation is solved using slowness.

Solute   A dissolved substance (e.g., a contaminant or a tracer in groundwater).

Spatial Moments   Statistical representations for the location, spread, and shape of a solute plume. In normalized forms, the zeroth moment defines the total mass, the first moment defines the location of the center of mass, and the second moment defines the spread of the plume.

Stacking   Averaging multiple radar or seismic traces collected from the same location to improve the signal-to-noise ratio.

Traveltime Tomography   Estimating the seismic velocity structure between wells based on the time for a type of wave (e.g., a compressional- or P-wave) to travel from the source to the receiver.

12

Geostatistics: Interpolation and Inverse Problems

Peter K. Kitanidis
Stanford University

12.1 Introduction

Hydrogeologic applications often require the estimation of functions from data. For example:

Interpolation From observations of water elevation at a number of wells, draw the contour map of the water table in an aquifer.

Inverse Problems Estimate transmissivity from head observations and other information.

Most applications involve quantities that vary in space but the methods may also apply to quantities that vary in time. A variety of methods have been developed to address such problems in hydrogeology but also elsewhere, interpolation and inverse problems being ubiquitous. Methods include splines (de Boor, 1978), objective analysis (Thiebaux and Pedder, 1987), etc. There are considerable formal similarities among available methods (Cressie, 1993), but even formally equivalent methods may differ in emphasis or in how they are applied and understood. The focus here is on *geostatistics* (Matheron, 1971), a practical methodology popular in hydrology and other geophysical sciences.

We will put in a nutshell results useful in solving interpolation and inverse problems. The presentation assumes familiarity with linear algebra and probability theory.

12.2 Preliminaries

12.2.1 Model

A model of variability is essential for the solution of interpolation problems. The unknown function, denoted as z, is a function of variable x (space and/or time coordinates). A practical and versatile model is:

$$z(x) = \underbrace{c(x) + \sum_{k=1}^{p} f_k(x)\beta_k}_{\text{deterministic}} + \underbrace{\varepsilon(x)}_{\text{random}} \tag{1}$$

where $c(x)$ and $f_1(x)$, ..., $f_p(x)$ are known functions. The bold letter x represents an array of point coordinates. The coefficients β_1, ..., β_p are unknown. The term $\varepsilon(x)$ is a *random function* with zero expected value.

The *deterministic part*, known also as the *mean*, or *drift*, or the *trend*, represents information about large-scale trends, zones, and other features that can be represented through deterministic functions. This part may include parameters β to be "fitted" from the data. However, a crucial requirement is that the number of β coefficients must be much smaller than the number of observations. The *random part* represents erratic, complex, and usually small-scale variability about which we know too little to model through deterministic functions but enough to describe through statistics (or probabilistic averages). In addition to stating that the expected value of ε is zero, we specify a covariance function, variogram, or generalized covariance function. These statistics furnish information about the variance and the degree of continuity of ε and are needed to obtain estimates with error bars.

12.2.2 Stochastic Approach

The objective is to find a function z, a hydrogeologic variable, given limited information. If nothing were known about z, the search should be over an ensemble that includes all functions. As more information becomes available, many of these functions are eliminated or become less likely. The stochastic approach involves assigning probabilities to these functions that form the ensemble.

Geostatistical estimation is for convenience broken into two phases:

Structural analysis. The search is narrowed down to functions that share certain characteristics, collectively known as *structure*, such as continuity and smoothness, scale of variability, etc. Information about the structure is represented here through equations for the deterministic part and the covariance function of the stochastic part (Equation [1]).

Conditioning. The search is further narrowed by eliminating functions that are not consistent with observations.

12.2.3 Least Squares

Consider the interpolation problem: estimate the value of z at a location x_0 from observations at locations x_1, ..., x_n. One approach is to fit parameters β_1, ..., β_p, and $z(x_0)$ from data $z(x_1)$, ..., $z(x_n)$ using Equation (1) and a weighted least squares criterion. Using this intuitive and time-honored approach, one obtains an estimate and its mean square error of estimation. It turns out that formally this is equivalent to the methods of linear geostatistics. However, in implementation, there are considerable differences between some trend analysis approaches (traditionally associated with least squares fitting) and geostatistics, including:

- Trend analysis attempts to capture the structure of the function through an elaborate deterministic part, whereas the random part is represented as lacking spatial correlation. However, the deterministic part is not well suited to represent information about small-scale and erratic variability.

In geostatistics, such variability may be represented through the structure of the random part that is appropriately parameterized through a covariance function.

- In geostatistics, a central issue is the selection of the covariance function or the variogram, which is accomplished in the phase that is known as structural analysis.

- In geostatistics, because one is not really interested in estimating the β coefficients, one may use simplified versions of the covariance function, which is what variograms and generalized covariance functions are. One can also use algorithms that are well suited for interpolation.

12.3 Ordinary Kriging

12.3.1 Interpolation

In many applications, it is appropriate to adopt the simplest and most conservative representation of the mean, which is as an unknown constant, i.e., referring to Equation (1),

$$z(x) = \beta + \varepsilon(x) \tag{2}$$

The information about the function structure is described only through the *variogram* (used instead of the traditional term *semivariogram*) defined as:

$$\gamma(x - x') = \frac{1}{2} E\left[\left(z(x) - z(x') \right)^2 \right] \tag{3}$$

where E denotes a probabilistic average or expected value. *Ordinary kriging*, the most common geostatistical method, gives an estimate and its mean square error of estimation of z at any location x_0 given the variogram γ and n observations $z(x_1)$, $z(x_2)$, …, $z(x_n)$. To infer the unmeasured value of z at a location x_0, a *linear estimator* is adopted, i.e., formula:

$$\hat{z}_0 = \sum_{i=1}^{n} \lambda_i z(x_i) \tag{4}$$

The weights λ_1, …, λ_n are chosen based on the requirements of (1) *unbiasedness*, i.e., the expected value of the estimation error $\hat{z}_0 - z(x_0)$ must be zero and (2) *best* or minimum variance estimation, i.e., the expected value of the square error must be as small as possible. These requirements result in a system of $n + 1$ linear equations with $n + 1$ unknowns, known as the kriging system:

$$-\sum_{j=1}^{n} \lambda_i \gamma(x_i - x_j) + \nu = -\gamma(x_i - x_0), \quad i = 1, 2, …, n$$

$$\sum_{j=1}^{n} \lambda_j = 1 \tag{5}$$

where ν is a Lagrange multiplier. The mean square error of estimation is given by the formula:

$$\sigma_0^2 = E\left[\left(\hat{z}_0 - z(x_0) \right)^2 \right] = -\nu + \sum_{i=1}^{n} \lambda_i \gamma(x_i - x_0) \tag{6}$$

Often, the reliability of the estimate is presented in the form of a 95% *confidence interval*. This is the interval that contains the actual value of $z(x_0)$ with probability 0.95. Under normality, i.e., Gaussian distribution of the estimation errors (a common implicit assumption in linear estimation), the confidence interval is approximately $[\hat{z}_0 - 2\sigma_0, \ \hat{z}_0 + 2\sigma_0]$.

In the *global neighborhood* approach, all observations are used. In the *local neighborhood* approach, one uses a subset of all observations; one may use a fixed number of observations that are closest to the point x_0, all observations within a certain radius, or a combination of these strategies. The global neighborhood approach is recommended except in special cases such as when there is a large number of observations. If the local neighborhood approach is used for interpolation, care should be taken to avoid introducing inadvertent discontinuities in the estimate.

12.3.2 Function Estimate

An expression for the estimate at any point x that is useful in interpolation and graphing a function is the following:

$$\hat{z}(x) = -\sum_{i=1}^{n} \gamma(x - x_i)\xi_i + \hat{\beta} \tag{7}$$

where the ξ coefficients and $\hat{\beta}$ are found from the solution of a single system of $n + 1$ equations with $n + 1$ unknowns:

$$-\sum_{j=1}^{n} \gamma(x_i - x_j)\xi_j + \hat{\beta} = z(x_i), \quad \text{for} \quad i = 1, \ldots, n$$

$$\sum_{j=1}^{n} \xi_j = 0 \tag{8}$$

This approach is useful when performing kriging on a fine grid, and expression (7) reveals how the form of the estimate depends on the form of the variogram.

The conditional covariance function R_c (i.e., the covariance function given the data) is not stationary and may be obtained explicitly as follows:

1. Form matrix of coefficients of system (8)

$$\begin{bmatrix} Q & X \\ X^T & 0 \end{bmatrix} \tag{9}$$

   where Q is matrix n by n with $Q_{ij} = -\gamma(x_i - x_j)$, X is n by 1 with $X_i = 1$, and exponent T denotes matrix transpose.

2. Compute matrix inverse

$$\begin{bmatrix} Q & X \\ X^T & 0 \end{bmatrix}^{-1} = \begin{bmatrix} P & a \\ a^T & B \end{bmatrix}$$

   where P is n by n, a is n by 1 and B is a scalar.

3. The conditional covariance function is given from:

$$R_c\left(x,x'\right)=-\gamma\left(x-x'\right)-\sum_{i=1}^{n}\sum_{j=1}^{n}\gamma\left(x-x_i\right)P_{ij}\gamma\left(x_j-x'\right)-B+\sum_{j=1}^{n}a_j\gamma\left(x_j-x'\right)+\sum_{i=1}^{n}\gamma\left(x-x_i\right)a_i \quad (10)$$

The mean square error is given from this formula by setting $x' = x$.

For illustration, the true function $z(x)$, the best estimate $\hat{z}(x)$, and the 95% confidence interval are shown in figure 12.1, for a hypothetical case of a function of one variable. The data are indicated by small open circles and the variogram $\gamma(h) = 1.4h$ was used for estimation. The best estimate reproduces the observations and is a simplified version of the actual function. The 95% confidence interval is indicative of the potential error (difference between estimate and actual function).
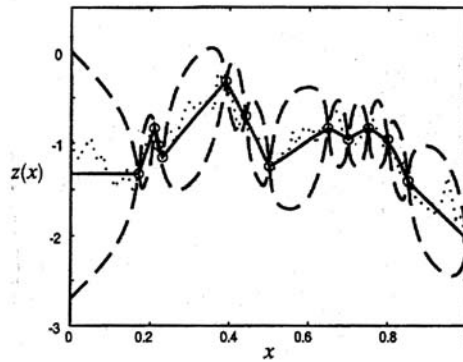
FIGURE 12.1 Best estimate (solid line), 95% confidence interval (dashed lines), versus actual (fine dotted line) z function.

12.3.3 Conditional Realizations

It is often useful to generate a large number of equally likely (given the available information) possible solutions in order to perform probabilistic risk analysis. These solutions are called *conditional realizations* or conditional sample functions or conditional simulations. The average of the ensemble of all conditional realizations is equal to the best estimate and the covariance is equal to the conditional covariance function. We focus on the prevalent case of Gaussian realizations.

12.3.3.1 Point Simulation

Step 1. Generate an unconditional realization $z(x;k)$, with zero mean, consistent with the assumed variogram (where the index k is used to remind that there are many possible solutions). This is a technical issue (Mantoglou and Wilson, 1982; Tompson et al., 1989; Dykaar and Kitanidis, 1992; Robin et al., 1993; Gutjahr et al., 1993) that goes beyond the scope of this chapter.

Step 2. Generate a conditional realization $z_c(x_0;k)$ at a given point x_0,

$$z_c\left(x_0;k\right)=z\left(x_0;k\right)+\sum_{i=1}^{n}\lambda_i\left(z\left(x_i\right)-z\left(x_i;k\right)\right) \quad (11)$$

where $z(x_i)$ is actual observation and the λ coefficients are the same ones that were used to obtain the best estimate.

12.3.3.2 Function Simulation

The functional form of $z_c(x;k)$ is developed similarly as the best estimate in Section 12.3.2. First, generate an unconditional realization $z(x;k)$, with zero mean. Then

$$z_c(x;k) = -\sum_{j=1}^{n} \gamma(x - x_j)\xi_j + \hat{\beta} + z(x;k) \qquad (12)$$

where the ξ and $\hat{\beta}$ coefficients are found from system:

$$-\sum_{j=1}^{n} \gamma(x_i - x_j)\xi_j + \hat{\beta} = z(x_i) - z(x_i;k), \quad \text{for} \quad i = 1, \ldots, n$$

$$\sum_{j=1}^{n} \xi_j = 0 \qquad (13)$$

For illustration, three conditional realizations (out of many possible) of function $z(x)$ are shown together with the best estimate $\hat{z}(x)$ in figure 12.2 for the same case as Figure 12.1. The conditional realizations have the structure described by the variogram and also reproduce the observations. Each conditional realization is equally likely to be the sought-after function. The average of all realizations is the estimate, which is a smooth function as shown in Figure 12.2. At any point x, 95% of the conditional realizations should fall within the confidence interval (see Cressie, 1993; Kitanidis, 1996).
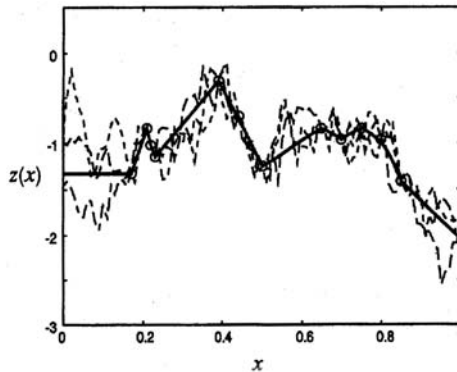
FIGURE 12.2 Three conditional realizations (discontinuous lines) and the best estimate (solid line).

12.4 Variogram

12.4.1 Variogram Model

To perform ordinary kriging, an expression for the variogram is needed. In the isotropic case, the variogram γ is a function of separation distance only, $h = ||x - x'||$, not the direction of the separation vector $x - x'$ between locations x and x'. For example, in three dimensions,

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \qquad x' = \begin{bmatrix} x_1' \\ x_2' \\ x_3' \end{bmatrix}, \qquad h = \sqrt{(x_1 - x_1')^2 + (x_2 - x_2')^2 + (x_3 - x_3')^2} \qquad (14)$$

We will consider the isotropic case first.

The variogram model must satisfy some mathematical requirements. In practice, the expression for the variogram is formed by combining a small number of simple mathematically acceptable expressions. Table 12.1 contains the most used isotropic variogram models.

TABLE 12.1 Common Variogram Models

Model	Expression, $\gamma(h)$	Parameters
Gaussian	$\sigma^2\left(1-\exp\left(-\dfrac{h^2}{L^2}\right)\right)$	$\sigma^2 > 0,\ L > 0$
Exponential	$\sigma^2\left(1-\exp\left(-\dfrac{h}{\ell}\right)\right)$	$\sigma^2 > 0,\ \ell > 0$
Spherical	$\begin{cases}\left(\dfrac{3}{2}\dfrac{h}{\alpha}-\dfrac{1}{2}\dfrac{h^3}{\alpha^3}\right)\sigma^2, & \text{for } 0 \le h \le \alpha \\[2mm] \sigma^2, & \text{for } h > \alpha\end{cases}$	$\sigma^2 > 0,\ \alpha > 0$
Nugget	$\begin{cases}C_0, & h > 0 \\[1mm] 0, & h = 0\end{cases}$	$C_0 > 0$
Linear	$\theta \cdot h$	$\theta > 0$
Power	$\theta \cdot h^s$	$\theta > 0,\ 2 > s > 0$

By adding variograms from Table 12.1, one can obtain other mathematically acceptable variograms, e.g., combining the linear and nugget effect variograms:

$$\gamma(h)=\begin{cases}C_0 + \theta h, & h > 0 \\[2mm] 0, & h = 0\end{cases} \tag{15}$$

with two parameters, $C_o \ge 0$ and $\theta \ge 0$. This way, one can construct a model that adequately and parsimoniously (i.e., without unnecessary terms and parameters) represents information about the structure of z.

In many applications, due to geologic stratification, hydrodynamic mechanisms, etc., the spatial structure may depend on direction as well as on distance. The simplest approach (Cressie, 1993; Kitanidis, 1997, describe more general cases) is to align the axes with the principal axes of anisotropy (e.g., parallel and perpendicular to layering) and to use as h:

$$h = \sqrt{\left(\frac{x_1 - x_1'}{l_1}\right)^2 + \left(\frac{x_2 - x_2'}{l_2}\right)^2 + \left(\frac{x_3 - x_3'}{l_3}\right)^2}$$

where l_1, l_2, l_3 are length parameters.

12.4.2 Variogram Selection

The variogram represents information about the structure of the function z and consequently should be consistent with: (1) prior information and (2) the observations.

Prior information, sometimes referred to as "soft data," is knowledge about the unknown that is additional to the observations. Prior information includes the experience and judgment of the analyst. In any specific application, there is always some prior information to the extent that even without observations one should be able to judge whether a variogram makes sense. Prior information is also important because certain features of the variogram cannot be established from the observations, no matter how dense. For example, if prior information suggests that function z should be smooth, then the Gaussian variogram (see Table 12.1) may be appropriate. If one expects that due to function "non-stationarity" (as when the likelihood of change in the hydrogeologic environment or geochemical regime increases rapidly with the distance) the estimate should not depend on measurements far away, one should choose a linear variogram. Certain decisions are easy to justify in the context of specific applications but, because here we deal with general methodologies, we will focus on the estimation of the variogram from data.

The process of developing a variogram from the data involves a number of steps and iterations. A variogram is selected tentatively based on an examination of the *experimental variogram*, as well as other information. The variogram is then calibrated (if it includes adjustable parameters) and tested on the basis of observations (or "hard data" or just "data") using *residuals*.

12.4.3 Experimental Variogram

For all $n(n-1)/2$ measurement pairs, plot the square difference $\frac{1}{2}[z(x_i) - z(x_i')]^2$ against the separation distance $h = \left\| x_i - x_i' \right\|$. The plotting produces a cloud of points known as the *raw variogram*; the *experimental variogram* is a relatively regular line through this scatter plot (see Figure 12.3).
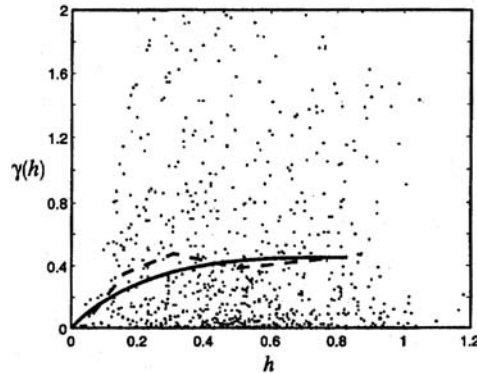
FIGURE 12.3 Raw (dots), experimental (dashed line), and model (solid line) variogram.

In the prevalent approach, the axis of separation distance is divided into consecutive intervals: the kth interval is denoted $[h_k^\ell, h_k^u]$ and contains N_k pairs of measurements $[z(x_i), z(x_i')]$. Compute:

$$\hat{\gamma}(h_k) = \frac{1}{2N_k} \sum_{i=1}^{N_k} \left[z(x_i) - z(x_i') \right]^2 \tag{16}$$

where index i refers to a pair of measurements $z(x_i)$ and $z(x_i')$ which meet the requirement

$$h_k^\ell < \left\| x_i - x_i' \right\| \leq h_k^u \tag{17}$$

The kth interval is represented by a single point h_k. Take h_k equal to the average value,

$$h_k = \frac{1}{N_k} \sum_{i=1}^{N_k} \left\| x_i - x_i' \right\| \tag{18}$$

Next, these points $[h_k, \hat{\gamma}(h_k)]$ are connected to form the experimental variogram (see Figure 12.3). Numerous modifications to this basic approach have been proposed to improve its robustness or to account for other effects (Cressie, 1993).

The appearance of the experimental variogram is affected by the choice of intervals. By choosing longer intervals, one averages over more points in each interval thus reducing jaggedness in the experimental variogram; however, one may sacrifice resolution, i.e., "straighten" the variogram. One must use judgment in selecting the intervals to achieve reasonable trade-offs between the objectives of reducing jaggedness and preserving resolution. It usually suffices to use three to six intervals, and it makes sense to use shorter intervals where the points of the raw variogram are less spread out (in the vertical direction).

An experimental variogram that depends on orientation and is thus useful in identifying anisotropy is the so-called directional variogram (Journel and Huigbregts, 1978).

The experimental variogram, though it may provide a reasonable preliminary estimate of the variogram as shown in Figure 12.3, has its limitations (Armstrong, 1984) and should not monopolize the analysis. The selection of the variogram to be used in kriging should rely on the examination of the residuals.

12.4.4 Residuals

Arrange the n measurements in a given sequence. Calculate the kriging estimate of z at the second point, x_2, given only the first measurement, x_1: $\hat{z}_2 = z(x_1)$ and $\sigma_2^2 = 2\gamma(x_1 - x_2)$. Calculate the actual error,

$$\delta_2 = z(x_2) - \hat{z}_2 \tag{19}$$

and normalize by the standard error

$$\varepsilon_2 = \frac{\delta_2}{\sigma_2} \tag{20}$$

Use the same procedure to construct the other residuals. For the kth measurement location, estimate through kriging the value of z using only the first $k-1$ measurements and normalize by the standard error. Thus,

$$\delta_k = z(x_k) - \hat{z}_k, \quad \text{for} \quad k = 2, \ldots, n \tag{21}$$

$$\varepsilon_k = \frac{\delta_k}{\sigma_k}, \quad \text{for} \quad k = 2, \ldots, n \tag{22}$$

Using the actual data, the variogram model, and kriging, one thus computes the actual (or *experimental*) residuals, which are $n-1$ δ or ε values.

According to the probabilistic model, the ε residuals satisfy:

$$E[\varepsilon_k] = 0, \quad k = 2, \ldots, n \tag{23}$$

$$E[\varepsilon_k \varepsilon_l] = \begin{cases} 1, & \text{if } k = \ell \\ 0, & \text{if } k \neq \ell \end{cases}, \quad k, \ell = 2, \ldots, n \tag{24}$$

and are called *orthonormal* (Kitanidis, 1991, 1997).

12.4.5 Variogram Testing

Model validation means testing the agreement of the model with the data. In simple terms, a validation test is like an experiment with the following steps:

1. Predict the outcome to the experiment using the theory
2. Observe the actual outcome of the experiment
3. Compare the predicted and observed outcomes.

The following tests are suggested to evaluate the model.

12.4.5.1 Q_1 Statistic

Experiment: Compute the average of the orthonormal residuals:

$$Q_1 = \frac{1}{n-1} \sum_{k=2}^{n} \varepsilon_k \tag{25}$$

Model Prediction: The model predicts that Q_1 is normally distributed with mean 0 and variance $\dfrac{1}{n-1}$.
Experimental Observation: Measure the experimental value of Q_1 (from the actual data).
Compare Model with Experiment: Agreement between the model and the data means that the experimental value of Q_1 is close to zero.

A reasonable rule is to reject the model if:

$$|Q_1| > \frac{2}{\sqrt{n-1}} \tag{26}$$

This rule involves a 5% probability that the correct model may be rejected.

12.4.5.2 Q_2 Statistic

The same procedure is followed with

$$Q_2 = \frac{1}{n-1} \sum_{k=2}^{n} \varepsilon_k^2 \tag{27}$$

Agreement between model and data means that the experimental value of Q_2 is near 1.
For $n > 40$, one may use (see Kitanidis, 1997, for procedure for smaller n): reject the model if

$$|Q_2 - 1| > 2.8 \sqrt{n-1} \tag{28}$$

12.4.5.3 Normality

The ε residuals should follow an approximately Gaussian distribution, because BLUE method makes most sense when estimation errors are approximately normal. Plotting the experimental $\varepsilon_2, \ldots, \varepsilon_n$ on normal probability paper one may visually detect departures from normality, or one may perform goodness-of-fit tests (Shapiro and Wilk, 1965, for example).

12.4.5.4 No Correlation

Finally, the experimental ε residuals should be uncorrelated, i.e., they should have the pure nugget effect variogram,

$$\gamma\left(x_i - x_j\right) = \begin{cases} 1, & \text{for } x_i \neq x_j \\ 0, & \text{for } x_i = x_j \end{cases} \tag{29}$$

One may thus perform a variogram analysis and test the hypothesis that the experimental variogram of $\varepsilon_2, \ldots, \varepsilon_n$ is indeed this γ. One can also plot the signs of the residuals seeking regular patterns that would contradict the model, which predicts that the residuals are uncorrelated.

12.4.6 Variogram Calibration

A measure of how close the model reproduces or fits the data is given by:

$$cR = Q_2 \exp\left(\frac{1}{n-1} \sum_{i=2}^{n} \ell n\left(\sigma_i^2\right)\right) \tag{30}$$

where σ_i^2 was defined in Section 12.4.4.

This quantity and Q_2 can be used to calibrate a variogram, i.e., to estimate its parameters (Kitanidis, 1991). Consider, for example, that the model of Equation (15) has been selected but its parameters C_0 and θ need to be estimated from the data. Select the values $C_0 \geq 0$ and $\theta \geq 0$ that minimize

$$cR = \exp\left(\frac{1}{n-1} \sum_{i=2}^{n} \ell n\left(\sigma_i^2\right)\right) \tag{31}$$

subject to the constraint

$$\frac{1}{n-1} \sum_{i=2}^{n} \varepsilon_i^2 = 1 \tag{32}$$

This method is related to more general methods of statistics (Corbeil and Searle, 1976; Rao, 1973; Kitanidis, 1983). Also, such a parameter estimation method is sometimes referred to as a "cross-validation" method.

12.4.7 Nonnegativity

Quite often z, the function of interest, is known to be nonnegative. For example, z may represent transmissivity or solute concentration. However, kriging does not account for this nonnegativity requirement. If the 95% confidence interval turns out unrealistic because it includes negative values, then the nonnegativity constraint needs to be enforced.

The simplest approach is to make a one-to-one variable transformation. The prevalent transformation is the logarithmic:

$$y\left(x\right) = \ln\left(z\left(x\right)\right) \tag{33}$$

Then, we proceed to perform kriging on the transformed variable. After the result is obtained, we back-transform the best estimate and the 95% confidence intervals or the conditional realizations. The results are guaranteed to be nonnegative. Issues of interpretation are discussed in Kitanidis (1997).

12.5 Kriging Variants

Consider estimation when z conforms to the general model of Equation (1). This formulation allows us to account for structural information by appropriately selecting the deterministic part in Equation (1), as the following two examples illustrate:

Example 1: If $z(\boldsymbol{x})$ is the (depth averaged) piezometric head in an aquifer where a linear trend surface is known to exist,

$$f_1 = 1, \qquad f_2 = \boldsymbol{x}_1, \qquad f_3 = \boldsymbol{x}_2 \tag{34}$$

where $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ are the two Cartesian coordinates of location $\boldsymbol{x}$.

Example 2: If z represents log conductivity and the domain includes two zones with different hydro-geologic properties, interzonal differences in the mean are accounted for through

$$f_1 = \begin{cases} 1, & \text{in zone 1} \\ 0, & \text{in zone 2} \end{cases}, \qquad f_2 = \begin{cases} 0, & \text{in zone 1} \\ 1, & \text{in zone 2} \end{cases} \tag{35}$$

The covariance function $R(\boldsymbol{x} - \boldsymbol{x}')$ of the random part is defined as

$$R(\boldsymbol{x} - \boldsymbol{x}') = E\big[\varepsilon(\boldsymbol{x})\varepsilon(\boldsymbol{x}')\big] \tag{36}$$

The estimate of z at location $\boldsymbol{x}_0$ is

$$\hat{z}_0 = c(\boldsymbol{x}_0) + \sum_{i=1}^{n} \lambda_i \big(z(\boldsymbol{x}_i) - c(\boldsymbol{x}_i)\big) \tag{37}$$

where the coefficients are found (using the previously mentioned unbiasedness and best requirements) from the kriging system of $n + p$ linear equations with $n + p$ unknowns:

$$\sum_{j=1}^{n} R(\boldsymbol{x}_i - \boldsymbol{x}_j)\lambda_j + \sum_{k=1}^{p} f_k(\boldsymbol{x}_i)v_k = R(\boldsymbol{x}_i - \boldsymbol{x}_0), \qquad i = 1, \ldots, n$$

$$\sum_{j=1}^{n} f_k(\boldsymbol{x}_j)\lambda_j = f_k(\boldsymbol{x}_0), \qquad k = 1, \ldots, p \tag{38}$$

where $v_1, \ldots, v_p$ are Lagrange multipliers. The mean square error is

$$E\left[\big(\hat{z}_0 - z(\boldsymbol{x}_0)\big)^2\right] = -\sum_{k=1}^{p} f_k(\boldsymbol{x}_o)v_k - \sum_{i=1}^{n} \lambda_i R(\boldsymbol{x}_i - \boldsymbol{x}_0) + R(0) \tag{39}$$

From the general case, the following special cases are obtained:

Ordinary Kriging: When $c(\boldsymbol{x}) = 0$, $p = 1$, $f_1 = 1$. The ordinary kriging equations are obtained if we set $-\gamma$ in place of R in Equations (38) and (39).

Kriging with External Drift and Simple Kriging: When $c(\boldsymbol{x})$ is given and there are no terms with unknown β coefficients. The kriging system is n equations with n unknowns and the ν coefficients vanish. The even more special case of $c(\boldsymbol{x})$ = constant, is known as simple kriging.

Universal Kriging: When $c(\boldsymbol{x})$ vanishes and there are p unknown drift coefficients. This is also equivalent to kriging with generalized covariances because one needs a simplified version of the actual covariance function (in ordinary kriging, one can use $-\gamma$ instead of the covariance function.) Issues of covariance estimation are discussed in Kitanidis (1993), and also the method of Section 12.7.3 can be applied.

12.6 Cokriging

Consider the case of N correlated functions, $z_1(\boldsymbol{x})$, ..., $z_N(\boldsymbol{x})$, such as log-transmissivity, head, and accretion rate. One may want to estimate one of these functions, say $z_1(\boldsymbol{x})$, from observations of $z_1(\boldsymbol{x})$, ..., $z_N(\boldsymbol{x})$ functions. The BLUE approach, known as *cokriging*, is similar to the one in kriging: after describing the join structure of the N functions, the best estimate is a linear function of the data with coefficients determined from the requirements of unbiasedness and minimum variance.

The join structure is described by:

Mean functions

$$E\big[z_k(\boldsymbol{x})\big] = m_k(\boldsymbol{x}), \qquad k = 1, \ldots, N \tag{40}$$

Each of the mean functions may be represented as the deterministic part in Equation (1).

Covariance (or autocovariance) functions:

$$E\Big[\big(z_k(\boldsymbol{x}) - m_k(\boldsymbol{x})\big)\big(z_k(\boldsymbol{x'}) - m_k(\boldsymbol{x'})\big)\Big] = R_{kk}(\boldsymbol{x}, \boldsymbol{x'}), \qquad k = 1, \ldots, N \tag{41}$$

Cross-covariance functions:

$$E\Big[\big(z_k(\boldsymbol{x}) - m_k(\boldsymbol{x})\big)\big(z_\ell(\boldsymbol{x'}) - m_\ell(\boldsymbol{x'})\big)\Big] = R_{k\ell}(\boldsymbol{x}, \boldsymbol{x'}),$$
$$\text{for} \quad k, \ell = 1, \ldots, N \text{ and } k \neq \ell \tag{42}$$

Mean and autocovariance functions describe the structure of individual functions. Cross-covariance functions describe the interdependence of pairs of functions.

The cokriging equations are incorporated into the general equations of the next section. Determination of the model (structural analysis) for a set of functions is a more advanced topic that cannot be covered here in detail. Nevertheless, the next section outlines a method for estimation of parameters of autocovariance and cross-covariance functions and examples of application of cokriging are given in Dagan (1985) and Hoeksema et al. (1989).

12.7 Generalized Linear Estimation Equations

All linear kriging and cokriging equations can be written in a unified way, using matrix notation, as follows.

12.7.1 Model

$$y = c + X\beta + \varepsilon \tag{43}$$

where y is the $n \times 1$ measurement vector, c is known $n \times 1$ vector, X is a known $n \times p$ matrix, β is a $p \times 1$ vector of parameters ("drift coefficients" in kriging), and ε is a random vector with zero mean and covariance matrix $Q_{yy}(\theta)$ (a function of some covariance parameters). Also,

$$s = c_s + X_s\beta + \varepsilon_s \tag{44}$$

where s is the $m \times 1$ vector of unknowns, c_s is known $m \times 1$ vector, X_s is a known $m \times p$ matrix, and ε_s is a random vector with zero mean, covariance matrix $Q_{ss}(\theta)$, and cross-covariance to ε that is $Q_{sy}(\theta)$.

12.7.2 BLUE

$$\hat{s} = c_s + \Lambda(y - c) \tag{45}$$

where:

$$\begin{bmatrix} Q_{yy} & X \\ X^T & 0 \end{bmatrix} \begin{bmatrix} \Lambda^T \\ M \end{bmatrix} = \begin{bmatrix} Q_{ys} \\ X_s^T \end{bmatrix} \tag{46}$$

The estimation error covariance matrix is:

$$E\left[(\hat{s} - s)(\hat{s} - s)^T \right] = -X_s M + Q_{ss} - Q_{sy}\Lambda^T \tag{47}$$

12.7.3 Parameter Estimation

Select covariance parameters θ that minimize:

$$\left(\left| Q_{yy} \right| \left| X^T Q_{yy}^{-1} X \right| \left| X^T X \right|^{-1} \right)^{\frac{1}{n-p}} \tag{48}$$

while at the same time satisfying:

$$Q_2 = \frac{1}{n-p} y^T \left(Q_{yy}^{-1} - Q_{yy}^{-1} X \left(X^T Q_{yy}^{-1} X \right)^{-1} X^T Q_{yy}^{-1} \right) y = 1 \tag{49}$$

where $|\ |$ applied on a square matrix indicates matrix determinant. Note that this method, which is a generalization of the method of minimizing cR in ordinary kriging, is appropriate for finding the parameters of generalized covariance functions and variograms (which are needed in kriging or cokriging) and not necessarily of ordinary covariance functions (see Kitanidis, 1987).

12.8 Inverse Problems

12.8.1 Definition

In order to be successful in modeling groundwater flow and transport, one needs to obtain representative values of the required parameters, such as transmissivity, storage coefficient, conductivity, dispersivity, retardation factor, reaction rate, etc. The task is particularly difficult when a parameter is known to vary in space but is not measured directly. For example, conductivity (or transmissivity) varies over the flow domain, but there are not enough direct observations to resolve its spatial variability.

The inverse problem of groundwater modeling is defined broadly as the estimation of parameters from observations of the system response. The quintessential problem is the determination of conductivity from observations of the hydraulic head. Usually, this goal is achieved through systematic adjustment of parameter values in order to match the observations (as illustrated in Carrera and Neuman, 1986; Cooley et al., 1986], hence the term *history matching* is also used. Many approaches have been proposed, and the literature on the subject is vast. There have been several literature reviews (Yeh, 1986; Carrera, 1987; Ginn and Cushman, 1990) and a book devoted to the subject (Sun, 1994).

In this chapter, the inverse problem is defined as:

Estimation of at least one spatial function $z(x)$ from observations of other quantities related to z through a mathematical model as well as direct measurements of z.

The method to be presented is an extension of the geostatistical methods for the solution of the interpolation problem. We will examine exclusively static estimation problems (where all observations are given at the same time, as opposed to methods where data are taken into account as they become available). The justification for the approach that follows is given in Kitanidis (1995).

12.8.2 Methodology

Consider that we want to estimate parameter $z(x)$, a spatial function such as the log conductivity. The function is represented though the model of Equation (1). After discretization, $z(x)$ is represented through an n by 1 vector s. The mean of s is

$$E[s] = X\beta \tag{50}$$

where X is a known $n \times p$ matrix, β are p unknown drift coefficients. Furthermore, s has a covariance matrix

$$E\left[(s - X\beta)(s - X\beta)^T \right] = Q(\theta) \tag{51}$$

that is considered a known function of parameters θ. If there are two or more spatially variable parameters that need to be estimated (such as log transmissivity and storage coefficient), this case is included in this formulation by considering s as the aggregation of all spatially variable parameters. The objective is to find the best estimate of s and to evaluate the error of estimation.

There may also be other unknown parameters, such as boundary conditions or the intensity of sources or sinks, but these will not be included here because we focus on the estimation of the spatial process and its structural parameters. The β and θ parameters are treated as unknown constants and are supposed to be few in number, certainly far fewer than the observations. The observations are related to the unknown spatial process and this relation is denoted through:

$$y = h(s) + v \tag{52}$$

where y is the $m \times 1$ vector of observations. The observation error v is random with normal distribution, zero mean, and covariance matrix R that is fixed or a known function of an expanded set of parameters θ. The standard deviations of the measurement errors, which are the square root of the diagonal elements of R, define how closely the observations should be reproduced.

An intuitive approach is to minimize the weighted least squares criterion:

$$\left(z - h(s)\right)^T R^{-1}\left(z - h(s)\right) + \left(s - X\beta\right)^T Q^{-1}\left(s - X\beta\right) \tag{53}$$

where the first term represents the objective of reproducing the observations and the second term the objective of being consistent with the structure. The minimization should be with respect to s and β vectors. The covariance parameters are to be imposed by the analyst based on prior information or inferred from the data through a process of cross-validation.

This approach is systematized into the two phases of a geostatistical approach:

1. Structural analysis, where the form of the mean and the covariance function are selected and the structural parameters are estimated.
2. Conditioning on the data, where vector s is estimated.

The implementation of the second step is an iterative cokriging approach. We start with an estimate of $\tilde{s}$ and we improve at each iteration. Find the derivative of h about s at $\tilde{s}$:

$$H = \left.\frac{\partial h}{\partial s}\right|_{s=\tilde{s}} \tag{54}$$

Then, assuming that the actual s is close to $\tilde{s}$, approximate

$$y = h(\tilde{s}) + H(s - \tilde{s}) + v$$

Define

$$\Sigma = HQH^T + R \tag{55}$$

Solve the cokriging system of equations:

$$\begin{bmatrix} \Sigma & HX \\ (HX)^T & 0 \end{bmatrix}\begin{bmatrix} \Lambda^T \\ M \end{bmatrix} = \begin{bmatrix} HQ \\ X^T \end{bmatrix} \tag{56}$$

where Λ is an $m \times n$ matrix of coefficients and M is a $p \times n$ matrix of multipliers. Then, the cokriging estimate is:

$$\hat{s} = \Lambda\left(y - h(\tilde{s}) + H\tilde{s}\right) \tag{57}$$

If $\hat{s}$ is practically equal to $\tilde{s}$, the algorithm has converged and the covariance matrix of estimation is

$$V = -XM + Q - QH^T\Lambda^T \tag{58}$$

Otherwise $\tilde{s}$ is set equal to $\hat{s}$ and the procedure is repeated. Examples are found in Kitanidis (1995) and Yeh et al. (1996).

The cokriging equations assume that the covariance parameters, θ, are known and in some approaches are not estimated in any formal way. They can be estimated (Kitanidis, 1995) by minimizing the criterion:

$$L = \ln\left|\Sigma\right| + \ln\left|X^T H^T \Sigma^{-1} HX\right|$$

$$+ \left(y - h(\hat{s}) + H\hat{s}\right)^T \left(\Sigma^{-1} - \Sigma^{-1} HX \left(X^T H^T \Sigma^{-1} HX\right)^{-1} X^T H^T \Sigma^{-1}\right) \left(y - h(\hat{s}) + H\hat{s}\right)$$

(59)

This approach is essentially a restricted maximum likelihood or cross-validation methodology. A variation to this approach is to minimize the sum of the first two terms subject to the constraint that the third term is equal to 1, which is the method of Section 12.7.3. The two approaches yield similar or identical results.

After the covariance parameters are updated, one may need to repeat the iterative procedure of estimating s.

The error computed in the iterative cokriging or "quasilinear" approach is valid provided that the final estimation error is small. In small-variance cases (Kitanidis and Vomvoris, 1983; Dagan, 1985; Hoeksema and Kitanidis, 1984, 1985, 1989; Rubin and Dagan, 1987a, 1987b; Wagner and Gorelick, 1989), the observation Equation (52) is linearized only once, about the mean (deterministic part). Sometimes, Monte Carlo methods can be used instead of the usual linearization (Hoeksema and Clapp, 1990). The approach is effectively the same with the method of Carrera and Neuman (1986) in the conditioning step but differs in the step of estimating θ. Conceptually and in simple terms, the relation of this approach to the least squares approach (Cooley and Naff, 1990, for example) is: in least squares the spatial variability is accounted for by including sufficient terms in the deterministic part; the covariance matrix Q is proportional to the identity matrix; and the problem is formulated to estimate the β coefficients.

12.8.3 Derivative Computation

Computationally, the greatest challenge in the application of this method with numerical models is the computation of the derivative matrix H in Equation (54). The direct approach of varying one element of s at a time is inefficient because usually there are many components in s and few components in y. More efficient "adjoint state" methods have been developed. This topic, although important, is beyond the scope of this chapter, and the reader is referred to Chavent et al. (1975), Neuman (1980), Townley and Wilson (1985), Carrera et al. (1990), and Sun (1994).

For Further Information

Cressie (1993) is a comprehensive reference on methods of statistical estimation of spatial functions. Journel and Huigbregts (1978) and Rendu (1981) describe geostatistics as applied in mining, including methods for the estimation of volume averages. Isaaks and Srivastava (1989) is another reference on basic geostatistics. Software is found in Deutsch and Journel (1992). Kitanidis (1997) is an introduction to estimation of spatial functions with applications in hydrogeology. De Marsily (1986) has a chapter on geostatistics including universal kriging and cokriging. A two-part paper (ASCE, 1990) reviews applications of geostatistics in hydrogeology. Statistical groundwater mechanics are covered in Dagan (1989) and Gelhar (1993). Rivoirard (1994) covers nonlinear geostatistics. Sun (1994) covers inverse methods in hydrogeology including methods for estimation and sensitivity analysis (computation of derivatives) in steady and unsteady state. Cooley and Naff (1990) describe the least squares approach to inverse modeling.

References

Armstrong, M. 1984. Common problems seen in variograms. *Mathematical Geology*. 16(3), 305-313.

ASCE Task Committee on Geohydrology. 1990. Review of geostatistics in geohydrology, Parts I and II. *ASCE J. of Hydraulic Engineering.* 116(5), 612-658.

Carrera, J. and Neuman, S. P. 1986. Estimation of aquifer parameters under transient and steady state conditions, 1. Maximum likelihood method incorporating prior information. *Water Resour. Res.* 22(2), 199-210.

Carrera, J. 1987. State of the art of the inverse problem applied to the flow and solute transport equations, in NATO ASI Ser., edited by E. Curtodio, A. Gurgui, and J. P. Lobo-Ferreira. D. Reidel, Hingham, MA.

Carrera, J., Navarrina, F., Vives, L., Heredia, J., and Medina, A. 1990. Computational aspects of the inverse problem, in *Computational Methods in Water Resources.* Computational Mechanics Publications, Billerica, MA, 513-523.

Chavent, G., Dupuy, M., and Lemonnier, P. 1975. History matching by use of optimal theory. *Soc. Pet. Eng. J.* 15(1), 74-86.

Cooley, R. L., Konikow, L. F., and Naff, R. L. 1986. Nonlinear-regression groundwater flow modeling of a deep regional aquifer system. *Water Resour. Res.* 22(13), 1759-1778.

Cooley, R. L. and Naff, R. L. 1990. Regression modeling of ground-water flow, USGS, Techniques of Water Resources Investigations, 03-B4. Washington, D.C.

Cressie, N. A. C. 1993. *Statistics for Spatial Data.* John Wiley & Sons, New York.

Corbeil, R. R. and Searle, S. R. 1976. Restricted maximum likelihood estimation of variance components in the mixed model. *Technometrics.* 18(1), 31-38.

Dagan, G. 1985. Stochastic modeling of groundwater flow by unconditional and conditional probabilities: The inverse problem. *Water Resour. Res.* 21(1), 65-72.

Dagan, G. 1989. *Flow and Transport in Porous Media.* Springer-Verlag, Berlin.

de Boor, C. 1978. *A Practical Guide to Splines.* Springer-Verlag, New York.

de Marsily, G. 1986. *Quantitative Hydrogeology.* Academic Press, New York.

Deutsch, C. V. and Journel, A. G. 1992. *GSLIB: Geostatistical Software Library and User's Guide.* Oxford University Press, New York.

Dykaar, B. B. and Kitanidis, P. K. 1992. Determination of the effective hydraulic conductivity for heterogeneous porous media using a numerical spectral approach 1. Method. *Water Resour. Res.* 28(4), 1155-1166.

Gelhar, L. W. 1993. *Stochastic Subsurface Hydrology.* Prentice Hall, Englewood Cliffs, NJ.

Ginn, T. R. and Cushman, J. H. 1990. Inverse methods for subsurface flow: A critical review of stochastic techniques. *Stoch. Hydrol. Hydraul.* 4, 1-26.

Gutjahr, A., Bullard, B., Hatch, S., and Hughson, L. 1994. Joint conditional simulations and the spectral approach for flow modeling. *Stoch. Hydrol. Hydraul.* 8 (1), 79-108.

Hoeksema, R. J. and Kitanidis, P. K. 1984. An application of the geostatistical approach to the inverse problem in two-dimensional groundwater modeling. *Water Resour. Res.* 20(7), 1003-1020.

Hoeksema, R. J. and Kitanidis, P. K. 1985. Comparison of Gaussian conditional mean and kriging estimation in the geostatistical solution of the inverse problem. *Water Resour. Res.* 21(6), 825-836.

Hoeksema, R. J. and Kitanidis, P. K. 1989. Prediction of transmissivities, heads, and seepage velocities using mathematical models and geostatistics. *Adv. in Water Resour.* 12(2), 90-102.

Hoeksema, R. J., Clapp, R. B., Thomas, A. L., Hunley, A. E., Farrow, N. D., and Dearstone, K. C. 1989. Cokriging model for estimation of water table elevation. *Water Resour. Res.* 25(3), 429-438.

Hoeksema, R. J. and Clapp, R. B. 1990. Calibration of groundwater flow models using Monte Carlo simulations and geostatistics, in *ModelCARE 90: Calibration and Reliability in Groundwater Modelling.* IAHS Publ. No 195, 33-42.

Isaaks, E. H. and Srivastava, R. M. 1989. *Applied Geostatistics.* Oxford University Press, New York.

Journel, A. G. and Huigbregts, Ch. J. 1978. *Mining Geostatistics.* Academic Press, New York.

Kitanidis, P. K. 1983. Statistical estimation of polynomial generalized covariance functions and hydrologic applications. *Water Resour. Res.* 19(4), 909-921.

Kitanidis, P. K. and Vomvoris, E. G. 1983. A geostatistical approach to the inverse problem in groundwater modeling (steady state) and one-dimensional simulations. *Water Resour. Res.* 19(3), 677-690.

Kitanidis, P. K. 1986. Parameter uncertainty in estimation of spatial functions: Bayesian analysis. *Water Resour. Res.* 22 (4), 499-507.

Kitanidis, P. K. 1991. Orthonormal residuals in geostatistics: model criticism and parameter estimation. *Mathematical Geology.* 23(5), 741-758.

Kitanidis, P. K. 1993. Generalized covariance functions in estimation. *Mathematical Geology.* 25(5), 525-540.

Kitanidis, P. K. 1995. Quasilinear geostatistical theory for inversing. *Water Resour. Res.* 31(10), 2411-2419.

Kitanidis, P. K. 1996. Analytical expressions of conditional mean, covariance, and sample functions in geostatistics. *Stoch. Hydrol. Hydraul.* 10(4), 279-294.

Kitanidis, P. K. 1997. *Introduction to Geostatistics: With Applications in Hydrogeology.* Cambridge Univ. Press, 249pp.

Mantoglou, A. and Wilson, J. L. 1982. The turning bands method for simulation of random fields using line generation by a spectral method. *Water Resour. Res.* 18(5), 1379-1394.

Matheron, G. 1971. *The Theory of Regionalized Variables and Its Applications.* Ecole de Mines, Fontainbleau, France.

Neuman, S. P. 1980. Adjoint-state finite element equations for parameter estimation, in *Finite Elements in Water Resources: Proceedings of Third International Conference.* Eds. S. W. Wang, C. V. Alonso, C. A. Brebbia, W. G. Gray, and G. F. Pinder. Springer-Verlag, New York.

Rao, C. R. 1973. *Linear Statistical Inference and Its Applications.* John Wiley & Sons, New York.

Rendu, J.-M. 1981. *An Introduction to Geostatistical Methods of Mineral Evaluation.* South African Institute of Mining and Metallurgy, Johannesburg.

Rivoirard, J. 1994. *Introduction to Disjunctive Kriging and Non-Linear Geostatistics.* Oxford University Press, New York.

Robin, M., Gutjahr, A., Sudicky, E., and Wilson, J. L. 1993. Cross-correlated random field generation with the direct Fourier transform method. *Water Resour. Res.* 29(7), 2385-2397.

Rubin, Y. and Dagan, G. 1987a. Stochastic identification of transmissivity and effective recharge in steady groundwater flow, 1. Theory. *Water Resour. Res.* 23(7), 1185-1192.

Rubin, Y. and Dagan, G. 1987b. Stochastic identification of transmissivity and effective recharge in steady groundwater flow, 2. Case study. *Water Resour. Res.* 23(7), 1193-1200.

Shapiro, S. S. and Wilk, M. B. 1965. An analysis of variance test for normality (Complete Samples). *Biometrika,* 52, 691-610.

Sun, N.-Z. 1994. *Inverse Problems in Groundwater Modeling.* Kluwer, Norwell, MA.

Thiebaux, H. J. and Pedder, M. A. 1987. *Spatial Objective Analysis: With Applications in Atmospheric Science.* Academic Press, London.

Tompson, A. F. B., Ababou, R., and Gelhar, L. W. 1989. Implementation of the three dimensional turning bands random field generator. *Water Resour. Res.* 25(10), 2227-2243.

Townley, L. R. and Wilson, J. L. 1985. Computationally efficient algorithms for parameter estimation and uncertainty propagation in numerical models of groundwater flow. *Water Resour. Res.* 21(12), 1851-1860.

Wagner, B. J. and Gorelick, S. M. 1989. Reliable aquifer remediation in the presence of spatially variable hydraulic conductivity: from data to design. *Water Resour. Res.* 25(10), 2211-2225.

Yeh, W. W.-G. 1986. Review of parameter identification procedures in groundwater hydrology: the inverse problem. *Water Resour. Res.* 22(1), 95-108.

Yeh, J. T.-C., Jin, M., and Hanna, S. 1996. An iterative stochastic inverse method: conditional effective transmissivity and hydraulic head fields. *Water Resour. Res.* 32(1), 85-92.

Glossary

BLUE Acronym for Best Linear Unbiased Estimation. Any estimation procedure where the estimate depends linearly on the data, with weights selected so that the estimation error has zero mean and minimum variance.

Cokriging A geostatistical estimation method where one function is estimated using measurements from another related function.

Conditional Realizations Functions that are possible solutions to the problem of finding a function from observations. They are consistent with the structure and with the data.

Deterministic Part Also known as mean, drift, or trend. The part of the function that is described through a deterministic expression, such as a polynomial or zones with uniform values.

Experimental Indicates computed from data.

Experimental Variogram An estimate of the variogram obtained from the data using a graphical procedure. It is usually shown as piecewise linear.

Generalized Covariance The part of the covariance that matters in universal kriging.

Geostatistics A method for the estimation of spatial variables. It is based on the theory of random functions.

Kriging Denotes any "geostatistical" interpolation or averaging method.

Ordinary kriging BLUE for the case that the mean is an unknown constant and the variogram is specified.

Orthonormal Normalized random variables, which are uncorrelated and have zero mean and unit variance.

Random Function Defined through probabilistic averages. Also known as *stochastic process*.

Random Part The part of the unknown function that is described as a random function with zero mean. Usually, the less "regular" part of the function.

Residuals Differences between model predictions and data. Useful in model testing and in parameter estimation.

Structure Common characteristics of an ensemble of functions. In this chapter, structure is quantified by expressions for the deterministic part and the covariance function of the random part.

Variogram Used synonymously with semivariogram. A measure of spatial structure. The expected value of half the square difference expressed as a function of the separation.

13
Groundwater Contaminants

Ernest R. Blatchley, III
and John E. Thompson
Purdue University

13.1 Introduction

Webster's New Collegiate Dictionary defines the verb "contaminate" as a process by which a material is made "inferior or impure by admixture" or made "unfit for use by the introduction of unwholesome or undesirable elements." While both of these definitions are accurate, they do not fully represent the connotation with which the word is used. For purposes of this chapter, the term "contaminant" will generally refer to an (undesirable) constituent which is introduced directly or indirectly as a result of human activity. Furthermore, the identification of groundwater contaminants and their effects will require knowledge of pre-existing or background conditions.

Groundwater contaminants exist in many forms, and contaminant classification schemes can be based on any of several physicochemical characteristics. For example, contaminants may be classified based on their preference for association with the aqueous phase or with particles. Contaminant distinction based on phase preference is important because the phase a contaminant associates with will affect its transport and toxicology. Furthermore, the form taken by a contaminant will also affect the choice of treatment processes which may be implemented to remediate a contaminated area. In general, the physicochemical characteristics of groundwater contaminants and the surrounding aquifer will play critical roles in determining their fate, transport, and effects.

The goal of this chapter is to define the nature and behavior of groundwater contaminants. Important physicochemical characteristics will be defined and described in terms of their potential effects on contaminants. The major contaminants classes will be examined and examples of important contaminants within each class will be provided.

13.2 Physicochemical Characteristics of Water

Water molecules are comprised of two hydrogen atoms and one oxygen atom. The oxygen atom, being more electronegative than hydrogen, preferentially attracts electrons, thereby leading to a nonuniform electron distribution around the water molecule. Together with the two pairs of unbound electrons carried by the oxygen atom, the water molecule displays an asymmetrical structure (see Figure 13.1). The hydrogen atoms form covalent bonds with the oxygen atom at an angle of 105°.
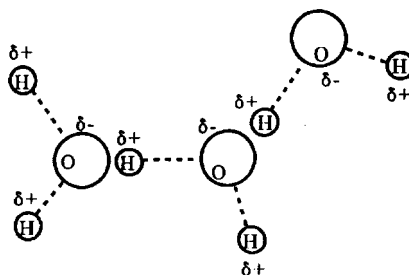
FIGURE 13.1 Schematic representation of the structure of the water molecule and the resultant hydrogen bonding. The terms $\delta+$ and $\delta-$ represent the positive and negative ends of the dipole.

The structure of water leads to several properties which are critical to its use as a resource and to the behavior of contaminants. Water is a highly polar molecule, with the negative end of the dipole being associated with the oxygen atom, and two strongly positive regions near the hydrogen atoms. One result of this structure is the ability of water molecules to attract each other through the formation of hydrogen bonds. These bonds, which are weak relative to covalent bonds, allow water to display properties which are somewhat anomalous based on its size and composition. For example, water has a higher boiling point than hydrogen sulfide (H_2S) even though sulfur, which lies immediately below oxygen on the periodic table, is heavier than oxygen. As a result, water exists as a liquid at room temperature and pressure, whereas H_2S is a gas under the same conditions. The polarity of water is critical to its ability to function as a solvent. The dipolar nature of water allows it to associate with positively or negatively charged species; in turn, this property allows water to separate (dissociate) ionizable constituents, leading to dissolution.

The ability of water to function as a solvent is not limited to ionizable constituents. Essentially all compounds have measurable solubility in water. However, compounds which do not form ions will express affinity for water molecules via another mechanism. An important example is the ability of water to dissolve polar molecules. If one examines the dipole moments of molecules in a homologous series of organic acids, decreasing polarity and aqueous solubility are observed as the molecule size increases. In essence, with the addition of each methyl group, the effects of the nonpolar group become less important because the electrons can distribute themselves over a larger molecule. Alternately, one could view the molecule as becoming more aliphatic as the chain length increases. Regardless of the viewpoint, the effect of polarity on water solubility is clear. In general terms, molecular solubility in water increases with polarity.

Natural or anthropogenic additions to water may also affect the solubility of constituents for water. For example, surfactants and detergents can dramatically increase the aqueous solubility of nonpolar compounds. In the case of surfactants, this can be accomplished through the formation of micelles, which exist as essentially a phase within a phase, often referred to as a "pseudo-phase" (see Figure 13.2). Micelles are agglomerations of organic molecules in which the polar ends of the surfactant molecules are directed outward, to associate with the bulk aqueous phase, and the relatively nonpolar ends of these molecules form the micelle core, yielding a relatively nonpolar region which promotes solubilization of nonpolar constituents. Detergents often serve as sequestering agents for Ca^{2+} and Mg^{2+} (the primary constituents of chemical hardness), thereby preventing interactions between hardness constituents and surfactants.

FIGURE 13.2 Two-dimensional schematic representation of a surfactant micelle in aqueous solution. The hydrophilic head is indicated by the filled circle (●); this end of the surfactant molecule will associate preferentially with the polar solvent (water). The nonpolar tails will be repelled by the water and allow the surfactant molecules to agglomerate into the micelles.
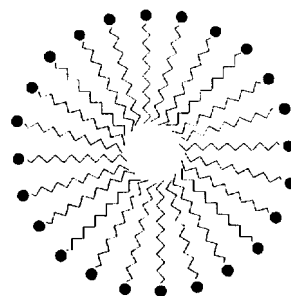
13.3 Factors Affecting Aqueous Solubility

The aqueous solubility of a contaminant compound will, to large degree, determine the transport, fate, and toxicology of that compound in a groundwater system. As such, it is important to define the characteristics of the system which will affect solubility. The primary factors which contribute to aqueous solubility are defined and reviewed below.

pH. Often referred to as a "master variable" in describing water composition, no single factor plays a more universal role in defining the characteristics of an aqueous system than pH. Water undergoes autoionization, resulting in the production of H^+ and OH^-

$$H_2O \rightleftarrows H^+ + OH^- \tag{1}$$

In reality, ions such as H^+ and OH^- do not exist in solution, at least not as written. As described previously, ions are surrounded by water molecules. Therefore, the notation H^+ is really a shorthand representation for the hydrated proton, which could be more accurately represented by $nH_2O \cdot H^+$, where the coefficient n represents the number of waters of hydration. For purposes of this chapter, shorthand notation will be used to represent H^+, OH^-, and all other ions. The strict definition of pH is as follows:

$$pH = -\log_{10}\left(H^+\right) \tag{2}$$

where (H^+) is defined as the activity of the hydrogen ion. For solutions of low ionic strength, activity may be closely approximated by molar concentration.

Because the activity of molecular water (solvent) in aqueous solution approaches unity as solute concentration approaches zero (i.e., dilute solutions), the equilibrium constant for reaction (1) can be defined as follows:

$$K_w = \left(H^+\right)\left(OH^-\right) \tag{3}$$

where K_w is defined as the ion product of water or the equilibrium "constant" for reaction (1). Like most acid-base reactions, equilibrium conditions in homogeneous solutions are established rapidly. However, it should be noted that the value of K_w is temperature dependent (see Table 13.1). Under some circumstances, the variations in K_w can have significant consequences in a system. For example, in a groundwater system, with a typical temperature of 13°C, the condition of neutrality (equal concentrations of H^+ and OH^-) will occur at a pH of approximately 7.2, whereas at room temperature, neutrality corresponds to pH = 7.0. In a physical sense, pH is a representation of the availability of hydrogen ions (protons) in solution; analytical determinations of pH are based on measurement of electrochemical potential using a hydrogen-ion selective electrode, which has been previously calibrated using standard solutions.

TABLE 13.1 Temperature Dependence of the Ion Product of Water (K_w)

Temperature (°C)	pK_w
0	14.9435
5	14.7338
10	14.5346
15	14.3463
20	14.1669
24	14.0000
25	13.9965
30	13.8330
35	13.6801
40	13.5348
45	13.3960
50	13.2617
55	13.1369
60	13.0171

Values as tabulated in the *CRC Handbook of Chemistry and Physics, 59th Edition* (1978). CRC Press, Boca Raton.

Solution pH has a dramatic effect on the characteristics of solubilized compounds which can donate or accept protons (acids and bases, respectively). Some compounds have the ability to act as both acids and bases. Many examples of these so-called "amphoteric" compounds can be found, none more important than water itself. Since K_w has been accurately quantified as a function of temperature, knowledge of pH and temperature yields knowledge of both (H^+) and (OH^-).

The effect of pH on aqueous chemistry is not limited to the autoionization of the water molecule; pH will determine the speciation of all acids and bases in solution. An important example in groundwater chemistry is the distribution of carbonate species. Gaseous carbon dioxide (CO_2) will dissolve in water to yield aqueous CO_2, followed by hydration to yield carbonic acid (H_2CO_3), which are often represented jointly by the fictitious species $H_2CO_3^*$. The equilibrium partitioning of gaseous and aqueous CO_2 are defined by Henry's law:

$$CO_{2(g)} \rightleftarrows CO_{2(aq)} \qquad H = \frac{\left(CO_{2(aq)}\right)}{\left(CO_{2(g)}\right)} \tag{4}$$
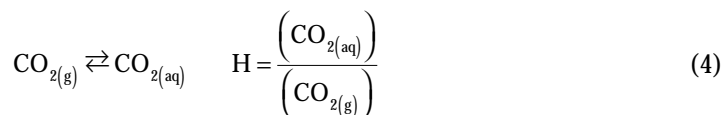
with gas phase activity (partial pressure) being expressed in the same units as aqueous phase activity. CO_2 hydration is defined by the following reaction:

$$CO_{2(g)} + H_2O \rightleftarrows H_2CO_3 \qquad K_h = \frac{\left(H_2CO_3\right)}{\left(CO_{2(aq)}\right)} \tag{5}$$
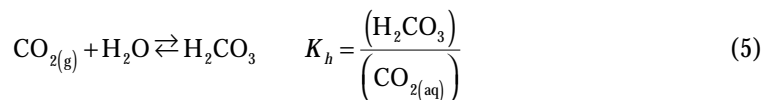
The equilibrium distribution of $CO_{2\,(aq)}$ and H_2CO_3 is such that the vast majority of the dissolved CO_2 attributed to these species is found in the form of $CO_{2\,(aq)}$. Therefore, the hypothetical compound $H_2CO_3^*$ is used to represent the combination of these two compounds. Carbonic acid can donate two protons and is therefore referred to as a diprotic acid:

$$H_2CO_3^* \rightleftarrows H^+ + HCO_3^- \qquad K_{a,1} = \frac{\left(H^+\right)\left(HCO_3^-\right)}{\left(H_2CO_3^*\right)} \tag{6}$$

$$HCO_3^- \rightleftarrows H^+ + CO_3^{2-} \qquad K_{a,1} = \frac{\left(H^+\right)\left(CO_3^{2-}\right)}{\left(HCO_3^-\right)} \tag{7}$$
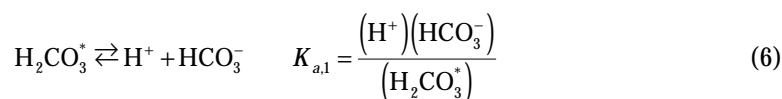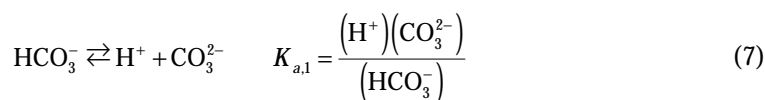
The equilibrium constants which describe these reactions are all functions of temperature. A compilation of these constants for the range of temperatures relevant to this subject can be found in Table 13.2. The information presented in Equations (4) through (7) and Table 13.2 reveal that with knowledge of pH and the concentration of any of the CO_2-based constituents, it is possible to quantify the concentrations (activities) of all CO_2-based constituents. As will be discussed later, the pH- dependence of the solubility of many metals in groundwater systems is closely related to the distribution of carbonate species.

TABLE 13.2 Temperature Dependence of Equilibrium Constants for the Carbonate System

Temp (°C)	$-\log_{10} H$	$pK_{a,1}$	$pK_{a,2}$
5	1.20	6.52	10.56
10	1.27	6.46	10.49
15	1.34	6.42	10.43
20	1.41	6.38	10.38
25	1.47	6.35	10.33
40	1.64	6.30	10.22
60	1.80	6.30	10.14

Data from Larson, T. E. and Buswell, A. M. 1942. Calcium carbonate saturation index and alkalinity interpretations, *Journal, AWWA*, 34, 1664.

Acid-base behavior determines the physicochemical characteristics of many other compounds. For example, H_2S is a relatively volatile, toxic compound. It is a product of some anaerobic biochemical processes, and can yield the sulfide ion (S^{2-}), which can play an important role in determining the solubility and mobility of metals. The link between H_2S and S^{2-} is described by the acid-base behavior of H_2S, another diprotic acid:

$$H_2S_{(aq)} \rightleftarrows H^+ + HS^- \qquad K_{a,1} = \frac{\left(H^+\right)\left(HS^-\right)}{\left(H_2S_{(aq)}\right)} \tag{8}$$

$$HS^- \rightleftarrows H^+ + S^{2-} \qquad K_{a,2} = \frac{\left(H^+\right)\left(S^{2-}\right)}{\left(HS^-\right)} \tag{9}$$
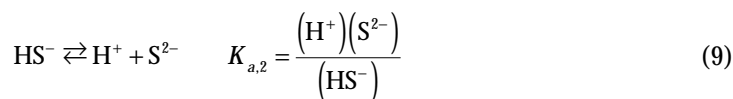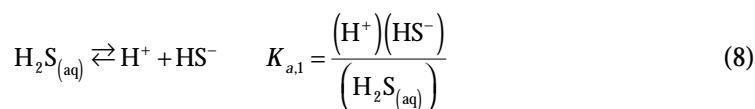
Dissociation of aqueous H_2S leads to the production of ionic species which have high solubilities relative to their parent compound. Furthermore, the sulfide ion (S^{2-}) forms precipitates with many metals which are extremely insoluble in aqueous solutions. The acid dissociation constants for the H_2S system (H_2S–HS^- pair and HS^-–S^{2-} pair) are listed together with those of many other important acid-base pairs in Table 13.3. These constants reveal the importance of pH in determining the characteristics and behavior of many compounds which display acid-base behavior in aqueous systems, such as groundwater.

For many of the same reasons that the term "pH" is often used to describe hydrogen-ion activity, the term "pC" is often used to describe the activity of a given constituent, where pC is defined as the negative $\log_{10}$ of the activity of that constituent. An important application for this term is in the development of pC-pH diagrams, which are used to illustrate acid-base behavior over a broad range of conditions. These diagrams are constructed based on knowledge of the equilibrium relationships for a given system and the total concentration of the compound(s) which display(s) acid/base behavior. An example of such a

TABLE 13.3 Acid Dissociation Constants for Common Ionizable Groundwater Constituents at 25°C

Compound (formula)	Conjugate Base (formula)	pK_a
Hydrochloric Acid (HCl)	Chloride (Cl^-)	~–3
Sulfuric Acid (H_2SO_4)	Bisulfate (HSO_4^-)	~–3
Nitric Acid (HNO_3)	Nitrate (NO_3^-)	–1
Bisulfate (HSO_4^-)	Sulfate (SO_4^{2-})	1.9
Phosphoric Acid (H_3PO_4)	Dihydrogen Phosphate ($H_2PO_4^-$)	2.1
Acetic Acid (CH_3COOH)	Acetate (CH_3COO^-)	4.7
Hydrogen Sulfide (H_2S)	Bisulfide (HS^-)	7.1
Dihydrogen Phosphate ($H_2PO_4^-$)	Hydrogen Phosphate (HPO_4^{2-})	7.2
Hydrogen Cyanide (HCN)	Cyanide (CN^-)	9.2
Ammonium (NH_4^+)	Ammonia (NH_3)	9.3
o-Silicic Acid ($Si(OH)_4$)	Silicate ($SiO(OH)_3^-$)	9.5
Silicate ($SiO(OH)_3^-$)	($SiO_2(OH)_2^{2-}$)	12.6
Bisulfide (HS^-)	Sulfide (S^{2-})	~17

Data from Stumm, W. and Morgan, J. J. 1996. *Aquatic Chemistry, 3rd Edition*, John Wiley & Sons, New York.

diagram is given for the carbonate system in Figure 13.3 for water exposed to the ambient atmosphere at 25°C, a so-called "open" system. Under the circumstances described in Figure 13.3, the constraint of equilibrium with atmospheric CO_2 dictates that the concentration of $H_2CO_3^*$ be fixed at approximately 10^{-5} molar, regardless of pH. At a pH between the first and second dissociation constants (i.e., $pK_{a,1} < pH < pK_{a,2}$), the majority of the carbonate present in the system will be found in the form of bicarbonate ion (HCO_3^-). At $pH > pK_{a,2}$ carbonate ion (CO_3^{2-}) will represent the predominant species. Under some circumstances, equilibration with atmospheric CO_2 cannot be assumed, such as an *in situ* groundwater. Under these circumstances (a "closed" system), the pC-pH diagram will be defined by the acid/base equilibria presented in Table 13.2 and the total amount of inorganic carbon (dissolved carbonate species) in solution. Figure 13.4 provides an example of the pH dependence of carbonate composition for a "closed" system, wherein exposure of the water to the ambient atmosphere is prevented. Since acid/base equilibria are established rapidly in most cases,* pC-pH diagrams represent a powerful tool for evaluating speciation of aqueous constituents which display acid/base behavior. The details of the development of these and other pC-pH diagrams can be found in texts which deal with the subject of aquatic chemistry (e.g., Stumm and Morgan, 1996).

Other aqueous constituents are also subject to pH effects. Principal among these are the metals, which generally exist in their oxidized forms in aqueous solutions containing measurable quantities of dissolved oxygen. The majority of these compounds exist as polyvalent cations and related complexes. Important examples among background constituents include calcium and magnesium, which are found in the forms of Ca^{2+} and Mg^{2+}, respectively, and related complexes. Ca^{2+} and Mg^{2+} will both form complexes and/or precipitates with hydroxide, bicarbonate, and carbonate. These anionic constituents are referred to as ligands in the complexation and precipitations reactions. Since the distribution of all three of these commonly found ligands are dependent on pH, the aqueous solubility of calcium and magnesium are also closely linked to pH. Similar logic can be used to describe the pH-dependent solubility behavior of many other metals in solution. A detailed discussion of this topic is beyond the scope of this text, but good summaries can be found in textbooks relating to the subject of aquatic chemistry. An excellent compilation of the equilibrium constants for complex formation and precipitation can be found in the reference series by Martell and Smith (1974–1989).

As in the case of acid-base reactions, the reactions which result in the formation of aqueous complexes in a homogeneous system (i.e., systems comprised only of an aqueous phase) are so rapid that in most

---

* Note that for acid/base distributions which rely on equilibration with a gas, the rate of equilibration between the gas and liquid phases will generally be much slower than the rate at which acid/base equilibria are reached in solution. Therefore, the assumption of the applicability of Henry's law (which defines gas-liquid partitioning at equilibrium) may not always be valid.
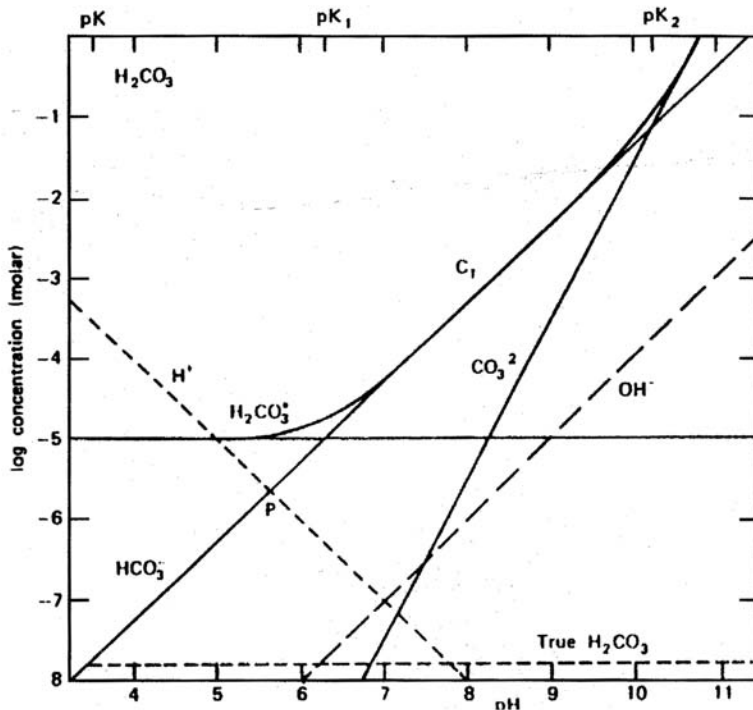
FIGURE 13.3 pC-pH diagram for the carbonate system under conditions of equilibrium with the atmosphere and 25°C. (From Stumm, W. and Morgan, J. J. 1981. *Aquatic Chemistry, 2nd Edition*, John Wiley & Sons, New York. With permission.)
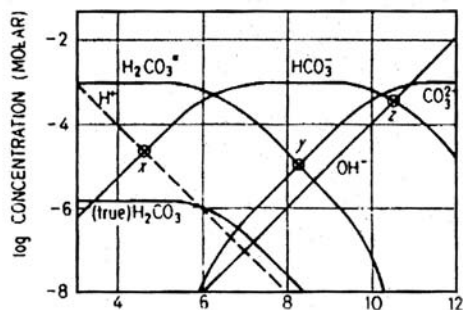
FIGURE 13.4 pC-pH diagram for the carbonate system without atmospheric exposure at 25°C. (From Stumm, W. and Morgan, J. J. 1981. *Aquatic Chemistry, 2nd Edition*, John Wiley & Sons, New York. With permission.)

circumstances equilibrium distributions can be assumed. However, the formation of precipitates and the reverse reaction (dissolution) involve at least two phases and equilibrium conditions do not always apply. Several factors contribute to the sometimes slow rates with which equilibria are achieved in these multiphase systems including transport limitations between phases and slow rates of crystallization among solids. Therefore, equilibrium calculations of solubility based on the presence of precipitates are not always valid and may result in errors when compared to actual conditions.

The pC-pH diagram can also be used as a tool for evaluation of precipitate formation. One example of how pC-pH diagrams may be used for this purpose is the so-called predominance area diagram, which provides an indication of the conditions under which a particular solid phase (precipitate) is expected to predominate. These diagrams can be used to predict the form of a solid and the limits of solubility
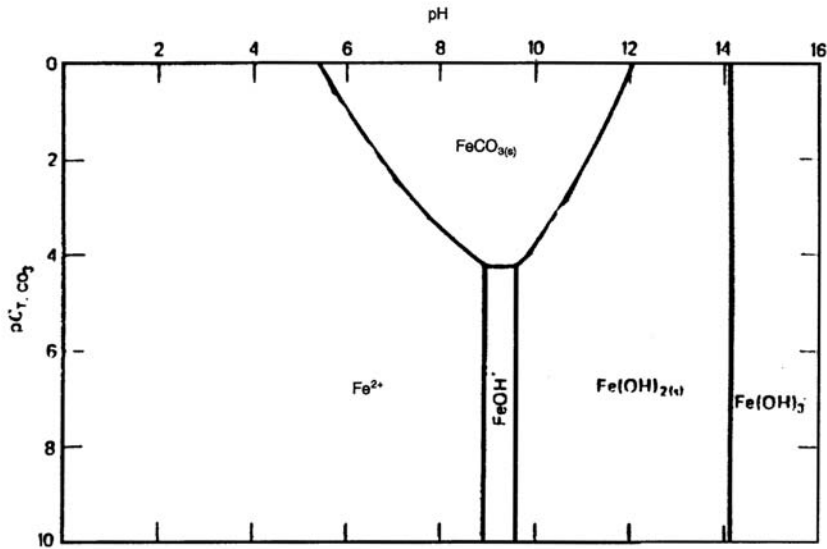
FIGURE 13.5 Predominance-area diagram for Fe(II) under conditions of exposure to carbonate and hydroxide and a total iron concentration of 10^{-5} M. (From Snoeyink, V. L. and Jenkins, D. 1980. *Water Chemistry*, John Wiley & Sons, New York. With permission.)

under aqueous conditions which will allow a metal to form precipitates with more than one ligand. An example of such a diagram is presented in Figure 13.5 for reduced iron, or Fe(II) at a total system concentration (C_t) of 10^{-5} M. Fe(II) will form precipitates in the presence of both hydroxide and carbonate. Since both ligands are likely to be present in groundwater systems, it is sometimes important to evaluate their combined effects on metal solubility. At pH < 10.5, the majority of precipitated Fe(II) in this example is expected to be present in the form of $FeCO_{3\ (s)}$. As pH increases above this value, more hydroxide ion becomes available and $Fe(OH)_{2\ (s)}$ is expected to predominate. Under either condition, it is likely that both forms of Fe(II) will be present, but one form will be expected to predominate. Figure 13.5 also defines the concentration of Fe(II) and related complexes which would be expected to be present in solution, in equilibrium with the solid phases, as a function of pH. The limit of solubility is defined by the lower borders of the solid regions within the predominance area diagram.

The characteristics and behavior of solids in solutions are strongly related to pH; in particular, the net surface charge displayed by a particle in solution is largely determined by its composition and the solution pH. At least three mechanisms exist whereby solids may express a surface charge in solutions. The first involves surface groups which display many of the same characteristics as dissolved constituents. For example, silica-based particles (e.g., sand) will have silanol groups; silanol groups display amphoteric behavior in aqueous solution:

$$\equiv Si-OH_2^+ \rightleftarrows \equiv Si-OH \rightleftarrows \equiv Si-O^- \tag{10}$$
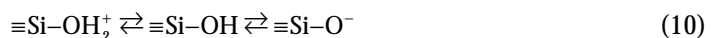
The charge expressed by this surface group will be determined by the availability of protons to adsorb to the site. Therefore, the charge expressed by this surface group will depend on the pH of the solution to which it is exposed. Many other surface groups on particles express amphoterism; furthermore, this behavior is not limited to inorganic surface groups.

A similar mechanism by which particles obtain surface charge is the adsorption of ionic constituents. The similarity stems from the fact that the behavior of amphoteric groups may be viewed as the equivalent of adsorption or desorption of protons. Similarly, other ionic constituents may adsorb to these sites resulting in a corresponding change of the surface charge.

A third mechanism by which particles express surface charge is isomorphous replacement of metals within crystal structures. For example, silica sand is commonly assumed to have a monomeric structure of SiO_2; however, imperfections in this structure, perhaps by the inclusion of metals with valence which is different than Si, results in a net surface charge.

The result of these processes is that particulates in aqueous solution often express a net surface charge. This charge is balanced by the migration of ions in solution to produce a region of local charge imbalance within the solution immediately adjacent to the particle:water interface. This region is often referred to as the electrical double-layer and plays an important role in governing the behavior of many constituents in groundwater systems.

The effect of pH in determining surface charge is profound. Titrimetric procedures exist whereby changes in surface charge may be quantified as a function of pH. These procedures allow for the determination of the zero point of charge (zpc), or the pH at which net surface charge is zero. A compilation of zpc values for some common particles is found in Table 13.4. At a pH above the zpc in a system where protons are the only adsorbable ion, the surface will express a net negative charge, whereas at a pH below the zpc, a net positive surface charge is expressed. From the values presented Table 13.4, it is evident that at near-neutral pH, most particles in solution express a net negative surface charge; however, important exceptions to this generalization can be found. This fact is important in that particles with like charges will repel each other, whereas those with opposite surface charge will experience a net attractive force. In fact, as a result of London and van der Waal's forces, even particles with like charges may experience a net attractive force if the surface charge is sufficiently small.

TABLE 13.4 Zero Points of Charge (pH_{zpc}) for Selected Inorganic Solids

Compound	pH_{zpc}
α-Al(OH)$_3$	5.0
CuO	9.5
Fe_3O_4	6.5
α-FeOOH	7.8
δ-MnO$_2$	2.8
SiO_2	2.0
Kaolinite	4.6
Montmorillonite	2.5

Data from Stumm, W. and Morgan, J. J. 1981. *Aquatic Chemistry, 2nd Edition*,
   John Wiley & Sons, New York.

Surface charges expressed by particles and soil media must be considered in an evaluation of the fate and transport of ionic and particulate contaminants. The fate of particles in soils has been the subject of considerable numerical and experimental research dealing with the process of sand filtration. The results of these research efforts have generated numerical models which can accurately represent the movement of particles in soil media under some conditions. Those interested in modeling particle transport in groundwater systems should refer to the text by Clark (1996), which provides an excellent critical review of the pertinent literature on the subject.

Sorption to Solids. In the general case, compounds present in groundwater systems will have the opportunity to associate with the aqueous phase and the solid phase(s). The partitioning of a constituent among these phases will depend on the characteristics of the constituent and the phases themselves.

Two categories of association among aqueous and solid phases exist. The first involves the accumulation of material at a water:solid interface and is termed adsorption. The second involves the intermingling of solute molecules with the molecules of the solid phase and is referred to as absorption. Absorption may be viewed as the dissolution of an aqueous constituent in a solid solvent. The distinction between the two processes is often vague and in a large number of cases, such a distinction is nearly impossible to make. Furthermore, in many cases both adsorption and absorption play a significant role in determining

the partitioning behavior of groundwater constituents among the solid and liquid phases. For purposes of this chapter, these terms will be collectively referred to as sorption.

Three general types of sorption can be described for groundwater systems. All three involve the formation of bonds between the constituent and the solid phase. The first type, physisorption, involves the formation of relatively weak physical bonds such as those attributable to London and van der Waal's forces. Under some conditions, the interactions between a solute and a solid surface may be more accurately described as a chemical bond, with the process being referred to as chemisorption. In chemisorption, the bond strength between a solute and the solid surface may approach that of a covalent bond. The third type of interaction involves an electrostatic interaction between surface groups and the solute and is often referred to as ion exchange.

Physisorption and ion exchange (which may be viewed as a special case of physisorption) tend to be readily reversible processes, since the forces which are responsible for the attraction of a solute to the solid are relatively weak. In the case of ion exchange, the factors which govern the affinity of a solute for the solid phase are comparatively well-defined. For ions of similar size, the valence of the ion will govern the magnitude of the coulombic force between the surface and solute; ions of higher valence will show a greater affinity for the solid surface. For ions of similar valence, the hydrated radius of the ion will dictate solid affinity, with ions having small hydrated radii showing preference over those with large radii. Interestingly, hydrated radius is inversely related to (unhydrated) ionic radius for elements of the same family on the periodic table. Therefore, ion exchange selectivity order may be defined for a particular solid medium. A representative example of a selectivity order is presented below for some common cations:

$$Ba^{2+} > Sr^{2+} > Ca^{2+} > Mg^{2+} > Cs^+ > K^+ > Na^+ > Li^+ \tag{11}$$
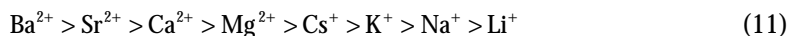
This order implies that at similar concentrations, ions will replace those ions which lie to their right in the selectivity order. As an example, Ca^{2+} will replace Na^+ under most conditions. The specific selectivity order displayed by a soil medium may vary from that defined in Equation (11) due to situation-specific circumstances. Furthermore, the reader should recognize that when solute concentrations are not similar, the tendency of ions with relatively high solution-phase concentration to replace ions associated with the solid phase will increase.

The importance of sorptive processes in determining the fate of groundwater contaminants may be illustrated by an example. Imagine a chemical spill comprised of three constituents: 1, 2, and 3, all at similar concentrations. Furthermore, we will assume that all three constituents compete for the same sorption sites within the soil matrix and that the affinity of the constituents is defined by $3 > 2 > 1$. All three compounds will be moved through the soil medium by advection, but their movement will be retarded by interactions with soil surfaces. Since constituent 3 shows the highest affinity for the soil matrix, it will be able to replace constituents 2 and 1. Therefore, on average, constituent 3 will spend more time in association with the soil matrix than constituents 2 or 1, and will therefore move through the matrix more slowly. A schematic representation of the movement of these constituents through this hypothetical system is provided in Figure 13.6. The separation of constituents 1, 2, and 3 based on their affinity for the stationary soil matrix is referred to as chromatography. This process is used on a much smaller scale and under more controlled conditions in the family of chromatographic analytical procedures, which are extremely important in the analysis of groundwater contaminants. Important analytical techniques which employ these principles include gas chromatography (GC), liquid chromatography (LC), and ion chromatography (IC). Many variations on these processes exist.

pε. Much in the same way that pH is used as a master variable in describing the availability of protons, *pε* is a master variable which is used to describe the availability of electrons for reaction. The analogy may also be carried further in that much as H^+ does not exist in solution, free electrons do not exist in solution. The mathematical definition of *pε* is as follows:
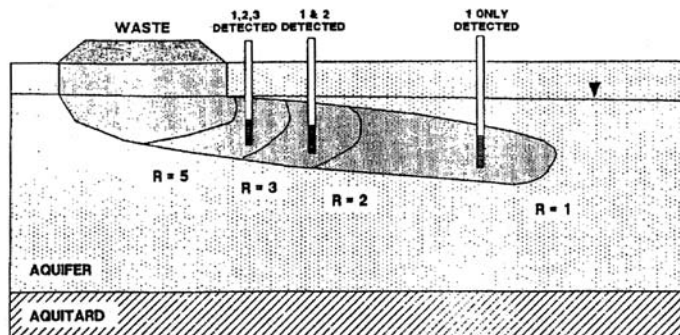
FIGURE 13.6 Schematic representation of the behavior of a mixture of pollutants (compounds 1, 2, and 3) in a groundwater system. All pollutants are assumed to be introduced from the same spill and are assumed to be present initially in roughly equal concentrations. The affinity of the pollutants for the soil matrix is defined by the order 3 > 2 > 1. (From U.S. Environmental Protection Agency. 1989. *Transport and Fate of Contaminants in the Subsurface*, EPA/625/4-89/019.)

$$p\epsilon = -\log_{10}\left(e^-\right) \tag{12}$$

In aqueous systems, $p\epsilon$ provides an indication of the tendency for electrons to be transferred. Under conditions of low $p\epsilon$, the activity of aqueous electrons is high, and there is significant pressure being exerted on aqueous constituents to undergo reduction. Conversely, under conditions of high $p\epsilon$, electron activity is low and the tendency is for oxidation reactions to be promoted. Oxidation and reduction reactions cannot take place in isolation (free electrons do not remain in solution for any significant period of time). Therefore, oxidation and reduction reactions must be coupled.

As in the case of pH, graphical tools exist to describe equilibrium conditions using $p\epsilon$ as a master variable. One such tool is the so-called $p\epsilon$-pC diagram. Figure 13.7 provides an example of such a diagram for a system containing dissolved iron at a total iron concentration of 10^{-4} M and a pH of 2. Under conditions of low $p\epsilon$, the system will be characterized as a reducing environment. In the case of the Fe(II)/Fe(III) system, this suggests that Fe(II) will be the dominant form. When $p\epsilon$ increases, a more oxidizing environment will persist. For the iron system at pH =2, at $p\epsilon > 13$, Fe(III) is expected to be the predominant form.

The $p\epsilon$-pC diagram is entirely analogous to the pC-pH diagram for description of acid/base equilibria. However, the $p\epsilon$-pC diagram is somewhat limited in its ability to present these equilibria in the sense that only a single pH can be represented. Another graphical tool for presentation of redox equilibria is the $p\epsilon$-pH diagram. Figure 13.8 provides an example of such a graph for the sulfur system at a total dissolved sulfur concentration of 10^{-2} M (Stumm and Morgan, 1981). This diagram provides an illustration of the conditions under which given sulfur species are expected to predominate. Though many similarities exist between pH and $p\epsilon$ and the equilibria they can be used to describe, a significant difference between equilibrium calculations based on acid/base transfer reactions and those corresponding to oxidation/reduction reactions is that whereas equilibrium conditions can often be assumed to apply for reactions involving proton transfer (i.e., acid/base reactions), the same cannot be said for reactions which involve electron transfer (i.e., oxidation/reduction reactions). Therefore, equilibrium calculations for redox reactions will not always yield an accurate representation of actual conditions in an aqueous system; however, these calculations do represent a powerful tool for defining the conditions one would expect to find if an infinite amount of time were allowed for the system to stabilize.

Compilations of aqueous redox chemical equilibria can be found in the literature. An excellent summary of the methods used to conduct redox equilibrium calculations and graphical representations of these equilibria can be found in Stumm and Morgan (1996).

Temperature. Thermodynamic considerations play an important role in evaluating the rates and equilibria of chemical reactions. In general, reaction rates will increase with temperature due to the
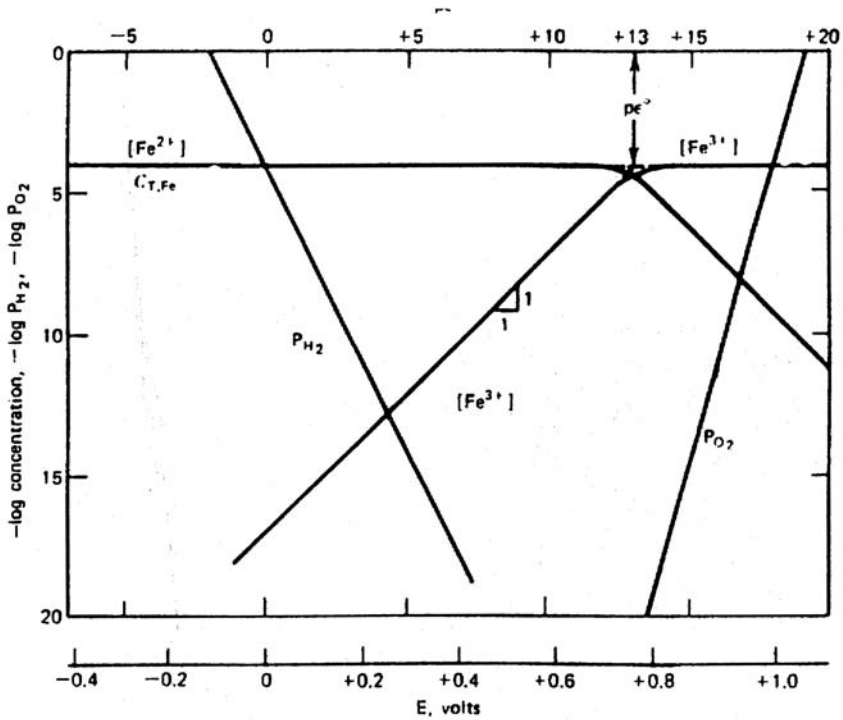
FIGURE 13.7 pε-pC diagram for iron at a total dissolved iron concentration of 10^{-4} M and pH = 2. (From Snoeyink, V. L. and Jenkins, D. 1980. *Water Chemistry*, John Wiley & Sons, New York. With permission.)
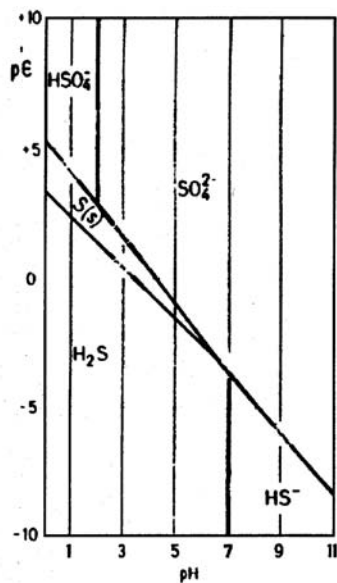
FIGURE 13.8 pε-pH diagram for sulfur at a total sulfur concentration of 10^{-2} M. (From Stumm, W. and Morgan, J. J. 1981. *Aquatic Chemistry, 2nd Edition*, John Wiley & Sons, New York. With permission.)

increased availability of thermal energy. This energy will promote collisions between molecules by increasing rates of diffusion and will reduce energy barriers which might otherwise limit reaction rates. Equilibrium conditions are achieved in a system when the rates of forward and reverse reactions are equal. If a given reaction is endothermic (i.e., heat is "consumed" by the reaction), then an increase in temperature would be expected to promote the forward reaction and shift the equilibrium to the right (an increase in "product" formation). Examples of the effects of heat on groundwater contaminants include changes in the equilibria and kinetics of precipitation and dissolution of minerals, and changes in the rates by which biological transformation processes proceed.

In most groundwater situations involving *in situ* contamination, temperature can often be assumed to be constant at a value of 1 to 2°C greater than the annual average ambient air temperature for the region (Freeze and Cherry, 1979). For example, the groundwater temperature in the midwestern United States (at depths greater than approximately 10 m, where diurnal and seasonal variations in air temperature have minimal effect) can generally be assumed to be approximately 13°C. Under these conditions, the role of temperature in governing contaminant transport and behavior is more of a buffer than a variable. However, those who deal with groundwater contaminants should be aware of the potential for thermal effects to influence system behavior. For example, shallow contamination in areas subject to substantial variations in climate (i.e., within the frost zone) may display variations in behavior as a function of the season. Similarly, other thermal sources and sinks within groundwater systems may play a role in determining the local behavior of contaminants. And finally, many remediation strategies call for removal of contaminants for *ex situ* treatment. Contaminants exposed to these conditions will obviously be subjected to a less stable thermal environment than those which remain *in situ*. Consequently, the behavior of groundwater contaminants may be different in the *ex situ* environment than *in situ*.

13.4 Inorganic Constituents

The compounds which comprise the inorganic fraction of groundwaters are largely attributable to interactions between source water and soils. As a source water moves through a soil, chemical reactions taking place at the soil:water interface will change the composition of both phases. In the previous section, the basic characteristics of these reactions were reviewed. Based on the working definition of "groundwater contaminants" provided in the introduction, the inorganic constituents we are interested in are those which have been introduced to a groundwater system as a result of human activity. Therefore, it is necessary to define background conditions in an unpolluted groundwater system to understand the significance of these reactions in terms of inorganic groundwater contaminants.

Source Water Composition. The accumulation of atmospheric moisture as a result of evapotranspiration may be viewed as a large-scale distillation process. However, the composition of precipitation, in the form of rain or snow, is not accurately represented by the composition of distilled water one might find in a laboratory. In particular, rain and snow will react with atmospheric constituents to determine its composition prior to reaching the earth's surface. A summary of measured compositions in rain and snow is provided in Table 13.5.

Many atmospheric gases display significant solubility in water. The typical composition of gases in a "clean" atmosphere (which also must be defined in relative terms as a result of human activities) is summarized in Table 13.6. Normally, precipitation can be assumed to be at equilibrium with atmospheric gases, as defined by Henry's law. Of the constituents listed in Table 13.6, several play significant roles in determining the chemistry of rain and snow. Primary among these is carbon dioxide (CO_2). As described previously, CO_2 will dissolve in water to form carbonic acid, which can dissociate to form bicarbonate (HCO_3^-) and carbonate (CO_3^{2-}) ions. The carbonate system plays a major role in governing groundwater composition through precipitation and complexation reactions. The carbonic acid cycle also plays a major role in determining the pH and buffering capacity of precipitation. Specifically, water at equilibrium with atmospheric CO_2 will have a pH of approximately 5.7. This value is often used as benchmark in

TABLE 13.5 Measurements of the Chemical Composition of Rain and Snow

Constituent	1	2	3	4	5
SiO_2	0.0		1.2	0.3	
Al (III)	0.01				
Ca^{2+}	0.0	0.65	1.2	0.8	3.3
Mg^{2+}	0.2	0.14	0.7	1.2	0.36
Na^+	0.6	0.56	0.0	9.4	0.97
K^+	0.6	0.11	0.0	0.0	0.23
NH_4^+	0.0				0.42
HCO_3	3		7	4	0.0
SO_4^2	1.6	2.18	0.7	7.6	6.1
Cl^-	0.2	0.57	0.8	17	2.0
NO_2^-	0.02		0.0	0.02	
NO_3^-	0.1	0.62	0.2	0.0	2.2
Total dissolved solids	4.8		8.2	38	
pH	5.6		6.4	5.5	4.4

1. Snow, Spooner Summit, U.S. Highway 50, Nevada (east of Lake Tahoe), altitude 7100 ft, Nov. 20, 1958. J. H. Feth, S. M. Rogers, and C. E. Roberson, *Chemical Composition of Snow in the Northern Sierra Nevada and Other Areas.* U.S. Geological Survey Water Supply Paper 1535J, 1964, 39 pp.

2. Average composition of rain from August 1962 to July 1963 at 27 points in North Carolina and Virginia. A. W. Gambell and D. W. Fisher, *Chemical Composition of Rainfall, Eastern N. Carolina and Southeastern Virginia:* U.S. Geological Survey Water Supply Paper 1535K, 1964, 41 pp.

3 and 4. Rain, Menlo Park, Calif., 7:00 P.M. Jan. 9 to 8:00 A.M. Jan. 10, 1958. Whitehead, and J. H. Feth, *Chemical Composition of Rain, Dry Fallout, and Bulk Precipitation at Menlo Park, Calif., 1957–1959, J. Geophys. Res.*, 69:3319-3333 (1964).

5. Station 526U, Belgium, European Atmospheric Chemistry Network. Average of 180 samples. L. Granat, On the Relation Between pH and the Chemical Composition in Atmospheric Precipitation, *Tellus*, 24, 550-556 (1972).

From Snoeyink, V. L. and Jenkins, D. 1980. *Water Chemistry*, John Wiley & Sons, New York. With permission.

the evaluation of the effects of atmospheric constituents which participate in the formation of acidic precipitation — precipitation with pH < 5.7 is often assumed to be under the influence of strong acids.

The influence of atmospheric composition on groundwater chemistry goes far beyond CO_2. Examples of other atmospheric constituents which play significant roles in the subsurface environment are provided below.

Oxygen (O_2) will dissolve in the aqueous phase to yield a liquid-phase concentration of approximately 8 to 10 mg/L, depending on altitude and temperature. In a groundwater system, the availability of oxygen (and other constituents) will play a major role in determining the extent and type (i.e., aerobic, anoxic, or anaerobic) of microbiological activity which will take place. Microbiologically mediated chemical transformations can play a significant role in determining groundwater composition; this fact is exploited in many soil and groundwater remediation strategies (see Chapters 25 and 26).

The oxides of nitrogen and sulfur (NO_x and SO_x), which are released to the atmosphere as a part of many natural processes (e.g., volcanic activity, microbiological activity) can react in the atmosphere to yield strong acids, such as HNO_3 and H_2SO_4. In an unpolluted atmosphere, the concentrations of these constituents are relatively low and do not play a significant role in determining precipitation chemistry. However, combustion of fossil fuels in mobile and stationary sources results in large increases in atmospheric NO_x and SO_x concentration. The scale on which these combustion processes are carried out allows them to influence rain and snow chemistry over geographic areas which span international boundaries.

TABLE 13.6 Typical Gas-Phase Composition in a "Clean" Atmosphere

Gas	Percentage by Volume	Partial Pressure (atm)
N_2	78.1	0.781
O_2	20.9	0.209
Ar	0.93	0.0093
H_2O	0.1–2.8	0.028
CO_2	0.03	0.0003
Ne	1.8×10^{-3}	1.8×10^{-5}
He	5.2×10^{-4}	5.2×10^{-6}
CH_4	1.5×10^{-4}	1.5×10^{-6}
Kr	1.1×10^{-4}	1.1×10^{-6}
CO	$(0.6–1) \times 10^{-4}$	$(0.6–1) \times 10^{-6}$
SO_2	1×10^{-4}	1×10^{-6}
N_2O	5×10^{-5}	5×10^{-7}
H_2	5×10^{-5}	5×10^{-7}
O_3	$(0.1–1.0) \times 10^{-5}$	$(0.1–1.0) \times 10^{-7}$
Xe	8.7×10^{-6}	8.7×10^{-8}
NO_2	$(0.05–2) \times 10^{-5}$	$(0.05–2) \times 10^{-8}$
Rn	6×10^{-18}	6×10^{-20}

Source: B. A. Mirtov, Gaseous composition of the atmosphere and its analysis. *Akad. Nauk. SSSR*, Inst. Prikl. Geofiz Moskva (translated by the Israel Program for Scientific Translations, published in Washington, U.S. Dept. of Commerce, Office of Technical Services, 1961, 209 pp.).

Radon (Rn) is a noble gas, which means that its electron structure is such that it tends to be quite inert with respect to conventional chemical reactions. However, radon has several naturally occurring isotopes, some of which have unstable nuclear structures, and therefore express radioactivity. In the natural atmosphere, the concentration of Rn is quite low and does not play a significant role in determining precipitation chemistry.

Groundwater composition will not be strictly governed by the composition of atmospheric gases; the subsurface may contain gases at partial pressures which are markedly different from those of the atmosphere. For example, locally intense microbiological activity can substantially alter the partial pressures of gases such as CO_2 and O_2. The reader should be aware of the potential for subsurface conditions to change gas-phase composition, and the resulting composition of groundwater with which it comes into contact. However, atmospheric composition provides a reference point in the analysis of groundwater composition and contamination.

Rain and snow are also effective scavengers of atmospheric particulate matter. Naturally occurring dust particles and aerosols will contribute to the inorganic composition of rain and snow. These particles are probably responsible for the presence of metals such as sodium (Na), potassium (K), calcium (Ca), and magnesium (Mg) in precipitation.

In a "polluted" atmosphere, the concentrations of gaseous and particulate constituents may be markedly different from those of a clean atmosphere. For comparison, Table 13.7 provides a listing of atmospheric composition in a clean atmosphere and in a polluted atmosphere, based on the trace constituents. It is interesting to note that most of these compounds are formed and can be found in unpolluted air, but the concentrations of many oxidizing compounds (e.g., ozone), acid-forming gases (e.g., oxides of nitrogen and sulfur), and particles can be substantially higher in a polluted atmosphere than in an unpolluted atmosphere. The nature of atmospheric contamination will be specific to an area. Therefore, the composition of a polluted atmosphere cannot be quantitatively generalized and local information is needed. However, for a given atmospheric composition, the increases in atmospheric constituent concentrations attributable to anthropogenic activities will translate to parallel changes in rainwater and snow chemistry in downwind areas. Those who work in the groundwater arena should be aware of the link between groundwater behavior and the surrounding environment, including the atmosphere.

TABLE 13.7 Typical Concentration Ranges of Trace Atmospheric Constituents in Clean and Polluted Atmospheres

Species (formula)	Concentration (ppbv)	
	Clean Troposphere	Polluted Troposphere
Sulfur Dioxide (SO_2)	1–10	20–200
Carbon Monoxide (CO)	120	1000–10,000
Nitric Oxide (NO)	0.01–0.05	50–750
Nitrous Oxide (N_2O)	0.1–0.5	50–250
Ozone (O_3)	20–80	100–500
Nitric Acid (HNO_3)	0.02–0.3	3–50
Ammonia (NH_3)	1	10–25
Formaldehyde (HCHO)	0.4	20–50
Formic Acid (HCOOH)		1–10
Nitrous Acid (HNO_2)	0.001	1–8
Peroxyacetyl Nitrate ($CH_3C(O)O_2NO_2$)		5–35
Non-Methane Hydrocarbons		500–1200

From Seinfeld, J. H. 1986. *Atmospheric Chemistry and Physics of Air Pollution*, John Wiley & Sons, New York. With permission.

Source Water:Soil Interactions. Percolation of rainwater and snow into a soil formation is responsible for the existence of groundwater. As water moves through soil, a complex set of reactions take place which alter the chemical composition of both phases (solid and liquid). The general categories of reactions have already been reviewed: acid/base, complexation, precipitation/dissolution, and oxidation/reduction. The local implications of these reactions will depend on the specific compositions of the solid and liquid phases which come into contact. For example, many geologic formations consist largely of calcium- and magnesium-based minerals, such as limestone. Dissolved CO_2 (and other acids) will interact with these minerals to promote equilibration between the phases. This results in mineral dissolution and an increase in the liquid-phase concentration of the metals which make up the dissolved minerals. In groundwater systems, the presence of Ca^{2+} and Mg^{2+} are important not only in terms of the dietary needs they serve to people who consume water from a groundwater source, but also in terms of industrial applications. Chemical hardness, which is defined as the sum of all polyvalent cations in an aqueous system, is almost exclusively attributable to Ca^{2+} and Mg^{2+}. The constituents of chemical hardness cause problems of fouling and scale formation in industrial systems, and form stable complexes with detergents, thereby increasing detergent usage. Mineral dissolution will also lead to increases in the concentrations of inorganic constituents.

A summary of the major, minor, and trace constituents, according to the designations defined by Freeze and Cherry (1979), in groundwater is provided in Table 13.8. As can be seen from this table, the ranges of constituent concentrations are quite broad. Not surprisingly, groundwater composition is highly variable between sites, and is largely determined by the differences in geology among the sites.

Inorganic groundwater contaminants may come in the form of increases in the concentrations of the constituents listed in Table 13.8 above their naturally occurring concentrations, or the introduction of constituents which otherwise would not be detectable in the system at all. A number of factors can contribute to the changes in solution chemistry in these systems, including changes in atmospheric composition, changes in soil chemistry, and human activities which directly alter the physical or chemical conditions of a soil matrix.

Possibly the most significant change in atmospheric behavior which can affect groundwater composition is the introduction of strong acids (e.g., H_2SO_4 and HNO_3) through combustion of fossil fuels. These acids reduce the pH of source waters (rain and snow) and promote dissolution of the soil matrix. In contaminated soils, dissolution is not limited to naturally occurring minerals. For example, many military sites have extensive heavy-metal contamination in their soils as a result of firing range activities. The heavy metals at these sites are dominated by lead (Pb), but other elements are also present including

TABLE 13.8 Major, Minor, and Trace Constituents, as defined by Freeze and Cherry (1979)

Major Constituents (greater than 5 mg/l)

Bicarbonate	Silicon
Calcium	Sodium
Chloride	Sulfate
Magnesium	Carbonic acid

Minor Constituents (0.01–10.0 mg/l)

Boron	Nitrate
Carbonate	Potassium
Fluoride	Strontium
Iron	

Trace Constituents (less than 0.1 mg/l)

Aluminum	Molybdenum
Antimony	Nickel
Arsenic	Niobium
Barium	Phosphate
Beryllium	Platinum
Bismuth	Radium
Bromide	Rubidium
Cadmium	Ruthenium
Cerium	Scandium
Cesium	Selenium
Chromium	Silver
Cobalt	Thallium
Copper	Thorium
Gallium	Tin
Germanium	Titanium
Gold	Tungsten
Indium	Uranium
Iodide	Vanadium
Lanthanum	Ytterbium
Lead	Yttrium
Lithium	Zinc
Manganese	Zirconium

zinc (Zn), copper (Cu), and antimony (Sb). The solubility of these metals in their zero-valent states is quite limited. Therefore, for these metals to represent a potential source of groundwater contaminants, they must generally undergo oxidation. A number of mechanisms exist whereby these metals may be oxidized in a soil including microbiological activity, contact with chemical oxidants, and galvanic interactions with other metals. As a result of these and other processes, zero-valent metals are transformed, usually at slow rates, to their more mobile oxidized forms, where dissolution by aqueous constituents can become significant.

Among heavy metals, lead has generally been the subject of the most extensive and comprehensive research, probably because of the large number of potential sources and its general ubiquity. Other examples of potential sources of lead in soils and groundwater include lead-based paints and leakage of lead-based fuels from underground storage tanks. While both of these products are now out of production in the U.S., significant reservoirs of contamination (or at least potential contamination) still exist. Furthermore, these compounds are still being produced in many areas of the world.

Another source of inorganic contamination is mining and related activities. In the past, a common mining technique involved the application of liquids to promote selective chemical extraction of valuable mineral constituents. These practices, which are no longer in use *in situ*, sometimes created significant local soil and groundwater contamination problems, some of which are quite long-lasting, and still exist today. Modern mining techniques are substantially improved in terms of their potential to contaminate, but are not exempt from the ability to cause groundwater problems.

The presence of mineral acidity in the form of elevated proton activity (low pH) leads to the dissolution of many precipitates, including those which involve hydroxide, carbonate, sulfide, and phosphate ligands. The mineral composition of many soils incorporates large amounts of these materials, which themselves act as buffers against changes in pH. Therefore, the presence of mineral acidity in soils and groundwater is somewhat unusual. However, processes do exist which can create locally low pH. One prominent example of such a process is the oxidation of pyrite (FeS_2) and other iron sulfide crystalline forms. As will be discussed below, the process is often linked to mining activities, and the resulting low pH waters from these operations are commonly referred to as "acid mine drainage." The stoichiometry of the process is described as follows:

$$FeS_{2(s)} + \frac{7}{2}O_2 + H_2O \rightleftarrows Fe^{2+} + 2SO_4^{2-} + 2H^+ \tag{13}$$

$$Fe^{2+} + \frac{1}{4}O_2 + H^+ \rightleftarrows Fe^{3+} + \frac{1}{2}H_2O \tag{14}$$

$$Fe^{3+} + 3H_2O \rightleftarrows Fe(OH)_{3(s)} + 3H^+ \tag{15}$$

$$FeS_{2(s)} + 14Fe^{3+} + 8H_2O \rightleftarrows 15Fe^{2+} + 2SO_4^{2-} + 16H^+ \tag{16}$$
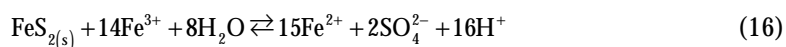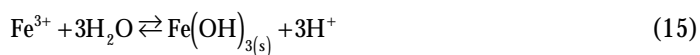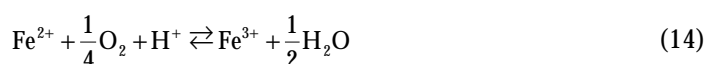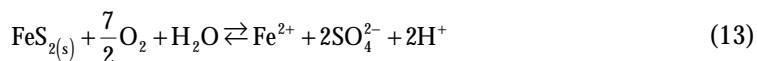
Reactions (13) through (16) do not accurately represent the mechanism by which pyrite oxidation leads to acidification. Temple and Delchamps (1953) hypothesized the mechanism illustrated in Figure 13.9 to describe the process and its rate limitations. Under the low-pH conditions which would be expected in waters subject to these reactions, the rate of Fe^{2+} oxidation is extremely slow. However, several species of bacteria are known to catalyze this reaction, including *Thiobacillus thiooxidans*, *Thiobacillus ferrooxidans*, and *Ferrobacillus ferrooxidans*. These microorganisms, which are categorized as microaerophiles, require oxygen to catalyze reaction (14), though they can survive under conditions of low oxygen concentration. The introduction of oxygen, often as a result of mining activities, initiates the process by dramatically increasing the rate of reaction (14). Once the process is started, the presence of oxygen is no longer a requirement for the development of mineral acidity. Therefore, sealing of mines which have experienced this problem does not represent an acceptable remediation approach.
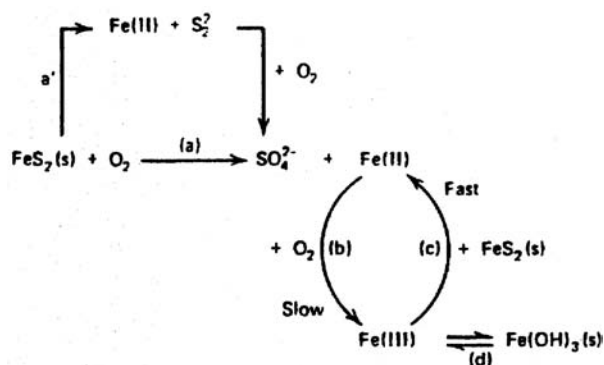
FIGURE 13.9 Mechanistic description of the processes responsible for pyrite oxidation, as hypothesized by Temple and Delchamps (1953).

The generation of acid mine drainage is often associated with coal mining activities since pyrite-bearing minerals are commonly found in conjunction with coal deposits. However, pyrite is also found in other mineral deposits, and similar acidification problems have been linked to acid mine drainage within copper and uranium mining facilities.

Another source of inorganic contamination is storage of salt for municipal operations, such as road deicing. Large salt piles, largely comprised of sodium chloride (NaCl), are often stored with little or no weather protection. Given the solubility of NaCl and related compounds in water, these situations provide a tremendous potential for salt introduction to a groundwater system. While the constituents of road salts do not generally represent an acute toxicity problem toward humans, they can increase salt concentrations to the point where a groundwater is not fit for human consumption. Additionally, the removal of sodium and potassium salts from groundwaters is difficult because they are so soluble in water — their affinity for water makes it difficult and expensive to separate them from the aqueous phase.

13.5 Organic Constituents

In uncontaminated soil/groundwater systems, the predominant organic constituents are humic substances. These compounds, which represent the products of natural decay processes from plant and animal matter, do not have a well-defined chemical structure. Therefore, humic substances are identified and differentiated according to a broad set of operational definitions. Three classes of humic substances exist:

1. Fulvic acid — constituents which are soluble at any pH
2. Humic acid — constituents which are insoluble at low pH, but otherwise soluble
3. Humin — constituents which are insoluble in water at any pH

In most groundwaters, the concentrations of humic substances (humic and fulvic acids) are quite low, normally less than several mg/l.

Though no structural definition exists for humic substances, extensive characterization work has been conducted on several classes of these compounds. Humic substances generally take the form of large heteroatomic organic polymers. The composition and structure of humic substances will differ depending of the source material and conditions under which it was generated. However, some generalizations have been developed regarding the functional groups which tend to exist within humic substances. For example, the structure presented in Figure 13.10 has been hypothesized to be representative of a monomer of a humic acid (Schnitzer and Khan, 1972). One of the most significant features of these hypothesized structural representations is the presence of oxygenated functional groups. These groups are believed to be largely responsible for the ability of humic substances to complex metals.

Many classes of synthetic organic compounds have been identified as groundwater contaminants, including pesticides, petroleum products, and (halogenated) solvents. Though it is not practical to list
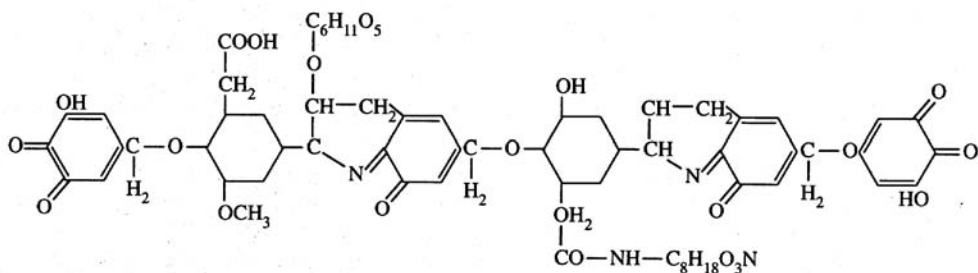
FIGURE 13.10 Representation of the structure of humic acid. (From Schnitzer, M. and Khan, S. U. 1972. *Humic Substances in the Environment*, Marcel Dekker, New York. With permission.)

all possible organic groundwater contaminants, a number of their important characteristics can be illustrated with examples from the aforementioned categories.

The mobility of groundwater contaminants is largely determined by their affinity for soil particles. Generally speaking, contaminants with high affinity for soil particles preferentially associate with the soil matrix, and therefore move relatively slowly (see Figure 13.6). Among organic compounds, relative affinity for soil vs. water can be evaluated using the octanol:water partitioning coefficient (K_{ow}). K_{ow} provides a measure of the equilibrium distribution of a specific compound between octanol and water phases in direct contact. K_{ow} values for a wide variety of important compounds, including many pesticides are presented in Table 13.9. Large values of K_{ow} are indicative of compounds which are likely to preferentially associate with organic constituents (because of their relatively low polarity), such as those found in many soils; small values indicate a preference for the more polar aqueous phase. Common sources of hydrophobic organics, such as those listed in Table 13.9 include land application, either directly or by incorporation with soil amendments (such as process sludges) and surface spills.

TABLE 13.9 Octanol-Water Partition Coefficients (K_{ow}) for Selected Soil and Groundwater Contaminants

Compound	$\log_{10} K_{ow}$
n-Octane	5.18
Acetone	−0.24
Benzene	2.13
Toluene	2.69
Xylenes	3.12–3.18
Phenol	1.45
Napthalene	3.36
Polychlorinated Biphenyls	4.09–8.23
2,3,7,8-Tetrachlorodibenzo-*p*-dioxin	6.64
Lindane	3.78
Dieldrin	5.48
p-*p*′-DDT	6.36
Parathion	3.81
Malathion	2.89
Atrazine	2.56
RDX	0.87

From Schwarzenbach, R. P., Gschwend, P. M., and Imboden, D. M. 1993. *Environmental Organic Chemistry*, John Wiley & Sons, New York. With permission.

Compounds which display extremely low aqueous solubility can exist as a separate liquid phase in groundwater systems, if present in sufficient quantities. In groundwater systems, these contaminant phases are referred to as nonaqueous phase liquids (NAPLs), and their behavior will be fundamentally different from that of the bulk aqueous phase. In the case of an NAPL with density less than the surrounding water, the nonaqueous phase is referred to as a light NAPL (LNAPL), and will generally be found at or near the phreatic surface (see Figure 13.11). Fluctuations in the height of the phreatic surface will cause the LNAPL to disperse. LNAPL spills often are attributable to leaking underground storage tanks (LUSTs) used for storage of petroleum hydrocarbons (fuels). Cleanup activities focus on location of the LNAPL phase and removal, by pumping if possible. Hydrocarbons which cannot be removed by these methods may be subjected to soil vapor extraction, *in situ* bioremediation, or a combination.

Low-solubility organic compounds with density greater than water can exist as a dense NAPL (DNAPL) phase; their characteristics are fundamentally different from those of LNAPLs. DNAPL compounds, because of their high density, tend to sink in groundwater systems. Often these compounds have lower viscosity than water — as a result, they tend to sink according to a pattern known as viscous fingering (see Figure 13.12). As they sink, they leave a trail of residual DNAPL compound among the solids (soil
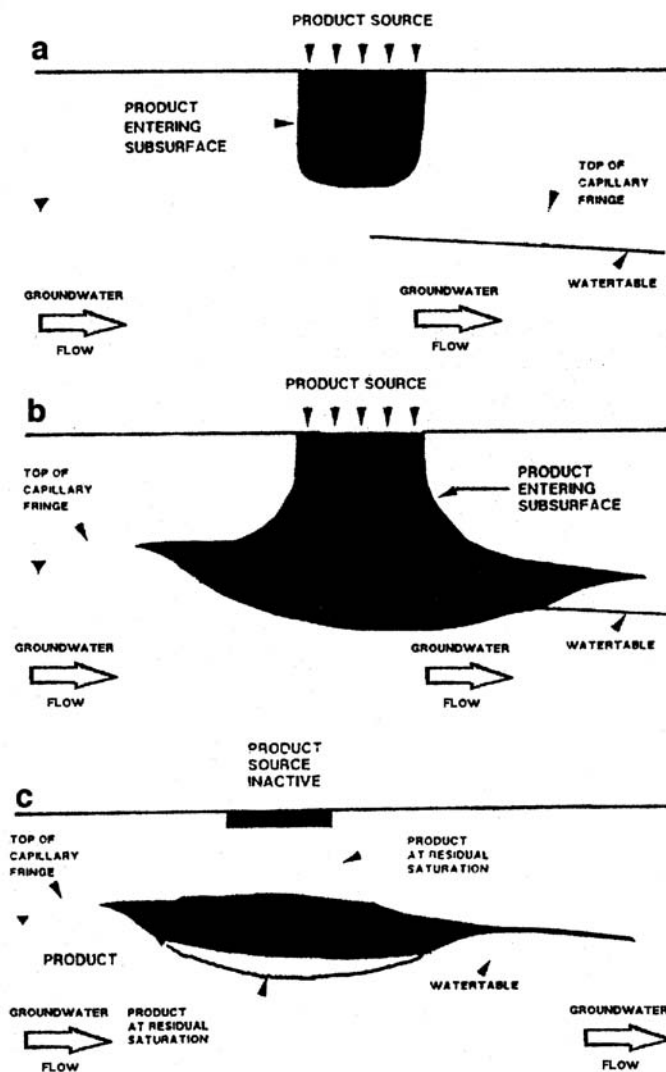
FIGURE 13.11 Schematic illustration of the behavior of LNAPL compounds. Under some conditions (a), the mass of an LNAPL spill is insufficient to allow penetration to the capillary fringe. With additional compound introduction (b), LNAPL product will reach the water table and begin to spread, though the compound will not penetrate far beyond the phreatic surface. If the source of LNAPL is eliminated (c), removal of LNAPL will allow for "rebound" of the water table. (From U.S. Environmental Protection Agency. 1989. *Transport and Fate of Contaminants in the Subsurface*, EPA/625/4-89/019.)

particles) and liquid (water) in their wake. A DNAPL phase will continue to sink until it completely disperses among the soil and water phases, or it reaches an impermeable boundary, such as a clay lens or bedrock (see Figure 13.12). DNAPL pools are extremely difficult to locate and are often difficult (or impossible) to remediate using existing technologies. Furthermore, because DNAPLs are often comprised of halogenated solvents, microbiological activity is slow to bring about changes in their molecular structure, and the products of microbial degradation involving these compounds are sometimes more toxic than their parent compounds. Possibly the most common examples of DNAPL compounds are trichloroethylene (TCE) and perchloroethylene (PCE). As in the case of LNAPLs, the source of DNAPL spills is often leakage or failure of underground storage tanks.
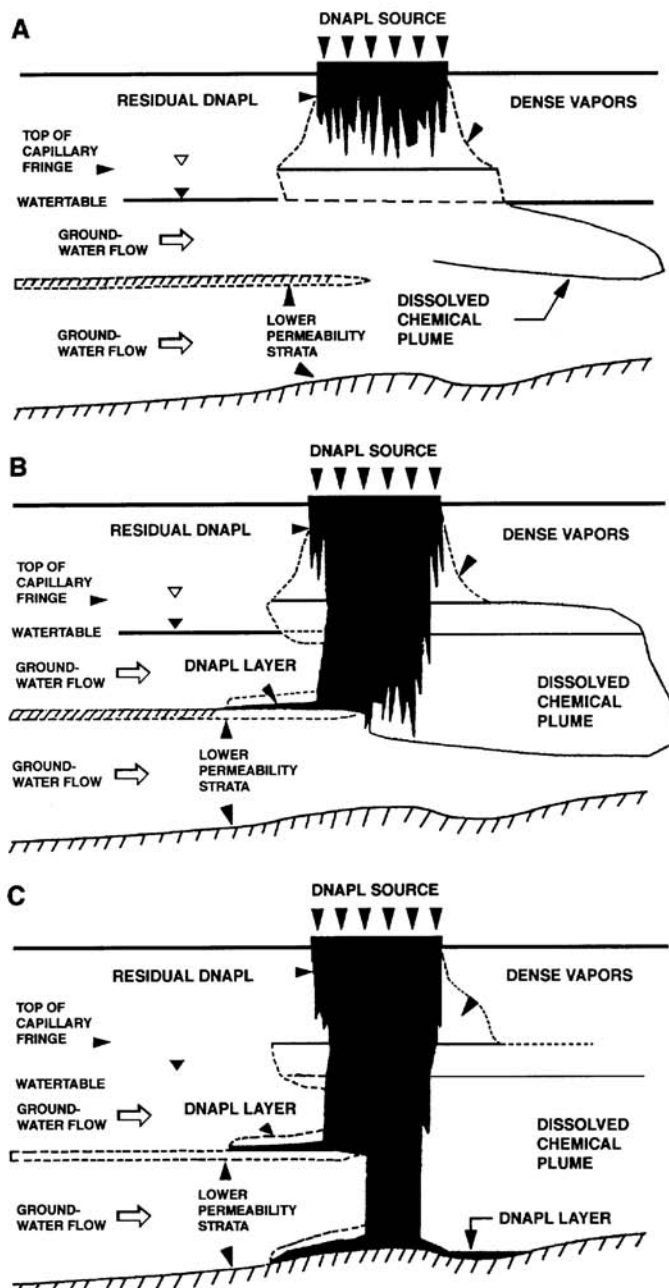
FIGURE 13.12 Schematic illustration of the behavior of DNAPL compounds. Under some conditions (A), the mass of a DNAPL spill is insufficient to allow penetration to the intact DNAPL to the capillary fringe; vertical movement of the DNAPL is by viscous fingering. With additional compound introduction (B, C), DNAPL product will reach the water table and continue to move vertically until it reaches an impermeable boundary. (From U.S. Environmental Protection Agency. 1989. *Transport and Fate of Contaminants in the Subsurface*, EPA/625/4-89/019.)

13.6 Radionuclides

13.6.1 Radioactive Decay

Each chemical element in the periodic table has a fixed number of protons, but can exist with a variety of neutrons. The term *isotope* refers to a nucleus with the same number of protons, but different numbers of neutrons. Radionuclides are typically written with the chemical symbol and a superscript to the left of the symbol that denotes the atomic mass (total number of protons and neutrons) of that isotope. For example, ^{12}C and ^{14}C are isotopes of the element carbon; each has 6 protons, but are associated with 6 and 8 neutrons, respectively. Elements that undergo radioactive decay as a result of an unstable nuclear structure are referred to as radionuclides. As part of the decay process, radionuclides release ionizing radiation in the form of alpha particles (α), beta particles (β), and gamma radiation (γ).

Alpha particles are composed of two protons and two neutrons. They interact with matter through electrostatic interactions between their positive charge and the negative charge of orbital electrons in the medium through which they travel. Because of this, α particles deposit large amounts of energy in a very small travel distance; the mean travel distance of a 4.7 million electron volt (MeV) α particle, such as that emitted from ^{237}Np, is only 3.3 cm through air (Knoll, 1989).

Beta decay results in the production of high-energy electrons or positrons. Interactions of β particles result in deposition of the β particle energy in the medium through which it travels. Beta particles can also undergo electron-nuclear reactions as well. Beta particles lose energy at a lower rate as compared to α particles, and therefore have greater penetrating power. Positrons are the antiparticle of the electron. Once a positron has deposited its energy in the medium, it will undergo an annihilation reaction with an electron, producing two 0.511 MeV photons. Electrons are expelled from the nucleus in the conversion of a neutron to a proton, while a positron is emitted for a proton to neutron conversion. Weakly interacting particles known as (anti)neutrinos are also produced in β decay, but are not the focus of this discussion because they have an extremely small interaction probability with matter.

Gamma rays are high-energy photons released by excited nuclei as they transition from an excited state in a radioactive decay process. The γ photons are usually emitted in conjunction with an α or β decay process, and interact primarily with atomic electrons through photoelectric absorption, Compton scattering, and pair production. Unlike α and β particles, which slow down gradually via Coulomb forces, γ radiation interacts by partial or complete energy transfer to atomic electrons in sudden, discrete interactions. As γ rays deposit energy in a medium, they transfer energy to atomic electrons, producing high-energy electrons in the medium that interact in the same manner as β decay high-energy electrons (Knoll, 1989). Gamma radiation has a much larger penetrating power relative to α or β radiation; the mean free path of a 0.66 MeV photon, such as that emitted by ^{137}Cs, is around 100 meters in air.

Each of the types of radiation described above is known as ionizing radiation. That is, they are capable of transferring enough energy to atomic electrons to allow the electrons to escape from the atom. As radiation interacts with the medium through which it travels, it creates a path of excitation (promotion of electron to a higher energy shell) and ionization in its wake.

The rate of radioactive decay of a group of atoms is proportional to the number of atoms of that isotope that are present in the group. This principle can be represented mathematically by defining a decay constant, λ, as the proportionality constant for decay. Equation (17) relates the time rate of change of the number of nuclei to the decay constant.

$$\frac{dN}{dt} = -\lambda N \tag{17}$$

This equation yields a simple analytical solution shown in Equation (18).

$$N = N_o e^{-\lambda t} \tag{18}$$

Using Equation (18), it is possible to determine the amount of time for half of a radionuclide to decay, as indicated by Equation (19). This time is referred to as the half-life of an isotope, $t_{1/2}$.

$$t_{1/2} = \frac{\ln 2}{\lambda} \tag{19}$$

As an element decays, the parent nucleus is transformed into a daughter nucleus, which may or may not be stable. If the daughter nucleus is unstable, it can then be considered a parent nucleus with an associated stable or unstable daughter nucleus. In this way, decay chains from a parent nucleus can involve numerous decays and radioactive emissions.

13.6.2 Definitions and Units

With these equations defined, it is necessary to understand the units which are commonly used to describe radionuclide activity. The activity, A, of a radioactive substance is defined as the number of decays of that radionuclide per unit time; the mathematical definition of activity is shown in Equation (20).

$$A = \lambda N \tag{20}$$

Activity is typically reported in units of Curies (Ci), where each Ci is defined as 3.7×10^{10} disintegrations per second. The SI unit for activity is the Becquerel (Bq), defined as one disintegration per second. It is important to note that a short half-life (high λ) radionuclide produces much higher activity relative to a long half-life (low λ) radionuclide when they are present in equimolar concentrations.

It is sometimes necessary to be able to calculate the amount of energy deposited in a medium exposed to radiation in order to analyze ionizing radiation effects. The amount of energy deposited per unit mass is referred to as dose, D, and is measured in Grays (Gy). A Gray is defined as one Joule of energy deposited per kg of material and is equivalent to 100 rads (radiation absorbed dose).

The biological damage done by ionizing radiation to organisms (e.g., humans) is called the dose equivalent, H, and is related to the dose by a quality factor, Q (see Equation [21]).

$$H = DQ \tag{21}$$

Dose equivalent is measured is Sieverts (Sv) using the product of the quality factor and dose in Grays (rems, or roentgen equivalent man, are also dose equivalent units where one rem is 0.01 Sv). The quality factor is related to the linear energy transfer (LET) of the radiation passing through the organism. High LET radiation does greater damage than lower LET radiation, and this is reflected by a higher quality factor for external α and β radiation as compared to γ radiation. Radiation energy is typically reported in units of millions of electron volts (MeV), while LET is typically given in units of keV per cm.

13.6.3 Important Radionuclides and their Sources

There exist a variety of sources of radionuclides which can be dissolved in groundwater. Sources are typically categorized as natural or anthropogenic. Natural sources of radionuclides include formation in the atmosphere caused by cosmic irradiation, and presence in mineral deposits of long-lived radioisotopes of heavy elements and their associated daughters. Anthropogenic sources consist of nuclear power plants, nuclear weapons explosions and testing, medical, research, industrial, and pharmaceutical radionuclide applications, and mining wastes. A number of specific radionuclides are known to represent potential problems in groundwater systems. Each is described below.

13.6.3.1 Radon (^{222}Rn)

The only long-lived isotope of radon is ^{222}Rn, which undergoes α decay with a half-life of 3.82 days. This radionuclide is produced by decay of ^{226}Ra, and is often associated with deposits that have significant

uranium concentrations (^{226}Ra and ^{222}Rn are part of the ^{238}U decay chain). ^{222}Rn represents a significant health concern in some homes, entering as a dissolved species in drinking water or directly from soil as a gas. Additional radiation is emitted from daughter products of ^{222}Rn as it decays to a stable element. ^{222}Rn is a naturally occurring radioisotope, and can be significant for uranium mining in terms of worker safety and waste disposal (Fetter, 1993).

13.6.3.2 Radium (^{226}Ra and ^{228}Ra)

^{226}Ra ($t_{1/2}$ = 1600 years) is a daughter of ^{230}Th, part of the ^{238}U decay chain, and decays subsequently to ^{222}Rn. Because ^{228}Ra ($t_{1/2}$ = 5.8 years) is a daughter of ^{232}Th, both radium radionuclides are associated with uranium- and thorium-containing deposits. Radium is more soluble than either thorium or uranium, and can therefore be transported in higher concentrations with an associated higher activity, particularly when considering the relatively short $t_{1/2}$ of ^{228}Ra. Because radium is associated with uranium deposits, it is therefore a potentially important problem in mining wastes (Freeze and Cherry, 1979).

13.6.3.3 Uranium (^{238}U and ^{235}U)

Uranium in mineral deposits is composed of 99.3% ^{238}U ($t_{1/2}$ = 4.5×10^9 years) and 0.7% ^{235}U ($t_{1/2}$ = 7.1×10^8 years). Because ^{235}U is fissile (split by low-energy neutrons), it is used as fuel for nuclear power reactors worldwide. Both uranium isotopes are long lived and part of long decay chains. Each has numerous radioactive daughter nuclei that produce emissions before decaying to stable states. Uranium is typically mined in open pits or underground and concentrated to 65 to 70% uranium via flotation, dissolution, and solvent extraction processes. Further purification to U_3O_8 and enrichment to increase the fraction of the ^{235}U isotope can achieve desired levels of ^{235}U, with 3 to 4% being typical for nuclear power reactors (Choppin et al., 1995).

13.6.3.4 Neptunium (^{237}Np)

^{237}Np can be found in nuclear waste as the product of nuclear reactions (in a nuclear power reactor) converting ^{238}U to ^{237}U, which then β decays to ^{237}Np. It has a $t_{1/2}$ of 2.2×10^6 years and undergoes α decay to ^{233}Pa, with associated daughter decays from that element and others. ^{237}Np is of great concern in high level and spent nuclear fuel disposal because it is only slightly sorbed (i.e., high mobility) in many groundwater environments and has a long half-life (USDOE, 1995).

13.6.3.5 Tritium (^{3}H)

Tritium undergoes β decay with a half-life of 12.43 years, producing ^{3}He as a daughter element. Because tritium can combine with hydrogen and oxygen to form tritiated water, it poses a significant threat in terms of groundwater transport. Tritium is especially useful in delineating groundwater recharge occurring since the 1963 spike of tritium released by the U.S. and Soviet Union prior to the implementation of an above-ground testing treaty (Fritz et al., 1991). Tritium concentrations in natural waters have increased nearly tenfold since the advent of the nuclear age (Fetter, 1993). Tritium is formed naturally in the atmosphere when high-energy neutrons bombard ^{14}N, producing ^{12}C and ^{3}H. It is also produced in large quantities at nuclear power plants, and in nuclear weapons testing and production.

13.6.3.6 Carbon (^{14}C)

Like tritium, ^{14}C can be used for dating purposes, in this case for organic material. It decays by β emission to ^{14}N, and has a $t_{1/2}$ of 5715 years. ^{14}C labeling of organic compounds is often done to provide sensitive detection methods for experiments with organic compounds. Because it can be oxidized to soluble or gaseous $^{14}CO_2$, or form a variety of organic compounds, ^{14}C can pose a significant transport threat as a radionuclide. It can be produced naturally in the atmosphere when low-energy neutrons bombard ^{14}N, producing ^{14}C and ^{1}H. Anthropogenic sources include nuclear power plants, nuclear weapons testing, and research laboratories.

13.6.3.7 Fission Fragments (^{90}Sr ^{99}Tc, ^{129}I, and ^{137}Cs)

Fission fragments are radioactive materials produced by fission of heavy isotopes such as ^{235}U. The fission reaction typically splits the heavy atom into two (sometimes three) lighter isotopes, with associated neutron, neutrino, and γ radiation emissions as well. The lighter isotopes are produced primarily in the atomic mass range of 80 to 155. Although a variety of fission fragments are produced, ^{90}Sr ^{99}Tc, ^{129}I, and ^{137}Cs are of primary concern due to the significant amounts that are produced and their relatively long half-lives.

^{90}Sr decays by β emission to ^{90}Y with a $t_{1/2}$ of 29.1 years. ^{90}Y then decays to stable ^{90}Zr via β emission. Strontium is a prominent fission fragment with no associated γ emissions. In high-level nuclear waste and spent fuel, it is one of the highest activity radioisotopes for the first 200 years after fission ceases. It can be particularly dangerous when ingested by humans because it is known to accumulate in or near the bones, resulting in greater irradiation of sensitive bone marrow.

^{99}Tc undergoes β decay to ^{99}Ru with a $t_{1/2}$ of 2.1×10^5 years. In addition to being a fission fragment, ^{99}Tc is also associated with medical and research wastes where it is used in diagnostic testing. ^{129}I is an important fission fragment because it β decays to stable ^{129}Xe with a long $t_{1/2}$ of 1.7×10^7 years. It is also used in medical diagnostic tests and can therefore be found in low-level waste as well as spent fuel.

^{137}Cs emits γ and β radiation while decaying to stable ^{137}Ba and has a $t_{1/2}$ of 30.3 years. Along with ^{90}Sr, ^{137}Cs is one of the high-activity isotopes in high-level nuclear waste for the first 200 years after fission ceases.

13.6.4 Radioactive Waste Disposal

There are four different classes of radioactive wastes defined in United States regulations. *High Level Waste* (HLW) is defined as radioactive waste produced in the reprocessing of nuclear fuel to recover usable uranium and plutonium. *Spent Nuclear Fuel* (SNF) is nuclear fuel that has been removed from a power reactor but has not been reprocessed. *Transuranic Waste* (TRU) consists of wastes from nuclear power or weapons productions operations with atomic number greater than 92 (uranium), being primarily α emitters with long half-lives; it is typically associated with nuclear weapons production, fuel fabrication, and reprocessing. *Low Level Waste* (LLW) is high volume, low radioactivity waste produced by industry, consumer products, research and medical labs, government, and nuclear power plants (Tang and Saling, 1990).

The United States does not reprocess commercial nuclear fuel, although Japan and most of Europe have chosen this option. As a result, U.S. wastes associated with the nuclear power industry are primarily in the form of SNF, but with significant amounts of HLW and TRU from government (defense) related activity. The following discussion will focus on radioactive waste disposal options being pursued in the U.S., but different strategies are being considered all over the world.

13.6.4.1 HLW and SNF Disposal

The U.S. government has agreed to take over the spent nuclear fuel being stored at commercial reactors throughout the country by 1998. Ultimate disposal of HLW and SNF is unlikely to occur in a timely fashion, though the ultimate responsibility lies with the federal government to solve the problem. The U.S. government accepted ultimate responsibility for taking the waste by passing the Nuclear Waste Policy Act of 1982; to pay for this, utilities are taxed 0.1¢ per kWh of electricity produced. As of 1994, over $10 billion had been collected in the fund to pay for waste disposal activities.

The current U.S. plan for ultimate disposal of HLW and SNF is to place it in the deep geologic repository at Yucca Mountain in Nevada. The repository will rely on an engineered barrier system (EBS) combined with site characteristics (geology) to contain the waste. A deep repository has the advantages of providing the best safety, flexibility, retrievability, convenience, maintenance, and feasibility relative to other disposal options (Faure, 1991). Because the Yucca Mountain site is still undergoing site characterization and construction, it will not be ready to accept waste in 1998. The entire HLW and SNF disposal plan is far behind schedule due to technical difficulties, public fear and protests, and resistance from the host state.

Because of this, there are plans to place the waste temporarily in monitored, retrievable storage (MRS). Once the Yucca repository is complete, waste will be transferred from the MRS to the final repository.

For the first few hundred years at the repository, ^{90}Sr and ^{137}Cs will be the highest activity components of the waste. However, it is generally assumed that the EBS and site characteristics will allow for complete containment over that time period. In the time frame of 10,000 years, isotopes such as ^{14}C, ^{99}Tc, and ^{129}I are believed to have the potential to escape containment because of their long half-lives (except ^{14}C), high solubility, and low sorption. ^{237}Np is another potential escape candidate because it has a long half-life and is only moderately sorbing in the site geology. For a time frame of one million years, ^{99}Tc, ^{129}I, and ^{237}Np pose the greatest threat for potential breakthrough from the repository.

13.6.4.2 TRU Disposal

TRU wastes have been held in interim storage since 1974 while an ultimate disposal solution was pursued. The wastes will be placed in the Waste Isolation Pilot Plant (WIPP), located in a salt bed repository in New Mexico. Site characterization was completed in 1978 and an environmental impact statement filed in 1980. Full construction started in 1984, and the site is on schedule to begin accepting waste as planned in 1998. TRU disposal plans are notable because it is the only successful radioactive waste disposal program in the U.S. that has not been significantly delayed or halted.

13.6.4.3 LLW Disposal

LLW disposal is the responsibility of the states, with technical assistance from the federal government. To alleviate some of the difficulty of each state constructing an LLW landfill, the states have formed compacts where numerous states can contribute waste to a single site. Much like federal HLW disposal efforts, the states are well behind schedule in meeting their planned opening dates for disposal sites. Since 1962, six LLW disposal sites have operated: Maxey Flats, KY, West Valley, NY, Sheffield, IL, Beatty, NV, Barnwell, SC, and Richland, WA. The LLW disposal program has been plagued by poor site characterization and location, drum breaching, cover breaching caused by subsidence, and lack of knowledge of microbial action to mobilize radionuclides. Maxey Flats, West Valley, and Sheffield have all had problems with tritium leaks from the disposal area (Newman, 1994). Future LLW disposal sites will use the lessons of the past to build safer containment areas for radionuclides. In particular, a greater emphasis will be placed on site characterization/selection, and greater restrictions on what types and forms (no liquids) of waste are being placed in the site.

13.6.5 Remediation

Remediation of radionuclide contaminated sites can be a costly and difficult endeavor. The chemical behavior of an atom in the environment is dictated by its atomic structure (not nuclear structure). Therefore, different isotopes of the same element behave identically in terms of chemical reactions. The essential difference between isotopes is that unstable elements will undergo radioactive decay. Because of this, radionuclide remediation techniques consist of the same methods which are used for inorganics and have been described previously in this chapter.

A fundamental difference between radionuclide and inorganic wastes is that radionuclides do not require direct contact to cause damage to an organism; radiation provides its own transport mechanism to reach a target. Because of this, providing for public and worker safety at a site is often more difficult than for conventional pollutants. In addition, remediation techniques can concentrate a radionuclide (e.g., ion exchange or activated carbon) to such an extent as to create a hazardous radiation field; appropriate shielding and transportation arrangements of the concentrate medium must be considered.

It is sometimes possible to use the unstable characteristics of radionuclides to facilitate remediation. For radionuclides with short half-lives, it may only be necessary to contain the waste for an appropriate period of time before the unstable nuclei have decayed below the background radiation. In this manner, it is possible to use the decay properties of the nucleus to facilitate safe disposal of some radionuclide wastes. The idea of transmutation (bombarding radioactive waste with neutrons or protons) has been

proposed as a way to stabilize radioactive nuclei, but has been discarded primarily because of high costs associated with this process.

13.7 Particulates

The definition of a particle is arbitrary and will depend on the application. For purposes of this presentation, particulate matter will refer to constituents which exist as distinct solid phase either within or apart from the groundwater itself. Though no absolute distinction is possible to define particles based on size, constituents with a size of 10 nm or greater may be considered to display particle behavior in groundwater system. Figure 13.13 provides a graphical representation of the spectrum of particle sizes which may be found in groundwaters, some important examples, and an indication of the physicochemical processes by which these constituents may be separated from the aqueous phase.
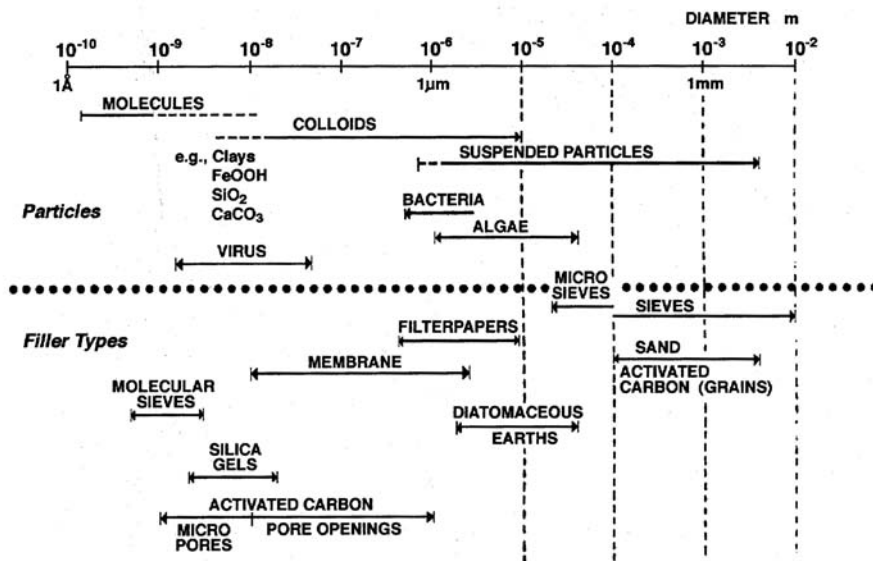
FIGURE 13.13 Spectrum of particle sizes to be seen in natural waters, including groundwaters. Physical processes for separation of particles of a given size from the aqueous phase are also indicated. (From Stumm, W. 1977. *Environ. Sci. Technol.* 11, 1066. With permission.)

Microorganisms play an important role as particulate contaminants in groundwater systems. Important categories of microbial groundwater contaminants include viruses, bacteria, and protozoan cysts. Based on the arbitrary definition given above, viruses are representative of the smallest particles which are likely to be present in groundwater systems. Their small size facilitates mobility in groundwater systems; however, even though viruses are smaller than the pore size openings common to groundwater systems, their movement in aquifers is not unimpeded by soil particles. Like all constituents in groundwater systems, movement is largely governed by association with soil grains. Viruses display surface characteristics (large surface:volume ratio, net surface charge) which allow them to associate with soil grains. In the general case, groundwater supplies are assumed to be subject to possible contamination by pathogenic viruses, and disinfection processes for public water supplies are designed to address this source of contamination. Microorganisms from the other categories (e.g., bacteria and protozoan cysts) will also move in groundwater systems; however, the larger size of microorganisms from these categories dictates that their movement will be generally slower than than that of viruses. Groundwater supplies which are under the direct influence of surface waters are subject to more stringent disinfection require-

ments for potabilization due to the possible presence of protozoan cysts such as *Giardia lamblia* and *Cryprosporidium parvum*, which are generally more difficult to inactivate than either bacteria or viruses.

Other, nonmicrobial colloidal particles can also exist and be transported in groundwater systems. Movement of these particles represents an important mechanism of transport for constituents which associate strongly with them. Important examples of compounds for which groundwater transport is enhanced by the presence of colloids include hydrophobic organics compounds, metals, and microorganisms.

For Further Information

Clark (1996) is a reference text in which transport phenomena are presented in terms of environmental processes. U.S. EPA (1989) provides basic information on many of the fundamental phenomena which govern the fate and transport of subsurface contaminants. Freeze and Cherry (1979) is perhaps the most frequently cited reference text on the subject of groundwater, and includes and excellent presentation of many important aspects of groundwater chemistry. Martell and Smith (1974–1989) represents a critical evaluation of stability (equilibrium) constants for complexation and precipitation reactions involving organic and inorganic aqueous constituents. This series of six volumes represents the most comprehensive summary of data on the subject. Snoeyink and Jenkins (1980) is an intermediate-level text on the subject of aquatic chemistry. Stumm and Morgan (1996) is the third edition of this advanced text on the subject of aquatic chemistry and is perhaps the most frequently cited reference on the subject.

References

Choppin, G., Rydberg, J., and Liljenzin, J. 1995. *Radiochemistry and Nuclear Chemistry, 2nd Edition*, Butterworth-Heinemann Limited, Oxford, UK.

Clark, M. M. 1996. *Transport Modeling for Environmental Engineers and Scientists*, John Wiley & Sons, New York.

CRC Handbook of Chemistry and Physics, 59th Edition. 1978. R. C. Weast and M. J. Astle (Eds.), CRC Press, Boca Raton, FL.

Faure, G. 1991. *Principles and Applications of Inorganic Geochemistry*, Macmillan Publishing Company, New York.

Fetter, C. 1993. *Contaminant Hydrogeology*, Macmillan Publishing Company, New York.

Freeze, R. and Cherry, J. 1979. *Groundwater*, Prentice-Hall, Englewood Cliffs, NJ.

Fritz, S. J., Drimmic, R. J., and Fritz, P. 1991. Characterizing shallow aquifers using tritium and ^{14}C: Periodic sampling based on tritium half-life, *Applied Geochemistry*, 6, 17-33.

Knoll, G. 1989. *Radiation Detection and Measurement, 2nd Edition*, John Wiley & Sons, New York.

Larson, T. E. and Buswell, A. M. 1942. Calcium carbonate saturation index and alkalinity interpretations, *Journal, AWWA*, 34, 1664.

Martell, A. E. and Smith, R. M. 1974–1989. *Critical Stability Constants*, Volumes 1-6, Plenum Press, New York.

Newman, A. 1994. Low-level radioactive wastes, high-level risk?, *Environmental Science & Technology*, 28, 11, 488A.

Schnitzer, M. and Khan, S. U. 1972. *Humic Substances in the Environment*, Marcel Dekker, New York.

Schwarzenbach, R. P., Gschwend, P. M., and Imboden, D. M. 1993. *Environmental Organic Chemistry*, John Wiley & Sons, New York.

Seinfeld, J. H. 1986. *Atmospheric Chemistry and Physics of Air Pollution*, John Wiley & Sons, New York.

Snoeyink, V. L. and Jenkins, D. 1980. *Water Chemistry*, John Wiley & Sons, New York.

Stumm, W. 1977. Chemical interaction in particle separation. *Environ. Sci. Technol.* 11, 1066.

Stumm, W. and Morgan, J. J. 1996. *Aquatic Chemistry, 3rd Edition*, John Wiley & Sons, New York.

Stumm, W. and Morgan, J. J. 1981. *Aquatic Chemistry, 2nd Edition*, John Wiley & Sons, New York.

Tang, Y. and Saling, J. 1990. *Radioactive Waste Management*, Hemisphere Publishing Company, New York.

Temple, K. L. and Delchamps, E. W. 1953. Autotrophic bacteria and the formation of acid in bituminous coal mines. *Appl. Microbiol.*, 1, 255.

U.S. Department of Energy (USDOE). 1995. *Total System Performance Assessment: An Evaluation of the Potential of the Yucca Mountain Repository,* prepared by TRW for USDOE, B00000000-01717-2200-00136, Rev. 01, Nov., 1995.

U.S. Environmental Protection Agency. 1989. *Transport and Fate of Contaminants in the Subsurface,* EPA/625/4-89/019.

Glossary

Acid/Base Reaction   A reaction in which proton transfer takes place.

Amphoterism   Ability of a solid surface or solute to display the characteristics of an acid or base.

Chemisorption   Association of solute (adsorbate) molecules with a solid (sorbent) as a result of relatively strong chemical bonds, similar in strength to covalent bonds.

Dense, Non-Aqueous Phase Liquid (DNAPL)   (Organic) liquids which are more dense than water, often being comprised of halogenated organic solvents.

Groundwater Contaminant   Groundwater constituent, often viewed as undesirable, introduced to the subsurface either directly or indirectly as a result of human activity.

High Level Waste (HLW)   Radioactive waste produced in the reprocessing of nuclear fuel.

Humic Substances   Stable organic products of the natural decay of plants and animals.

Ion Exchange   Association of charged solutes (ions or complexes) with a solid (sorbent) as a result of electrostatic interactions.

Light Nonaqueous Phase Liquid (LNAPL)   (Organic) liquids which are less dense than water, often being comprised of hydrocarbons of hydrocarbon mixtures, such as those which result from petroleum refinement.

Low Level Waste (LLW)   High volume, low radioactivity waste produced from many industrial and commercial processes.

Micelle   Pseudo-phase agglomeration of (organic) surfactant molecules in which polar "ends" of the surfactant molecules are directed outward and the nonpolar "ends" are directed inward.

Nonaqueous Phase Liquid (NAPL)   Usually comprised of organic compounds having extremely low aqueous solubilities.

NO_x   Nitrogen oxides — common atmospheric constituents produced by a number of natural and anthropogenic processes, including combustion of fossil fuels.

Oxidation/Reduction (Redox) Reaction   Reaction in which electron transfer takes place.

Particle   Constituents which exist as a distinct solid phase within or apart from a solvent; oftentimes particles are operationally defined as constituents greater than a given size (e.g., 10 nm).

pϵ   Negative $\log_{10}$ of electron (e^-) activity; a representation of electron availability in solution.

pH   Negative $\log_{10}$ of proton (H^+) activity; a representation of proton availability in solution.

Physisorption   Association of solute (adsorbate) molecules with a solid (sorbent) as a result of relatively weak physical bonds.

Polarity   A relative term describing the distribution of electrons (or electronic charge) around a molecule. Highly polar molecules generally contain atoms of widely varying electronegativity.

Radionuclides   Unstable elemental isotopes which release ionizing radiation.

SO_x   Sulfur oxides — common atmospheric constituents produced by a number of natural and anthropogenic processes, including combustion of fossil fuels.

Transuranic Waste (TRU)   Wastes from nuclear power or weapons production having an atomic number greater than 92.

Zero Point of Charge (zpc)   pH at which net surface charge is zero.

14

Nonreactive Contaminant Transport in the Saturated Zone

Vivek Kapoor
Georgia Institute of Technology

14.1 Introduction

Dealing with groundwater pollution problems requires anticipating how the concentration of a contaminant introduced in natural geological material will change in space and time. For example, in designing waste-disposal in geological medium, the barrier provided by the natural medium can be an important parameter in determining its long-term feasibility. In aquifer decontamination, it is seldom possible to completely remove all of the contamination, and the evolution of the residual contaminant concentration can play an important role in determining the feasibility of the decontamination scheme. The aerobic degradation of hydrocarbons in groundwater will be controlled by the dissolved oxygen concentration and the concentration of other nutrients, in addition to the contaminant concentration. Such practical matters have motivated developing an understanding of the evolution of the concentration of solute introduced in water flowing through natural porous medium.

This chapter summarizes the current understanding of transport of a dissolved or suspended nonreactive solute in saturated porous medium. The basic transport mechanisms and their representation in the transport equation that forms the basis of a continuum treatment of solute transport in porous medium are first discussed in the section on advection and local dispersion. The role of aquifer property heterogeneity in spreading contaminants at rates greater than what would be effected by local dispersion

alone in a homogeneous medium is then addressed. The complex distribution of concentration in heterogeneous media is illustrated. The description of heterogeneity as random fields is illustrated by providing a numerical example that shows the interactions between heterogeneous advection, local dispersion, spreading, and dilution, that need to be considered in assessing contaminant transport in the field.

14.2 Advection and Local Dispersion

14.2.1 Seepage Velocity

As solute in saturated porous medium is advected by the fluid flowing through it, the fluid velocity needs to be described to study contaminant transport. The pore-scale fluid velocity is spatially varying due to three interrelated reasons: (1) the no-slip condition renders the fluid velocity zero at the solid surfaces and the larger velocities occur away from the solid surfaces; (2) complex pore-scale geometry and interconnectedness; (3) variability of pore-sizes. Because of these factors, describing in detail the fluid velocity at the pore-scale is a difficult task. However, the flux of fluid over cross-sections containing many pores-spaces, q_i (units: $L^3/L^2/T$) , has been experimentally shown to follow Darcy's law (described in an earlier chapter in the handbook), which, on inferring the attendant empirical constitutive parameter of permeability (k, units L^2), provides a framework for assessing flux in porous media in practical applications. The seepage velocity, v_i, is the specific discharge divided by the porosity (n) of the medium ($v_i = q_i/n$) and has the same units as specific discharge.

14.2.2 Local Dispersion

The seepage velocity v_i does not describe the pore-scale variability of the fluid velocity, i.e., solute undergoes spatially complex advection on account of pore-scale fluid velocity variability around v_i. Additionally, solute also undergo molecular diffusion. The effect of the interaction of molecular diffusion and the complex pore-scale velocity (not described by the seepage velocity v_i) on the solute concentration is called *local dispersion.*

The mass flux (mass being transported per unit cross-sectional area of a porous medium, per unit time), F_i, (units: $M/L^2/T$) is represented by

$$F_i = n\left(v_i c - d_{ij} \frac{\partial c}{\partial x_j} \right) \tag{1}$$

The first term in Equation (1) represents the mass flux associated with the seepage velocity, and the second term represents a Fickian model of mass flux associated with the complex pore-scale variability and molecular diffusion, where d_{ij} (units: L^2/T) is the local dispersion coefficient tensor, that needs to be experimentally determined. The mass per unit volume of solution is denoted by c, the solute concentration. It is implicit in Equation (1) that the concentration c and the seepage velocity v_i are attributes of the fluid that are averaged over many pores contained in a representative elementary volume (REV), and that we make no attempt to describe the nonuniformity of fluid properties over volumes smaller than the REV. The convention of summation over a repeated index is followed in Equation (1) and the rest of the chapter.

For a locally isotropic medium the classic Scheidigger equations (Bear, 1972) give for the dispersion tensor:

$$d_{ij} = \left(\alpha_T \sqrt{v_i v_i} + d^m \right) \delta_{ij} + \frac{\left(\alpha_L - \alpha_T \right)}{\sqrt{v_i v_i}} v_i v_j \tag{2}$$

where $\sqrt{v_i v_i}$ is the seepage velocity magnitude, α_L is the longitudinal local dispersivity, α_T the transverse local dispersivity (units: L), d^m (units: L^2/T) is the effective coefficient of molecular diffusion of the solute in the porous medium, and δ_{ij} is Kronecker's delta ($\delta_{ij} = 0$ for $i \neq j$ and $\delta_{ij} = 1$ for $i \neq j$). The values of α_L, α_T, and d^m need to be determined experimentally. The tortuous flow pathways in the porous medium result in d^m being less than d, the coefficient of molecular diffusion in free solution. The longitudinal dispersivity tends to be less than a centimeter in natural granular material, and the transverse dispersivity even smaller. The reader is referred to the work of Harleman and Rumer (1963a,b), Klotz et al. (1980), and Eidsath et al. (1983) for experimental documentation of the local dispersion coefficients. Bear (1972), Dullien (1979), and Greenkorn (1983) present results of theoretical investigations of the appropriateness of the Fickian assumption and the linear relationship between seepage velocity and dispersion coefficients. For locally anisotropic media, more general models of the local dispersion coefficients have been proposed, but are seldom used, because of the enormous number of empirical constitutive parameters.

Local dispersion in porous media is fundamentally different from the classic Taylor (1953) capillary dispersion phenomena, because the local dispersion coefficients (Equation [2]) do not exhibit an inverse relationship with molecular diffusion at large values of the pore-scale Peclet number $\text{Pe}_{pore} = \sqrt{k v_i v_i}/d$, where k, the permeability of the porous medium, provides a characteristic squared pore length scale. When $\text{Pe}_{pore} \gg 1$, molecular diffusion does not control porous-media local dispersion coefficients. This difference is because in the Taylor (1953) capillary dispersion problem the fluid flow is unidirectional and molecular diffusion singularly causes the solute to sample different velocities, while in multidimensional variable fluid velocities at the pore-scale, even without molecular diffusion, the velocity experienced by solute changes along the tortuous flow paths. However, it would be incorrect to conclude that molecular diffusion coefficients do not have a role to play in transport in porous media, because, the molecular diffusion time-scale over a characteristic pore length scale, which can be estimated as the permeability divided by the molecular diffusion coefficient (k/d), is typically not very large. Taking a nominal value of the molecular diffusion coefficient to be 10^{-5} cm^2/sec, and considering the range of permeability between that of a fine glacial till and gravel (10^{-16} cm^2 to 10^{-3} cm^2) (Freeze and Cherry, 1979), k/d will vary between 10^{-11} and 100 seconds. In many practical applications, the time-scales of interest are much larger than that, so the molecular diffusion induced sampling of different velocities in pore spaces occurs. Therefore, it may be more appropriately stated that when $\text{Pe}_{pore} \gg 1$, transport processes at the REV scale are not limited by molecular diffusion in many practical applications.

14.2.3 Transport Equation

The solute mass conservation statement may be expressed as

$$\frac{\partial(nc)}{\partial t} + \frac{\partial F_i}{\partial x_i} = nr \tag{3}$$

where r is the time rate at which solute mass is being introduced per unit volume of the fluid. Substituting the expression for F_i (Equation [1]) in the solute mass conservation statement gives the advection-dispersion equation (Bear, 1972; Marsily, 1986; Bear and Verruijt, 1987; etc.):

$$\frac{\partial(nc)}{\partial t} + \frac{\partial(nv_i c)}{\partial x_i} - \frac{\partial}{\partial x_i}\left(nd_{ij}\frac{\partial c}{\partial x_j}\right) = nr \tag{4}$$

In this chapter the focus is on nonreactive solutes. The influences of sorption and irreversible transformations are presented in other chapters of this handbook.

The time-variation of concentration at a fixed point is referred to as a concentration breakthrough curve. Also, the spatial variation of the concentration at a given time is called a snapshot curve. The local dispersion coefficients are typically inferred by analyzing concentration breakthrough curves observed

in the laboratory. The usefulness of Equation (4) to understand and model the concentration averaged over the REV stems from our ability to infer the constitutive parameters like the dispersion coefficients, and its ability to predict the spatial–temporal evolution of the concentration under various initial and boundary conditions.

The transport equation (Equation [4]) is capable of predicting the concentrations that are averaged over many pore-spaces. However, it is incorrect to assume that the fluid concentration is completely uniform at scales smaller than the REV, as it is precisely the correlation between sub REV concentration variations and pore-scale velocity variations that gives rise to the local dispersive flux. In the experimental study of Rashidi et al. (1996) microscopic (pore-scale) measurements of concentrations and velocity were made to directly calculate the dispersive flux and infer the dispersion coefficient. Such a directly calculated dispersion coefficient was found to be close to what was inferred from the concentration breakthrough curve. This directly verifies the feasibility of representing the effects of pore-scale variability and molecular diffusion in the local dispersion coefficients. Averaging the concentration over an REV results in a concentration field that varies smoothly in space in a homogeneous porous medium and is amenable to the differential formulation (Equation [4]), which is the basis for most practical investigations of concentrations in the subsurface.

An intuitive understanding about the nature of the transport equation (Equation [4]) can be gained from examining its implications for the moments of a solute body, which are defined as:

Zero*th* moment:
$$\int_\Omega c(\boldsymbol{x},t)\,d\Omega$$

First moment:
$$X_i = \frac{\int_\Omega x_i c(\boldsymbol{x},t)\,d\Omega}{\int_\Omega c(\boldsymbol{x},t)\,d\Omega}$$

Second centered moment:
$$R_{ij}^2 = \frac{\int_\Omega (x_i - X_i)(x_j - X_j)c(\boldsymbol{x},t)\,d\Omega}{\int_\Omega c(\boldsymbol{x},t)\,d\Omega}$$

Ω denotes the spatial domain. In a homogeneous porous medium with a spatially constant seepage velocity v_i, for a pulse input far away from boundaries, in the absence of any source/sink terms ($r = 0$) it follows from Equation (4) that

$$\int_\Omega c(\boldsymbol{x},t)\,d\Omega = \frac{M}{n} \qquad \frac{dX_i}{dt} = v_i \qquad \frac{1}{2}\frac{dR_{ij}^2}{dt} = d_{ij} \tag{5}$$

M is the total mass of solute. In a homogeneous porous media flow field the seepage velocity determines the rate at which the center of mass of a solute body moves, and the local dispersion coefficients control the spatial extent over which it gets distributed. These fundamental features of advection-local dispersion can be simply experimentally illustrated in a homogeneous one-dimensional laboratory column of saturated porous media. The concentration distribution resulting from the introduction of a thin pulse of solute in a unidirectional homogeneous flow ($v_2 = 0$, $v_3 = 0$) is given by

$$c\left(x_1, t\right) = \frac{M}{An\sqrt{2\pi}\,R_{11}(t)}\exp\left[-\frac{\left(x_1 - X_1(t)\right)^2}{2R_{11}^2(t)}\right] \tag{6}$$

A is the cross-sectional area of the porous media column, and M is the total mass of solute introduced, and $dX_1(t)/dt = v_1$, $dR_{11}^2(t)/dt = 2d_{11}$. The bell-shaped spatial variation of concentration described by Equation (6) is called a Gaussian. For homogeneous porous media with spatially uniform flows, as long as the source term r does not involve a nonlinear function of c, classical Fourier and Laplace transformation techniques yield analytical solutions to the transport equation (Equation [4]) for a variety of initial and boundary conditions.

On inferring the seepage velocity v_i and the local dispersion coefficients d_{ij}, the porous continuum scale specification of the transport problem becomes complete. This formulation is practically important because the scales of sampling of groundwater are typically much larger than the pore-scale. As long as the properties of the porous medium change slowly compared to the REV scale (the usual continuum assumption) the porous continuum model is expected to provide a correct description of solute transport. This description of solute transport has been complete for some time and hundreds of measurements of local dispersivities exist. Work continues on the theoretically difficult problem of predicting the values of dispersion coefficients based on models of pore-scale geometry (e.g., Mei, 1992).

In recent times, the practical difficulties in accounting for the effects of heterogeneity of porous media have occupied much attention. In field applications the effects of heterogeneity can be very strong for the contaminant transport problem. The following sections briefly describe the effects of enhanced spreading, the possibility of a disconnect between dilution and spreading, and the complex spatial temporal distribution of the solute concentration.

14.3 Enhanced Spreading due to Heterogeneity

The systematic quantitative assessment of flow and transport rests on our ability to infer its constitutive properties (e.g., hydraulic conductivity, porosity, local dispersion coefficients). The minimum scale entertained in such investigations is the REV, because the porous continuum model requires averaging over it. This framework is quite successful in the homogeneous laboratory setting. However, when the constitutive properties of natural porous media are inferred at distinct points in space, they exhibit intense spatial variability. The complex geological processes responsible for the formation of natural aquifer material do not result in material with spatially uniform constitutive properties. What are the implications of this ubiquitous heterogeneity on solute transport? In a field application, what are the heterogeneity-induced effects one needs to anticipate? These questions are discussed in the following pages.

A clear distinction needs to be made between heterogeneity of porous media constitutive properties and the pore-scale variability. This is because in most practical applications the interest lies in concentrations and flows averaged over many pore volumes, while in contrast, seldom does the averaging occur over scales much larger than the scale of geological variability. As a consequence of geological heterogeneity, the solute concentration distribution in the field is quite complex (as revealed in the Cape Cod bromide tracer test, Leblanc et al., 1991; Garabedian et al., 1991), whereas in laboratory tests with homogeneous porous media and measurements involving averaging over many pore volumes, the concentration varies relatively smoothly in space, consistent with the classical solutions of constant coefficient transport equations.

14.3.1 Field Observations of Macrodispersion

An overwhelming effect of geological heterogeneity is that it causes solute to spread much faster than that implied by the local dispersion coefficients. If a heterogeneous porous medium is subjected to a hydraulic gradient, the ensuing seepage velocity field will vary in space. As different portions of the solute

body will experience different seepage velocities, it will spread out much more rapidly than in a hypothetical homogeneous porous medium.

If x_1 denotes the bulk flow direction in the field, it has been found that $dR_{11}^2(t)/dt$ is much larger than twice the local dispersion coefficient d_{11} (see Equation [5]). The heterogeneity-induced enhanced spreading is also highly anisotropic, with the effect in the transverse horizontal dimension being much smaller than that along the flow direction, and the effects in the vertical dimension being even smaller. The Cape Cod (Leblanc et al., 1991; Garabedian et al., 1991), Columbus, Mississippi (Boggs et al., 1992; Adams and Gelhar, 1992), Borden (Mackay et al., 1986; Freyberg, 1986; Sudicky, 1986) tracer tests are examples of carefully controlled and intensely monitored tracer tests in which the effects of enhanced field-scale spreading have been documented.

From a modeling perspective, if field heterogeneity is not explicitly incorporated in a model in detail (as is almost always the case), the local dispersion coefficients d_{ij} are simply incapable of representing the spreading effects caused by heterogeneity. For example, if a naive attempt was made to employ an average velocity v and predict the spatial second moments of plumes using Equation (4), in conjunction with local dispersion coefficients (and the associated values of the local dispersivities α_L which is typically less than a centimeter and α_T which is even smaller) the spatial second moments of the solute body would be grossly underestimated (especially along the mean flow direction).

The observations of enhanced spreading has created a modeling practice of replacing local dispersion coefficients d_{ij} by often much larger "macrodispersion" coefficients D_{ij}, in the hope that they may somehow represent the effects of heterogeneity on the spreading rates. The macrodispersion coefficients are pragmatically defined as half the rate of change of the spatial second moment in a field setting:

$$\frac{1}{2}\frac{dR_{ij}^2}{dt} = D_{ij} \quad \left(\text{Macrodispersion coefficients}\right) \tag{7}$$

The corresponding macrodispersivities are simply defined as the macrodispersion coefficient divided by the mean velocity (v). Based on this definition the dispersivity in the direction of flow (D_{11}/v) for the Cape Cod tracer test is about one meter, the horizontal transverse dispersivity is a couple of centimeters, and the vertical transverse dispersivity is a couple of millimeters (Garabedian et al., 1991). These macrodispersivities and the corresponding macrodispersion coefficients have been and continue to be the topic of intense theoretical research (Cushman, 1990; Gelhar, 1993) because of their central role in determining the spatial extent over which contamination spreads in time. Gelhar et al. (1992) present a critical review of spreading rates observed in the field and the implied macrodispersivities.

In modeling contaminant transport in the field, it is difficult to explicitly incorporate the effects of small-scale heterogeneity, and thereby model the detailed concentration distributions. The effort to model the REV averaged c(**x**,t) is abandoned in most practical applications and some large-scale spatial average of the concentration, C(**x**,t), is modeled. For the rate of growth of the spatial second moment of C(**x**,t) to be consistent with the field observations, it is often assumed that the large-scale flux is still of the form in Equation (1), replacing the variable seepage velocity by a large-scale average velocity, and replacing the local dispersion coefficients d_{ij} by the macrodispersion coefficients D_{ij}. In this modeling practice, the Fickian model of transport is assumed to apply to the averaged description of concentration: C(**x**,t). This assumption can be valid only when the scale of spatial variations of C(**x**,t) is much larger than the scales of variations of the unresolved small-scale heterogeneity, whose influence is sought to be parameterized in the macrodispersion coefficients D_{ij}. If that assumption is not met, nonlocal representations of the flux of C(**x**,t) may be required (Deng et al., 1993; Cushman et al., 1995).

Flow variability at scales comparable to or larger than the scales characteristic to the spatial variations of C(**x**,t) act to displace and possibly cause a large-scale distortion of C(**x**,t), and do not always enhance the rate of growth of its spatial second moment like small-scale flow variability (Rajaram and Gelhar, 1993). Therefore, in modeling transport at a given site it is desirable to understand and incorporate the large-scale geological structures explicitly into the model (Anderson, 1989; Webb and Anderson, 1996).

The analysis of the Columbus tracer test (Adams and Gelhar, 1992) shows how large-scale groundwater flow structure can profoundly influence the shape of a tracer plume, and that an explicit recognition of the large-scale structure results in a significantly different assessment of the macrodispersion coefficient relevant to quantify the dispersive flux.

The mathematical description of the heterogeneity of aquifer properties is of central interest in predicting what happens in the field. In doing so, the geological structural controls and the large-scale geological features should be incorporated in any practical model. However, the small-scale geological variability is impossible to measure in exact detail. At best, one can characterize its statistical properties (intensity of fluctuations, spatial persistence) in a region. This reality of small-scale variability, its well-documented dramatic impact on plume spreading rates, and our ignorance about its detailed nature have led to its representation by a multidimensional stochastic process (random field), and an examination of the relationship between solute concentrations and parameters describing the intensity and spatial persistence of variability. A comprehensive review of this rapidly evolving topic is not possible in the space of this chapter (for that the reader is referred to Cushman, 1990 and Gelhar, 1993). What follows is an example of representing hydraulic properties as a realization of a multidimensional stochastic process and an illustration of its impacts on transport. The objective is to provide illustrations of the types of effects that can be experienced in the field, and therefore should be acknowledged in field-scale assessments.

14.4 Stochastic Model

The overwhelming effect of heterogeneity in field concentration data has lead to attempts to systematically explore the relationship between heterogeneity and transport phenomena. In the process of explaining transport phenomena in aquifers, a better understanding and more realistic descriptions of flow in heterogeneous porous media have been developed, as its variability, along with the local dispersion coefficients, control the concentrations observed in the field.

In dealing with a hydrogeologic constitutive parameter that has been observed to vary in space, acknowledging our limited ability to map its variability in detail, that constitutive parameter can be represented as a random field. Although practical data limitations make it challenging to assess the parameters for describing even the first two statistical moments of such random field representations, the random field description provides a way to systematically acknowledge the heterogeneity of the aquifer and our ignorance about all its details.

14.4.1 Random Field Representation of Heterogeneity

To illustrate the effects of heterogeneity we present the evolution of a pulse input in a statistically stationary random hydraulic conductivity (K) field. In the example, $\ln K$ is taken to be a stationary random field, and the variations of $\ln K$ around its ensemble mean $\ln K_G$, $f(\mathbf{x}) = \ln K(\mathbf{x}) - \ln K_G$, are assumed to be spatially stationary and characterized by a spatial covariance function $R_{ff}(\xi) \equiv E[f(\mathbf{x})f(\mathbf{x} \pm \xi)]$, where $E[.]$ denotes the statistical expectation operator. In any practical application the form and parameters of the covariance function require collecting data at that site. The assumption of stationarity is one of convenience. If data exhibit discernible trends in the mean, variance, and spatial persistence parameters, they need to be accounted for (Rajaram and McLaughlin, 1990).

For illustration purposes, we present an example in which

$$R_{ff}\left(\xi\right) = \sigma_f^2 \prod_{i=1}^{m} \exp\left[-\left(\xi_i / \lambda_i\right)^2\right]$$

where m is the number of spatial dimensions, which is taken to be 2. The variance of $\ln K$ is denoted by σ_f^2. From the available data, an empirical covariance function of $\ln K$ can be constructed, and the distance

in any principal direction at which the covariance falls to $\sigma_f^2 e^{-1}$ is the "correlation scale λ_i" in that direction. Another associated scale is the "integral scale I_i," which is such that $\sigma_f^2 I_i$ is the area underneath the covariance function when plotted versus the ith spatial separation axis (with the separation distance in the other directions being 0). In the two-dimensional example shown here $\sigma_f^2 = 1$, $I_1 = 100$ cm and $I_2 = 10$ cm, $K_G = 100$ cm/d and the overall rate of drop of hydraulic head in the x_1 direction is taken to be 0.01, and the porosity n is taken to be a constant value of 0.3.

The simulated hydraulic conductivity field shown in Figure 14.1 is one realization of a multidimensional stochastic process (random field). If measurements are used only to construct a covariance model and simulate such realizations they are unconditional realizations insofar as they are not required to attain the measured values at the points of measurement. If the simulated realization attains the measured values at the points of measurement, it is a conditioned realization. Both these types of realizations of hydraulic conductivity fields can be used to understand how transport occurs in a heterogeneous aquifer. It is inappropriate to use ordinary kriging to simulate a hydraulic conductivity field to perform detailed numerical simulations of flow and transport (employing local dispersion coefficients) because the simulated hydraulic conductivity field will usually lack the microstructure that influences transport, on account of being a conditional mean (with the assumption of Gaussianity). Eggleston et al. (1996) provide an example of these different types of simulations of hydraulic conductivity for the Cape Cod tracer test site.
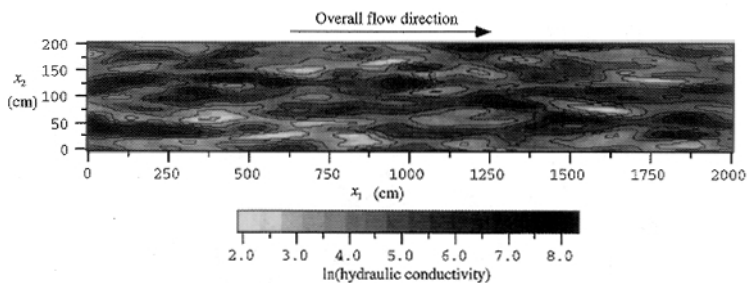
FIGURE 14.1 Simulated random hydraulic conductivity field using the spectral algorithm of Dykaar and Kitanidis (1992).

14.4.2 Flow and Transport Simulation

The flow problem is solved with no flux boundary conditions at $x_2 = 0$ and 200 cm, employing a standard successive over-relaxation scheme (Ames, 1992) that converges rapidly when the spatial discretization adequately resolves the spatial derivatives of the lnK field ($\Delta x_i \leq 2/\max(\partial \ln K/\partial x_i)$ ensures that the equations resulting from discretization have a unique solution). A strip of solute is introduced, centered at 100 cm, with the maximum concentration being 1 unit. A transport simulation is done for $\alpha_L = 0.5$ cm, and $\alpha_T = 0.1$ cm. In Figure 14.2 are shown shaded and contoured plots of the concentration field, employing a standard centered, explicit, second-order finite difference scheme (Ames, 1992).

Figure 14.3 shows the evolution of the x_1 spatial second moment of the solute body in the heterogeneous domain, and what would occur without heterogeneity. It is not surprising that the heterogeneity resulted in faster rates of growth of spatial moments compared to the homogeneous case. In practical applications, this enhanced rate of growth is typically accounted for by simply using a suitably larger "macrodispersion" coefficient, instead of the local dispersion coefficient that works well in laboratory-scale homogeneous specimens. However, there is a major qualitative difference between the observations in the laboratory and that in the heterogeneous field (which the simulation illustrates). In the homogeneous laboratory experiments, the concentration breakthroughs are smooth; however, in the field the concentration distribution is highly irregular. Therefore all that the macrodispersion coefficient-based methodology can hope to accomplish is to describe the bulk dimensions (i.e., spatial second moments) of the concentration distributions, and not their details.
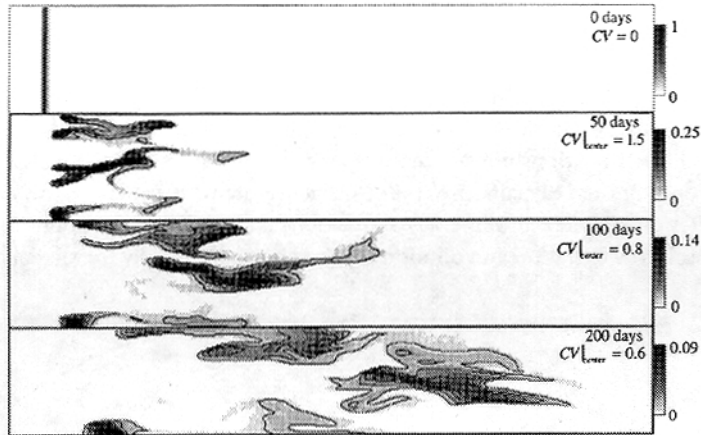
FIGURE 14.2 Evolution of solute concentration in the saturated random hydraulic conductivity field shown in Figure 14.1. The heterogeneity causes enhanced spreading and a complex spatial distribution of concentration. The average of the concentration in the x_2 direction was found, and the mean squared deviation of the concentration around that average value was also found. The resultant coefficient of variation (CV = root mean square deviation divided by the mean) provides a nondimensional measure of concentration fluctuations. Initially the CV rises, and eventually it undergoes a slow decrease.
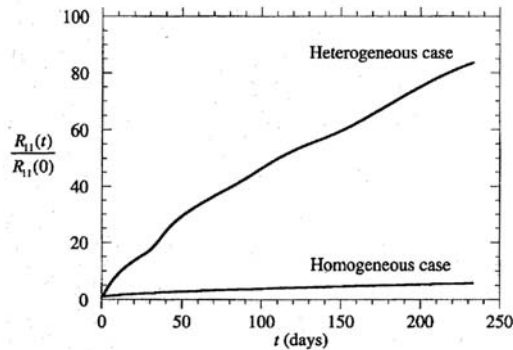
FIGURE 14.3 Enhanced spreading due to heterogeneity.

It is difficult to argue that the detailed variations of the concentration field are of no important consequence, because sampling and exposure volumes can be smaller than the scale of heterogeneity of aquifers, and they seldom affect the averaging implicit in the macrodispersion coefficient-based description. Therefore, it would be quite attractive to be able to effect a coarse description $C(\mathbf{x},t)$, possibly employing macrodispersion coefficients, and also be able to put an error bar around that coarse description. That is the motivation for developing an understanding of concentration fluctuations and dilution in addition to the spatial moments of the concentration distribution.

14.5 Spreading versus Dilution

For a pulse input in a homogeneous porous media with a spatially constant seepage velocity, the peak concentration at any time, $c_p(t)$, is inversely proportional to the spatial extent of the plume (Equation [6]). In a homogeneous one-dimensional laboratory column of saturated porous media, that feature can be used to infer the dispersion coefficients. If a thin pulse of solute is introduced at one end of a homogeneous saturated column of porous media, measuring the breakthrough concentration at the downstream end reveals that the peak concentration is inversely proportional to R_{11}, the radius of gyration

(i.e., square root of the centered spatial second moment) of the solute body along the direction of flow, as depicted in Figure 14.4a. Similarly, in a multidimensional homogeneous porous media with a uniform hydraulic gradient, the peak concentration is inversely proportional to the product of the radii of gyration in the different principal directions $R_{11}R_{22}R_{33}$. These features are basic properties of a constant coefficient advection-diffusion equation. In dealing with contaminant transport in the field, the natural question arises whether this inverse relationship between the peak concentration and the spatial second moments is followed. Is the drop of the peak concentration commensurate with growth of the spatial moments of the solute body? Results of detailed numerical simulations and an intensely monitored field tracer test indicate that the peak concentration and dilution may not be quantified by the spatial second moments.
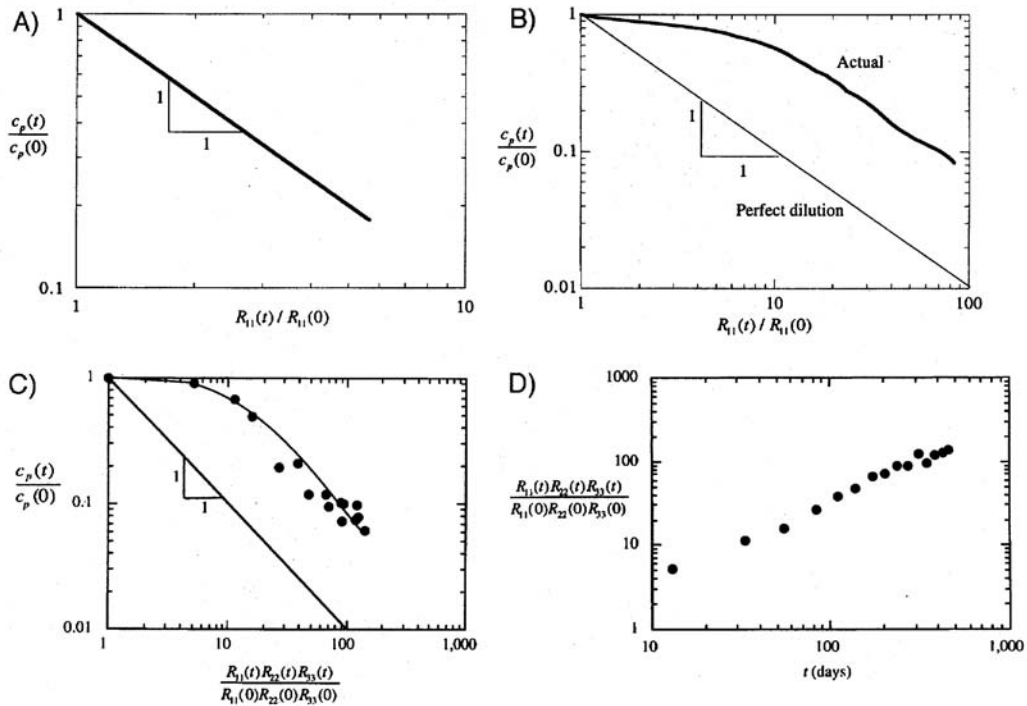
FIGURE 14.4 Spreading versus dilution. (a) Peak concentration in homogeneous hydraulic conductivity field. (b) Peak concentration in heterogeneous hydraulic conductivity field. (c) Peak concentration for the Cape Code bromide tracer test. (d) Spreading for the Cape Cod bromide tracer test. (Data for figures c and d from LeBlanc, D. R., Garabedian, S. P., Hess, K. M., Gelhar, L. W., Quadri, R. D., Stollenwerk, K. G., and Wood, W. W. 1991. Large-scale natural gradient tracer test in sand and gravel, Cape Cod, Massachusetts: 1. Experimental design and observed tracer movement. *Water Resour. Res.* 27(5), 895-910; Garabedian, S. P., LeBlanc, D. R., Gelhar, L. W., and Celia, M. A. 1991. Large-scale natural gradient tracer test in sand and gravel, Cape Cod, Massachusetts: 2. Analysis of spatial moments for a nonreactive tracer. *Water Resour. Res.* 27(5), 911-924.)

14.5.1 Disconnect Between Spreading and Dilution

The transport simulation in the random flow field (Figure 14.2) is such that the x_1 spatial second moment of the solute mass is growing with time, and at a larger rate than what would be effected by local dispersion alone in a homogeneous porous media. For that setting, the question posed here is: Does an n-fold increase in R_{11} in the heterogeneous media result in an n-fold decrease of the peak concentration? Figure 14.4b presents the peak concentration as a function of R_{11}. At the end of the simulation time about a hundred-fold increase of R_{11} from its initial value has occurred, but is accompanied by only a tenfold decrease in the peak concentration (Figure 14.4b). That plot clearly shows that at early times dilution

and spreading are quite disconnected: the first tenfold increase in R_{11} is only accompanied by a 1.67-fold decrease in the peak concentration (to 60% of its initial value). At early time the rapid growth of R_{11} is not accompanied by a commensurate decrease of the peak concentration. For the hypothetical pure advection case, of course, the peak concentration is invariant (in a divergence free flow field). Figure 14.4b also indicates that at large time an inverse dependence of the peak concentration with R_{11} is approached, i.e., after the early time disconnect between dilution and spreading, at large time, dilution and spreading become connected again.

What is the time-scale over which dilution and spreading remain disconnected? In answering this question it needs to be recognized that no dilution can be effected without local dispersion (Kitanidis, 1994), notwithstanding the heterogeneity-induced growth of the spatial second moments. Therefore local dispersion will play an important role in determining that time-scale. A time-scale over which the multidimensional local dispersion process occurs over a multi-dimensional (m-dimensional) spatial zone with dimensions equal to the correlation scales λ_i of the heterogeneity can be formulated as

$$t_d = \left[\sum_{i=1}^{m} \frac{d_{ii}}{\lambda_i^2} \right]^{-1} \tag{8}$$

Evaluating the local dispersion coefficients corresponding to the mean x_1 velocity (v) of the simulation, ($d_{11} = v\alpha_L$, $d_{22} = v\alpha_T$), this time-scale is 250 days, which is roughly the total simulation time. Therefore it is not surprising that for most of the simulation, dilution and spreading were disconnected (Figure 14.4b) in the heterogeneous hydraulic conductivity field.

14.5.2 Cape Cod Bromide Tracer Data

The disconnect between spreading and dilution is also exhibited by the Cape Cod bromide tracer data shown in Figure 14.4c (adapted from Figure 14.8, Kapoor and Gelhar, 1994, with data from Leblanc et al., 1991; Garabedian et al., 1991), which is qualitatively similar to the result of numerical simulation shown in Figure 14.4b. Over about 500 days a 200-fold increase in $R_{11}R_{22}R_{33}$ is accompanied by a 20-fold decrease in the peak concentration (Figures 14.4c and 14.4d). At early time the disconnect between the spatial second moment and dilution is quite pronounced. A tenfold increase of $R_{11}R_{22}R_{33}$ is accompanied by a decrease in the peak concentration by a factor of 1.4 (to 70% of its initial value). At large time it appears that dilution and spreading become more connected.

What is the time-scale, over which dilution and spreading are disconnected at the Cape Code site? The local dispersion time-scale characteristic to the heterogeneity correlation scale (defined in Equation [8]) t_d, is equal to 165 days, using the following parameters on the heterogeneity scale and local dispersion (Garabedian et al., 1991): $\lambda_1 = \lambda_2 = 260$ cm, $\lambda_3 = 19$ cm, $d_{11} = v\alpha_L$, $d_{22} = d_{33} = v\alpha_T$, $v = 42$ cm/d, $\alpha_L = .5$ cm, $\alpha_T = .05$ cm. Note that 165 days corresponds to when $R_{11}R_{22}R_{33}$ has increased by a factor of 50 from its initial value (Figure 14.4d), and that is when dilution and $R_{11}R_{22}R_{33}$ appear to get significantly connected (Figure 14.4c).

The purpose of the preceding discussion is to caution against the use of large *effective*, or *macro* dispersion coefficients to assess dilution (peak concentration). At best, these large dispersion coefficients can describe the spread of the solute body (R_{ij}). The cardinal role played by local dispersion in effecting dilution, and the smallness of the local coefficients, and the preceding numerical simulation and field tracer test data indicate that in real applications dilution may lag considerably behind spreading of contaminants.

Much theoretical work has been done on predicting the spatial second moments at early and large times. In contrast the description of the early time disconnect and large time connection between dilution and spreading is relatively recent (Kapoor and Gelhar, 1994) and needs to be explored for a broad variety of heterogeneity descriptions to develop practical guidelines on assessing dilution.

14.6 Concentration Fluctuations

14.6.1 Irregularity of Concentration Distributions

The example simulation (Figure 14.2) illustrates that although the spatial second moment of the solute body grows at a rate much faster than what would be caused by local dispersion alone in a hypothetical homogeneous aquifer (Figure 14.3), the actual concentrations do not always decrease at a rate that is commensurate with the rate of growth of the overall spatial dimension of the solute body (Figure 14.4b). While most modeling studies focus on describing some kind of spatial mean concentration employing macrodispersion coefficients (or effective dispersion coefficients), the peak concentration and the general variability of concentrations around the mean concentration may also be of interest, especially if the averaging implicit in defining the mean concentration is larger than the scale over which the contaminant concentration is of possible concern. It would be quite attractive if a simple method could be found to supplement the mean concentration by a concentration standard deviation. Before such a method can be developed, the general nature of concentration fluctuations needs to be recognized.

14.6.2 Rise and Fall of the Concentration Coefficient of Variation

Do plumes resulting from a pulse input into a heterogeneous aquifer become increasingly irregular with time? The answer to that question, at early times, is yes, and at large times the concentration irregularity actually decreases with time (Kapoor and Gelhar, 1994). This can be seen in the simulation results presented in Figure 14.2. The irregularity of the concentration is examined by computing averages in the x_2 direction, $\langle c \rangle$, and also the concentration variance, defined as

$$\left\langle \left(c - \langle c \rangle \right)^2 \right\rangle$$

The concentration coefficient of variation, CV, defined as

$$CV = \sqrt{\left\langle \left(c - \langle c \rangle \right)^2 \right\rangle} / \langle c \rangle$$

is an important nondimensional measure of the irregularity of concentration. When CV $\ll$ 1 then the mean concentration is an adequate quantifier of actual concentration values. The CV at the center of mass of the solute body is shown in Figure 14.2. It is shown in the figure that the CV increases from its initial value of zero and later slowly decreases with time. For pure advection the CV would keep unrealistically increasing unboundedly with time. The decrease of the CV with time is singularly brought about by local dispersion (Kapoor and Gelhar, 1994). This rise and fall of the concentration CV is consistent with the early time disconnect between dilution and the reconnection of dilution and spreading, at large time, due to local dispersion.

Investigations of concentration fluctuations are much fewer in number than the mean concentration and the spreading rates. However, field tracer tests clearly show highly irregular concentration distributions and a disconnect between spreading and dilution rates. Further analysis of concentration fluctuations need to be conducted, and methods to predict the concentration CV and dilution need to be developed and tested.

For Further Information

Bear and Verruijt (1987) cover groundwater flow and contaminant transport from a modeling perspective. Freeze and Cherry (1979) provide a textbook coverage of groundwater contamination. Gelhar (1993)

provides a systematic introduction to the topic of analyzing flow and contaminant transport in heterogeneous porous media. Gelhar et al. (1992) provide a critical review of data on field-scale dispersion coefficients.

References

Adams, E. E. and Gelhar, L. W. 1992. Field study of dispersion in a heterogeneous aquifer 2. Spatial moment analysis. *Water Resour. Res.* 28(12), 3293-3307.

Ames, W. F. 1992. *Numerical Methods for Partial Differential Equations*. Academic Press, San Diego, CA.

Anderson, M. P. 1989. Hydrogeologic facies models to delineate large-scale spatial trends in glacial and glaciofluvial sediments. *Geol. Soc. Am. Bull.* 101, 501-511.

Bear, J. 1972. *Dynamics of Fluids in Porous Media*. American Elsevier, New York.

Bear, J. and Verruijt, A. 1987. *Modeling Groundwater Flow and Pollution*. Reidel, Boston.

Boggs, J. M., Young, S. C., Beard, L. M., Gelhar, L. W., Rehfeldt, K. R., and Adams, E. E. 1992. Field dispersion in a heterogeneous aquifer. *Water Resour. Res.* 28(12), 3281-3291.

Cushman, J. H. (Ed.). 1990. *Dynamics of Fluids in Hierarchical Porous Media*. Academic Press, San Diego, CA.

Cushman, J. H., Hu, B. X., and Deng, F.-W. 1995. Nonlocal transport with physical and chemical heterogeneity: localization errors. *Water Resour. Res.* 31(9), 2219-2237.

Deng, F.-W., Cushman, J. H., and Delleur, J. W. 1993. A fast Fourier transform stochastic analysis of the contaminant transport problem. *Water Resour. Res.* 29(9), 3241-3247.

Dullien, F. A. L. 1979. *Porous Media, Fluid Transport and Pore Structure*. Academic Press, New York.

Dykaar, B. B. and Kitanidis, P. K. 1992. Determination of the effective hydraulic conductivity for heterogeneous porous media using a numerical spectral approach, 1, Method. *Water Resour. Res.* 28(4), 1155-1166.

Eidsath, A., Carbonell, R. G., Whitaker, S., and Herrmann, L. R. 1983. Dispersion in pulsed systems, III. Comparison between theory and experiments in packed beds. *Chemical Engineering Science.* 38, 1803-1816.

Eggleston, J. R., Rojstaczer, S. A., and Peirce, J. J. 1996. Identification of hydraulic conductivity structure in sand and gravel aquifers: Cape Cod data set. *Water Resour. Res.* 32(5), 1209-1222.

Freeze, R. A. and Cherry, J. A. 1979. *Groundwater.* Prentice-Hall, Englewood Cliffs, NJ.

Freyberg, D. L. 1986. A natural gradient experiment on solute transport in a sand aquifer, 2, Spatial moments and the advection dispersion of nonreactive tracers. *Water Resour. Res.* 22(13), 2031-2046.

Garabedian, S. P., LeBlanc, D. R., Gelhar, L. W., and Celia, M. A. 1991. Large-scale natural gradient tracer test in sand and gravel, Cape Cod, Massachusetts: 2. Analysis of spatial moments for a nonreactive tracer. *Water Resour. Res.* 27(5), 911-924.

Gelhar, L. W., Welty, C., and Rehfeldt, K. R. 1992. A critical review of data on field-scale dispersion in aquifers. *Water Resour. Res.* 28(7), 1955-1974.

Gelhar, L. W. 1993. *Stochastic Subsurface Hydrology*. Prentice Hall, Englewood Cliffs, NJ.

Greenkorn, R. A. 1983. *Flow Phenomena in Porous Media*. Marcel Dekker, New York.

Harleman, D. R. F. and Rumer, R. R. 1963a. Longitudinal and lateral dispersion in an isotropic porous medium. *J. Fluid Mech.* 16, 385-394.

Harleman, D. R. F., Melhorn, P. F., and Rumer, R. R. 1963b. Dispersion-permeability correlation in porous media. *J. Hydraul. Div. ASCE.* 67, 67-85.

Kapoor, V. and Gelhar, L. W. 1994. Transport in three-dimensionally heterogeneous aquifers: 2. Predictions and observations of concentration fluctuations. *Water Resour. Res.* 30(7), 1789-1801.

Kitanidis, P. K. 1994. The concept of the dilution index. *Water Resour. Res.* 30(7), 2011-2026.

Klotz, D., Seiler, K.-P., Moser, H., and Neumaier, F. 1980. Dispersivity and velocity relationships from laboratory and field experiments. *J. Hydrol.* 45, 169-184.

LeBlanc, D. R., Garabedian, S. P., Hess, K. M., Gelhar, L. W., Quadri, R. D., Stollenwerk, K. G., and Wood, W. W. 1991. Large-scale natural gradient tracer test in sand and gravel, Cape Cod, Massachusetts: 1. Experimental design and observed tracer movement. *Water Resour. Res.* 27(5), 895-910.

Mackay, D. M., Freyberg, D. L., Roberts, P. V., and Cherry, J. A. 1986. A natural gradient experiment on solute transport in a sand aquifer, 1, Approach and overview of plume migration. *Water Resour. Res.* 22(13), 2017-2029.

Marsily, G., de. 1986. *Quantitative Hydrogeology, Groundwater Hydrology for Engineers.* Academic Press, Orlando, FL.

Mei, C. C. 1992. Method of homogenization applied to dispersion in porous media. *Transport in Porous Media.* 9, 261-274.

Rajaram, H. and McLaughlin, D. 1990. Identification of large-scale spatial trends in hydrologic data. *Water Resour. Res.* 26(10), 2411-2423.

Rajaram, H. and Gelhar, L. W. 1993. Plume-scale dependent dispersion in heterogeneous aquifers, 2, Eulerian analysis and three-dimensional aquifers, *Water Resour. Res.* 29(9), 3261-3276.

Rashidi, M., Peurrung, L., Tompson, A. F. B., and Kulp, T. J. 1996. Experimental analysis of pore-scale flow and transport in porous media. *Advances in Water Resources.* 19(3), 163-180.

Sudicky, E. A. 1986. A natural gradient experiment on solute transport in a sand aquifer: Spatial variability of hydraulic conductivity and its role in the dispersion process. *Water Resour. Res.* 22(13), 2069-2082.

Taylor, G. I. 1953. Dispersion of soluble matter in solvent flowing slowly through a tube. *Proc. Royal Soc. London.* 219(Series A), 186-203.

Webb, E. K. and Anderson, M. P. 1996. Simulation of preferential flow in three-dimensional heterogeneous conductivity fields with realistic internal architecture. *Water Resour. Res.* 32(3), 533-545.

Glossary

Coefficient of Variation   The standard deviation divided by the mean. The concentration coefficient of variation initially increases with time. Local dispersion can halt the growth of the concentration coefficient of variation, and also cause it to decrease with time at large time.

Concentration Fluctuations   Variations of concentration around its smoothly varying mean.

Dilution   Attenuation of concentrations of a solute undergoing transport. Dilution is distinct from spreading. Spreading can occur without any dilution.

Gaussian   Bell-shaped spatial distribution of solute concentration, resulting from the introduction of a pulse of solute at a point in a uniform flow field.

Geological Heterogeneity   The textural variations exhibited in natural porous media, which give rise to spatial variability of their constitutive properties.

Local Dispersion   The effect of interaction of molecular diffusion and the pore-scale velocity variations on the solute concentration. The mass flux associated with local dispersion in the presence of concentration gradients causes the solute concentration to decrease with time.

Macrodispersion   The enhanced rate of spreading of solute introduced in a heterogeneous aquifer, relative to what would occur in a hypothetical homogeneous porous medium.

Second Spatial Moment   A measure of the spread of solute mass around its center of mass.

Seepage Velocity   The fluid velocity averaged over many pores contained in a representative elementary volume (REV) of the porous continuum.

Spreading   Growth of the spatial second moments of a concentration distribution.

15

Reactive Contaminant Transport in the Saturated Zone

John H. Cushman
Purdue University

T. R. Ginn
University of California–Davis

15.1 Introduction

We extend the discussion of the previous chapter to a setting wherein the dissolved chemical reacts with its environment. The catalogue of natural geochemical reactions and transformations associated with man-made contaminants that have entered the saturated zone is essentially limitless, but for the most part these reactions can be categorized according to speed, reversibility, and phase association. A relevant subset of the classification scheme of Rubin (1983) includes the four cases: (1) reactions that are fast (equilibrium) and reversible involving aqueous-aqueous or (2) aqueous-solid phase interactions; and (3) reactions that are not fast (nonequilibrium) involving aqueous-aqueous or (4) aqueous-solid phase interactions. The terms *fast* and *not fast* are defined relative to the other transformation processes involved, including other reactions and convective and dispersive transport on the observation time scale. Aqueous-aqueous (categories 1 and 3) are termed *homogeneous* and aqueous-nonaqueous (2 and 4) are termed

heterogeneous (regarding phase; not to be confused with spatial variability). This scheme is useful because there is a distinct mathematical approach for each of these categories. The chemical mechanisms driving general reactions and the mass balance approaches for formulating representative transformation models are beyond the present scope, and are summarized at various levels in, for instance, Rubin (1983), Fetter (1993), Marsily (1986), Sposito (1994), and Lichtner et al. (1996). Here we will be restricted to some of the simpler (but representative) reaction mechanisms in categories 2 (sorption) and 4 (biodegradation), and focus on scaling issues associated with transport in heterogeneous porous media. We will begin with a brief summary of reaction transformation formulations and then describe concepts of heterogeneity, scale, and upscaling. The remainder of the chapter describes current methods for scaling plume behavior involving linear reactions (Eulerian approaches), and for scaling contaminant arrival behavior involving nonlinear reactions (Lagrangian approaches).

We start with a simple illustrative example of an equilibrium heterogeneous reaction, involving the single electrolyte NaCl. Dissolved chloride is a simple anion and generally moves in accordance with the local pore-water velocity, much as the water molecules themselves do (with some exception due to the repulsive electrostatic potential between the anion and the mineral surfaces in, for example, quartzitic or kaolinitic materials). Thus chloride is often reasonably treated as an inert or passive tracer using methods described in the previous chapter. Sodium, however, tends to be both transported and exchanged with other cations on the negatively charged mineral surfaces. That is, aqueous sodium reacts with the solid phase by cation exchange, one of a number of mechanisms by which aqueous species may associate with the mineral surfaces. Other mechanisms include precipitation (e.g., Novak and Sevougian, 1993), surface complexation (e.g., Dzombak and Morel, 1990), hydrophobic sorption of organic species to organic solids (Karickhoff, 1983), colloid filtration (Tien et al., 1979), and intergranular diffusion (Wood et al., 1990).

Reaction transformations are formally captured via augmenting the basic solute transport rate equation (Equation [1] of Chapter 14) with a generalized mass loss rate, $nr = [\partial C/\partial t]_{\text{transformation}}$, here written as mass transfer to the solid phase, with constant porosity n:

$$\frac{\partial C}{\partial t} + \frac{\partial (V_i C)}{\partial x_i} - \frac{\partial}{\partial x_i}\left(d_{ij}\frac{\partial C}{\partial x_j}\right) = \left(\frac{\partial C}{\partial t}\right)_{\text{transformation}} \equiv -\frac{\rho_b}{n}\frac{\partial S}{\partial t} \tag{1a}$$

where *C* is the aqueous concentration [M/L³],

   S is the sorbed concentration per unit mass of solid [M/M],

   ρ_b is the bulk density of the aquifer material [M/L³],

   n is the porosity.

A majority of sorption and other heterogeneous transformations may be represented by specifying this mass transformation rate as a function of the departure of the system from an equilibrium state. The first-order form is:

$$\frac{\partial S}{\partial t} = K_r\left(K_d C - S\right) \tag{1b}$$

where K_r is the desorption rate [1/T] and K_d is the equilibrium distribution coefficient [*L/M*]. Equation (1b) specifies a linear nonequilibrium ("not fast") reversible mass transformation rate that approximates numerous sorption processes (of category 4) in natural media. Alternatively, one may express the right-hand side of Equation (1b) as [$K_f C - K_r S$], where $K_f = K_r K_d$ is the forward sorption rate [1/T]. Reactive solutes undergoing transport according to Equation (1) exhibit attenuation. If K_r is very large so that Equation (2) reaches steady-state much more quickly than local conditions change [due to other mass fluxes in Equation [1a]), then the *local equilibrium assumption* is invoked and the chemical partitioning is taken to occur instantaneously. We may then use the steady-state solution to Equation (1b), $S = K_d C$,

as an algebraic relation for S in terms of C in the derivative on the right-hand side of Equation (1a) (to express the dependence of S on C). Differentiating and grouping terms gives from Equation (1a)

$$\frac{\partial C}{\partial t} + \frac{1}{R}\frac{\partial\left(V_i C\right)}{\partial x_i} - \frac{1}{R}\frac{\partial}{\partial x_i}\left(d_{ij}\frac{\partial C}{\partial x_j}\right) = 0 \tag{2}$$

where $R = 1 + \rho_b K_d/n$ is known as the retardation factor. When transport is governed by an equation involving equilibrium reactions such as Equation (2), the solute(s) are retarded (as opposed to attenuated), as is clear upon recognizing that R reduces the transport parameters V_i and d_{ij}.

Equations (1) and (2) illustrate the simplest case of heterogeneous mass transformations of the nonequilibrium and equilibrium types. This example illustrates the different mathematical approaches for the two cases. For nonequilibrium reactions we specify a ("kinetic") function for

$$\frac{\partial S}{\partial t} \equiv r\left(S, C\right)$$

and for equilibrium we specify the algebraic relation (termed "isotherm") of the form $S = r(C)$, so that

$$\frac{\partial S}{\partial t} \equiv r'\left(C\right)\frac{\partial C}{\partial t}$$

and $r(C)$ becomes involved in the transport operator. More complex forms of r arise when the mass transformations are nonlinear in C and/or in S, when multiple processes are in effect, when the non-aqueous phase (for S) undergoes transformations, and when multiple interacting solutes are involved, to note only a few cases.

In the case of multiply interacting solutes and solid-phase species, an evolution equation is written for each component, and these equations are coupled through the (multiple) reaction terms appearing. Such systems present special issues requiring attention in their solution depending on the character of the reactions (cf. review in Yeh and Tripathi, 1989). When all interactions are at equilibrium the transport equations are coupled with a set of algebraic equations relating mobile to immobile species, that complicates solution via integration schemes designed for the transport operator alone. Recent approaches use iterative solutions between the algebraic and differential equations (Yeh and Tripathi, 1989) or a decoupling of the system (Rubin, 1990). In the other limiting case where all the reactions have a finite rate, then a sink term appears for each species in the evolution equation and the solution can proceed conventionally through operator-splitting (e.g., Chiang et al., 1990; Valocchi and Malmstead, 1992) or explicit integration schemes (e.g., Widdowson et al., 1988). Finally, a mixture of both nonequilibrium and equilibrium transformations results in a *differential-algebraic equation* (DAE) system which may be solved, under particular conditions regarding the index of the system (Gear and Petzold, 1984; Gear, 1988), through specialized iterative or noniterative decoupling approaches (e.g., Steefel and MacQuarrie, 1996).

We now turn our attention to the available approaches for dealing with reactions in the presence of physical and/or chemical heterogeneities in natural media. The goal of the scaling approaches is to derive maximum information about the solute fate and transport when data on the spatial variability of the natural media is limited. We describe two such approaches, a Eulerian stochastic-analytic method that is useful for describing expected plume behavior under linear reactions, given statistical (correlation) properties of the spatially variable aquifer flow and reaction characteristics; and a Lagrangian-streamtube approach useful for predicting solute arrivals at observation locations (as opposed to plume evolution) when the reactions involve nonlinear transformations. Beyond the current scope is a third approach that is quite useful when available aquifer data includes conditioning data on flow and reaction property

values on significant scales, such as identification and incorporation of hydrofacies geometry (Bierkens and Weerts, 1994; Scheibe and Freyberg 1995).

15.2 Physical, Chemical, and Microbiological Heterogeneity

It is clear that transport models for multiple species of reactive and degradable chemicals in a homogeneous environment must be very complicated. Couple this already complicated transport phenomena with heterogeneity, and the resultant problem increases in complexity by an order of magnitude. To date, only the simplest of such problems have been solved analytically, and even Monte Carlo methods have been applied to only moderately complex systems.

In the previous chapter physical heterogeneity was discussed in the context of a random hydraulic conductivity field. There are two other commonly overlooked forms of physical heterogeneity: a random porosity field and a random "local" dispersivity field. Both are commonly considered of secondary importance to flow and transport, though recent Monte Carlo simulations (Hassan et al., 1997) suggest correlation between conductivity and porosity may play a significant role in transport. As in Chapter 14, however, we will equate physical heterogeneity with a random conductivity field.

Superimposed on the random conductivity field is a random geochemical field. Few experiments have been conducted on this field, so little is currently known about its correlation structure, including its correlation with conductivity. Random distributions of mineral type, diagenesis, weathering, and random organic matter decay and deposition give rise to this natural heterogeneity. Superimposed over both the physically and chemically heterogeneous fields is a dynamic random distribution of microbial populations. The major difference between physical/chemical heterogeneity and microbial heterogeneity is the dynamic nature of the latter. Microbes are transported and adsorbed much as are chemicals and other colloidal particles, though the adsorptive mechanisms are different.

15.3 Scale of Observation and Constitutive Concepts

Engineered porous media are generally designed to be homogeneous when viewed over the scale on which they are to be used in a given technology. Natural porous formations, on the other hand, are almost always inhomogeneous on any relevant scale of observation. As mentioned in the previous section these natural formations are not only inhomogeneous physically, but chemically and biologically as well. However, the degree and type of these heterogeneities depends to a large extent on the scale of observation. It is for this reason that "effective" parameters such as so-called macrodispersivity have little meaning in natural systems.

Systems which are not homogeneous on any relevant scale of observation are often said to have a "continuous hierarchy" of scales, or to be of "evolving heterogeneity." Special tools are often required to deal with these types of media, because any rational upscaling technique will lead to a nonlocal constitutive theory. Continuous hierarchy is associated with extensive "self-organization" which is evident in long-range correlations among both dynamic and static properties. Information transfer upon upscaling in such systems can be compressed (scaled), but not decomposed (Cushman, 1990). Constitutive variables are, in general, wave-vector and frequency dependent and result from nonlocality being expressed upon upscaling.

A primary goal of modelers of transport in heterogeneous media is to develop functional forms for constitutive variables and to know the regions of space-time that are important to their definition. If the constitutive theory depends upon what is happening at a point (here a point has meaning only with respect to the window and scale of resolution associated with the measurement process) in space-time, or a very small neighborhood of the point, then the theory is said to be local. On the other hand, if information is needed to define the constitutive theory from regions of space-time distinct from a neighborhood of the space-time point where evaluation of the theory is to be made, then the model is said to be nonlocal. These ideas will be made more concrete in subsequent sections.

15.4 Upscaling

The concept of "upscaling" is one of the most important in modern hydrological modeling circles. Upscaling simply means taking information on scales smaller than those of interest, homogenizing it, and studying the resultant "larger" scale system. The rationale behind the need to upscale is that we simply cannot make enough detailed measurements at the smaller scales to be of much use on the larger scales wherein relevant transport processes are taking place. An excellent example of this process was provided in Chapter 14, in which the local scale convective-dispersion equation was used to obtain a macroscale equation for the mean concentration and its fluctuations.

There are numerous methods that may be used to upscale the physics in hydrology, including matched asymptotic expansions, Taylor-Aris-Brenner method of moments, renormalization group theory, transform methods, volume averaging, particle tracking, etc. Many of these are illustrated in Cushman (1990). In this chapter we use transform and particle tracking methods.

15.5 Linearity and Nonequilibrium Eulerian Stochastic Models

15.5.1 Deterministic Nonequilibrium Sorption

It is assumed the chemical is transported by convection and local dispersion under steady, saturated, incompressible groundwater flow in a nondeformable porous medium, of constant porosity. The chemical undergoes first-order linear nonequilibrium kinetics with its surroundings. We assume a natural scale exists, the "Darcy" scale (as in Chapter 14), and we wish to upscale without knowledge of a scale separation between the Darcy and the "Reservoir" scale. The Darcy-scale transport is then governed by Equation (1). It is assumed the mean-flow is in the x_1-direction, that K_r and K_d are deterministic constants, and as a result of randomness in ln K, V_i, S, and C are random. Solving for the mean and fluctuation equations for Equation (1) and using a transform analysis gives (Cushman et al., 1996), assuming triplet correlation fluxes can be neglected,

$$
\frac{\partial \overline{C}}{\partial t} + \overline{V}_1 \frac{\partial \overline{C}}{\partial x_1} - \frac{\partial}{\partial x_j} \int_0^t \iint_{R^3} D_{ij}(x-y, t-t') \frac{\partial \overline{C}(y,t')}{\partial y_i} \, dy \, dt'
$$

$$
= -K_r K_d \int_0^t \left[\delta(t-t') - e^{-K_r(t-t')} \right] \overline{C}(x,t') \, dt' + K_r e^{-K_r t} S_o
$$

(3)

where

$$
D_{ij}(x-y, t-t') = d_{ij}\delta(x-y, t-t') + \overline{v_i v_j}(x-y) G(x-y, t-t')
$$

(4)

and

$$
\overset{\circ}{G}(k, \omega) = \left[\omega\left(1 + \frac{K_r K_d}{\omega + K_r}\right) + d_i k_i^2 + i k_1 \overline{V}_1 \right]^{-1}
$$

(5)

where $\overline{v_i v_j}$ is the assumed stationary fluctuating-velocity covariance, S_o and C_o are initial concentrations, the ° represents Fourier-Laplace transform where t is dual to ω and x is dual to k.

The form of the mean Equation (3) is quite different from the original local-scale Equation (1). This dispersive flux is nonlocal (third term on the left-hand side of Equation [3]), as is the source (left-hand

side of [3]) term. The upscaling process is the source of nonlocality and it results from a lack of information on the small-scale details of the transport process.

15.5.2 Random Nonequilibrium Sorption

Here the local-scale model is exactly as the previous case except it is assumed K_d is a random spatial field which may be correlated with ln K. In this case, the resultant mean equation is more complicated and given by (Hu et al., 1995)

$$
\frac{\partial \overline{C}}{\partial t} + \overline{V}_1 \frac{\partial \overline{C}}{\partial x_1} - \frac{\partial}{\partial x_j} \int_0^t \int_{R^3} D_{ij}(x-y,t-t') \frac{\partial}{\partial y_i} \overline{C}(y,t') \, dy \, dt'
$$

$$
- \frac{\partial}{\partial x_j} \int_0^t \int_{R^3} G(x-y,t-t') \overline{k_d v_j}(x-y) \overline{C}(y,t') \, dy \, dt'
$$

$$
= -K_r \left\{ \overline{K}_d \overline{C} - e^{-K_r t} S_0 - K_r \overline{K}_d \int_0^t e^{-K_r(t-t')} \overline{C}(x,t') \, dt' - \int_0^t \left[\delta(t-t') - K_r e^{-K_r(t-t')} \right] \right. \tag{6}
$$

$$
\left[\int_0^{t'} G(x-y,t'-t'') \overline{k_d k_d}(x-y) \overline{C}(y,t'') \, dy \, dt'' \right.
$$

$$
\left. + \int_0^{t'} \int_{R^3} B(x-y,t'-t'') \overline{k_d v_l}(x-y) \frac{\partial \overline{C}}{\partial y_i}(y,t'') \, dy \, dt'' \right] dt \right\}
$$

with

$$
D_{ij}(x-y,t-t') = d_{ij}\delta(x-y,t-t') + \overline{v_i v_j}(x-y)B(x-y,t-t') \tag{7}
$$

$$
\overset{\circ}{B}{}^{-1}(k,\omega) = \omega \left(1 + \frac{K_r \overline{K}_d}{\omega + K_r} \right) + d_i k_i^2 + i k_1 \overline{V}_1 \tag{8}
$$

and

$$
\overset{\circ}{G}(k,\omega) = \overset{\circ}{B}(k,\omega) \frac{\omega K_r}{\omega + K_r} \tag{9}
$$

The addition of randomness in K_d again increases the complexity of the mean equation. In this case, in addition to nonlocal sources/sinks and dispersion, there is also nonlocality in the convective flux (fourth term on the left-hand side of Equation [6]).

15.5.3 Random Nonequilibrium Sorption and Equilibrium Microbial Decay

It is possible to generalize the previous results to the case where there is linear first-order microbial degradation. In this case, the local-scale Equation (1) is replaced by

$$\frac{\partial C}{\partial t} + \frac{\partial S}{\partial t} = \frac{\partial}{\partial x_j} d_{ij} \frac{\partial C}{\partial x_i} - \frac{\partial (V_i C)}{\partial x_i} - K_c C - K_s S \tag{10}$$

which is coupled with Equation (2). Here K_c and K_s are the solute degradation rates in the solution and sorbed phases. The mean equation can be obtained for this problem, but it is too complicated to present here. See Hu et al. (1997) for its presentation and development. In the next subsection we will, however, present some numerical solutions to it.

15.5.4 Numerical Results for Random K_d and No Degradation

To apply Equation (6), one needs to know three covariance functions $\overline{v_i v_j}$, $\overline{k_d k_d}$, and $\overline{k_d v_i}$ where v_i and k_d are the fluctuating velocity in the ith-direction and the fluctuating partition coefficient, respectively. Since it is difficult, if not impossible, to measure v_i, it is common to functionally relate v_i to f (the fluctuating log-conductivity) and then to measure the covariance of f and its cross-covariance with k_d. First-order expressions relating v_i to f and v_i to k_d are (Gelhar and Axness, 1983; Hu et al., 1995)

$$\widehat{\overline{v_i v_j}}(k) = \left(\frac{Q_g}{n}\right)^2 \left[\delta_{i1} - \frac{k_1 k_i}{|k|^2}\right]^2 \widehat{\overline{ff}}(k) \tag{11}$$

and

$$\widehat{\overline{v_i k_d}}(k) = \frac{Q_g}{n} \left[\delta_{i1} - \frac{k_1 k_i}{|k|^2}\right] \widehat{\overline{f k_d}}(k) \tag{12}$$

where Q_g is the product of the geometric mean conductivity with the mean gradient in the x_1-direction, n is the porosity, and the caret indicates Fourier space transform. Both $\overline{ff}$ and $\overline{fk_d}$ must be obtained from field experiments. For illustrative purposes, we assume various correlation structures for these covariances. For perfect positive correlation (Model A) or perfect negative correlation (Model B) to first order, we have

$$\widehat{\overline{k_d k_d}}(k) = \left(K_d^G\right)^2 e^{\sigma_f^2} \widehat{\overline{ff}}(k) \tag{13}$$

and

$$\widehat{\overline{k_d f}}(k) = \pm K_d^G e^{\sigma_f^2} \widehat{\overline{ff}}(k) \tag{14}$$

with plus for Model A and minus for Model B, where K_d^G is the geometric mean of $K_d(x)$. For uncorrelated K_d (Model C), to first order,

$$\widehat{\overline{k_d k_d}}(k) = \left(K_d^G\right)^2 e^{\sigma_W^2} \widehat{\overline{WW}}(k) \tag{15}$$

and

$$\overline{\hat{fk}_d}(k) = 0 \tag{16}$$

where W is a normally distributed random space function with zero mean, variance σ_W^2, and covariance

$$\overline{WW}(z) = \sigma_W^2 e^{-z/l_W} \tag{17}$$

where l_w is the integral scale of W and l is the correlation length of ln K.

We take the initial concentration C_o to be constant, C_m over a rectangular prism $2a_1 \times 2a_2$ and assume $S_0 \equiv 0$. The log-conductivity is chosen to be exponential

$$\overline{ff}(z) = \sigma_f^2 \exp\left[-|z|^2 / l^2 \right] \tag{18}$$

Model D will correspond to deterministic and constant K_d.

In Figure 15.1 (from Hu et al., 1995], Figure 3), we plot the various spatial moments (see Chapter 14) as a function of nondimensional time $t' = t\overline{V}/\overline{R}l$ where $\overline{R} = 1 + \overline{K}_d$. It is seen that the first moment is little affected by the models irrespective of K_r. Relative to the uncorrelated model, negative correlation between the chemical heterogeneity and physical heterogeneity increases the second longitudinal moment, and positive correlation decreases it. The difference between the three models becomes larger as K_r increases. Deterministic K_d can produce significantly different results from random K_d. Figure 15.1C shows that the difference in the second transverse moments between the four models is small, but distinct at large time.
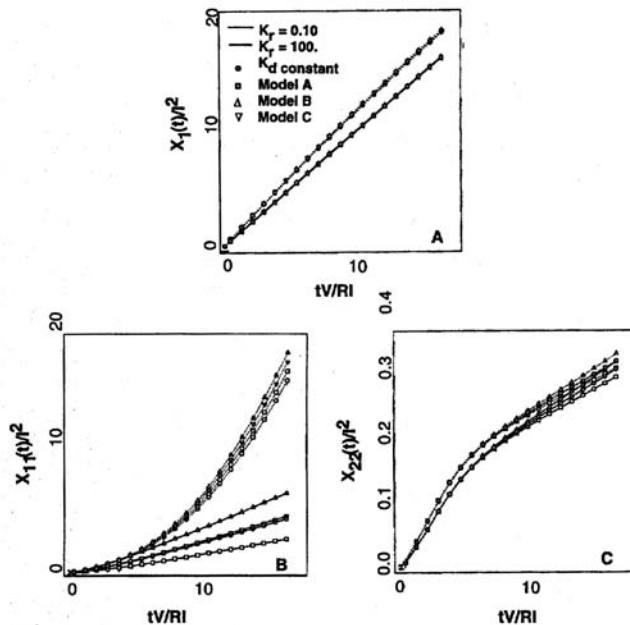
FIGURE 15.1 Spatial moments for various models with $K_d^G = 0.2$: (A) first moment, (B) second longitudinal moment, (C) second transverse moment. (Adapted from Hu, B. S., Deng, F. W., and Cushman, J. H. 1995. A nonlocal theory of reactive transport with physical and chemical heterogeneity: linear nonequilibrium sorption with random K_d. *Water Resour. Res.* 31(9):2219-2237.)

15.5.5 Numerical Results with Random Chemistry and Microbiology

The mean concentration for this case can be obtained given a much larger number of covariance functions,

$$\overline{v_i v_j}, \ \overline{v_i k_d}, \ \overline{v_i k_c}, \ \overline{v_i k_s}, \ \overline{k_d k_d}, \ \overline{k_d k_s}, \ \overline{k_d k_c}, \ \overline{k_c k_c}, \ \overline{k_s k_s}, \text{ and } \overline{k_s k_c}$$

We call the case without degradation model D. In this case, $K_s = K_c \equiv 0$. We again must relate v_i, k_d, k_c, and k_s to f and in general, it follows to first order (with $a = d, c, s.$)

$$\widehat{\overline{v_i k_a}}(k) = \left(\frac{Q_g}{n}\right)\left[\delta_{i1} - \frac{k_1 k_i}{|k|^2}\right]\widehat{\overline{f k_a}}(k) \tag{19}$$

For illustrative purposes, we assume k_c and k_s are gaussian and

$$\overline{k_c k_s} = \sigma_c \sigma_s \exp\left[-r^2/l_c^2\right] \tag{20}$$

We assume positive (Model A), negative (Model B), and uncorrelated (Model C) covariance structures in the form

$$\overline{k_c f} = \pm\sigma_c \sigma_f \exp\left[-r^2/l_{cf}^2\right],$$

$$\overline{k_s f} = \pm\sigma_s \sigma_f \exp\left[-r^2/l_f^2\right],$$

where $+$ is for Model A, and $-$ for Model B, and k_d is linearly negatively correlated with f. Use the same initial data as the previous case with $l_c = l_{cf} = l_f = 1.0$ m, $\sigma_f^2 = 0.2$, $\overline{V}_1 = 1.0$ m/day, $d_1 = 0.05$ m²/day, $d_2 = 0.005$ m²/day, $c_m = 1$, $\overline{K}_d = 1$/day, $\overline{K}_c = 0.01$/day, $\overline{K}_s = 0.003$/day, $\sigma_c^2 = 0.01$, and $\sigma_s^2 = 0.003$.

Figure 15.2 plots mean concentrations contours for the various models at 80 days. Models A, B, and C produce similar mean distributions. The shape of Model D's contours are similar to the others, but they are larger in area and have higher mean concentrations, owing to a lack of degradation.

15.6 Lagrangian Stochastic Models for Nonlinear Reactions

Mass transfer equations representing natural biological and chemical transformations in saturated sub-surface systems typically involve strongly nonlinear transformations among multiple species. Eulerian methods that result in ensemble-averaged expected plume behavior have been successful in upscaling the effects of idealized heterogeneities on the fate and transport of linearly sorbing solutes at both equilibrium (e.g., Berglund and Cvetkovic, 1996) and nonequilibrium rates (e.g., above, cf. also Quinodoz and Valocchi, 1993; Dagan and Cvetkovic, 1993; Andricevic, 1995), and solutes undergoing linear decay (e.g., Hu et al., 1995; Miralles-Wilhelm and Gelhar, 1996). These methods are, however, generally inappropriate for application to reactive transport where nonlinearities exist in nonequilibrium transformations or in multicomponent (coupled) dependencies. Sudicky et al. (1990) points out that the ergodic requirement (that, simply, the solute domain is large relative to the dominant scale of heterogeneity) is invalid for a nonconservative (e.g., sorbing or degraded) solute whose domain shrinks with time. Also, the use of a macrodispersion parameter effectively represents as mixing what may actually be plume deformation due to nonuniform convections, as described in the previous chapter. In the case of dually-limiting reactants (interior and exterior to the plume, such that reactions only occur at the mixing fringe)
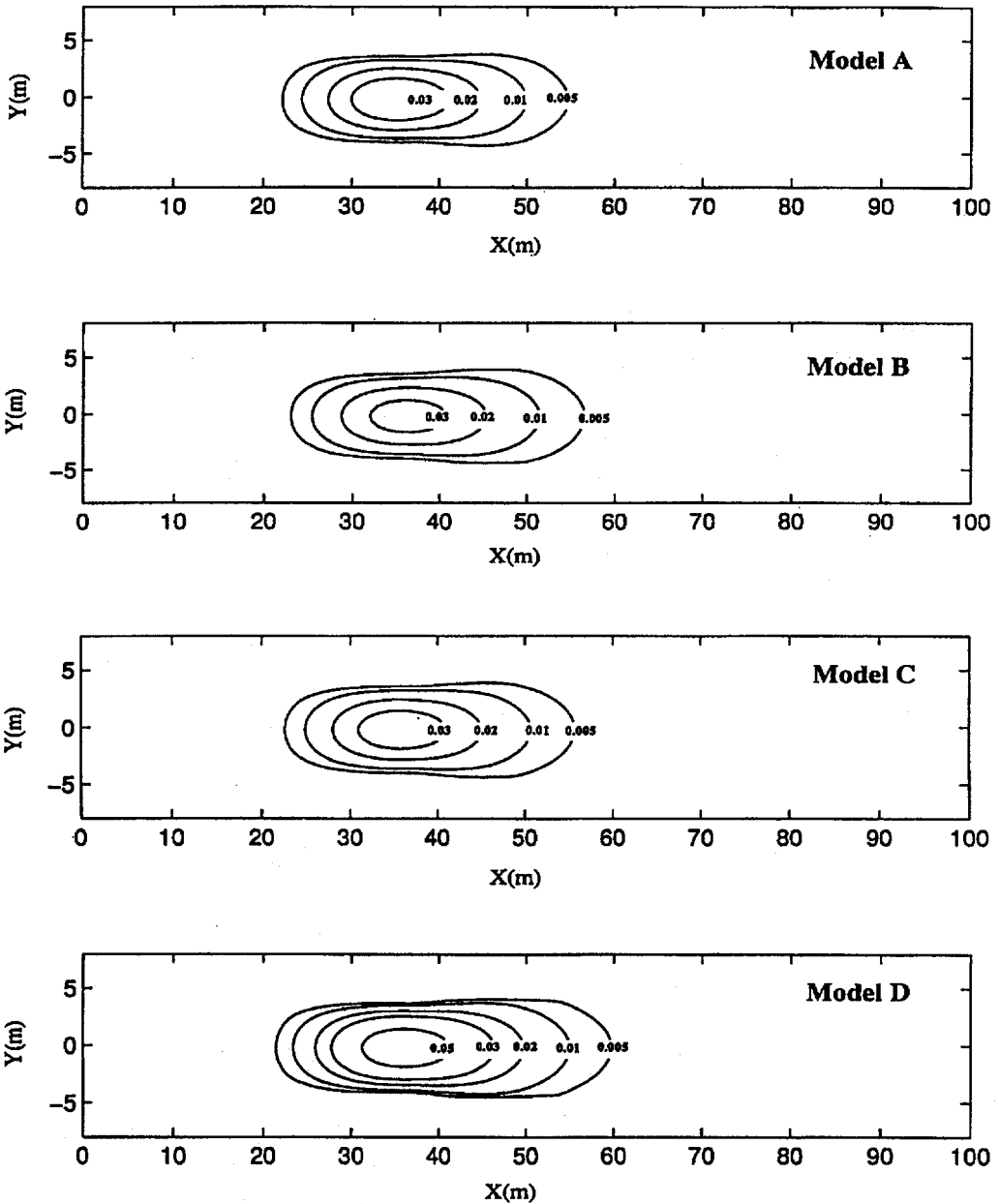
FIGURE 15.2 Concentration contours for various models at t = 80 days for $K_r = \overline{K_d} = 1/\text{day}$, $\overline{K_c} = 0.01/\text{day}$, and $\overline{K_s} = 0.003/\text{day}$. (Adapted from Hu, B. X., Cushman, J. H., and Deng, F. W. 1998. Nonlocal theory of reactive transport with physical, chemical, and biological heterogeneity. *Adv. Water Resour.* In press.)

this artificial mixing will significantly overestimate the reaction sink. The scaling procedure, as an equation averaging operation, is valid when the averaging operator commutes with the (linear) kinetic reaction operator, so that the average of quantities *within* the reaction term may be obtained. When nonlinearities exist in the kinetics of transformation, the averaging and reaction operations do not generally commute, and so the scaling approach fails to generate useful equations for average plume behavior.

An approach to scaling complex reactive transport in heterogeneous materials may be obtained when one first applies the reaction operation to solute parcels as they follow nonuniform convective paths, and then averages the resulting reacted solute convections over the appropriate convective paths contributing to an observation. In this approach, known as the solute flux-based (Cvetkovic and Dagan, 1996) or stochastic-convective reaction (SCR) method (Simmons et al., 1995), the reactions are cast in terms of the solute residence-time during convection. The method is based in the stochastic-convective representation of dispersive transport of a passive tracer (Simmons, 1982). The travel-time of the solute is defined as the time of residence of solute moving through the aquifer, from the point of spill or contamination entry to the system, to the observation location or regulatory boundary. Solute particles that enter the system and arrive at the observation location follow particular paths, or streamtubes, and the set of streamtubes delivering solutes to an observation sampler (point), line (well), or plane (regulatory boundary) makes up an *ensemble* of travel paths, and so the travel-time describing the ensemble is a distributed quantity, characterized by a distribution function. When the reactive transport model along any given streamtube can be expressed in terms of travel-time, the solution to the model is in terms of travel-time, and the solute arrival at the observation is recovered by averaging the solution of the model against the travel-time distribution function.

The advantages of this approach are (1) the reliance upon travel-time, a global measure of the effects of locally varying Eulerian velocities, as opposed to estimates of the full nonuniform velocity field, and (2) the solution for each streamtube equation system can be obtained *before* averaging, thus eliminating the operator commutation problem with scaling nonlinearly reactive transport. This scaling approach is usually applied to provide predictions of observed solute arrivals at observation samples, or arrivals at an imaginary control plane. It is a direct although cumbersome matter to use the SCR approach to reconstruct estimates of plume evolution, which may be done if the travel-time distribution function is given or estimated at many points in the aquifer. Similar streamtube principles have been used to gain computational speedup in two- and three-dimensional reactive transport simulations in a deterministic setting (Thiele et al., 1995a,b).

The method proceeds in three basic steps. First, the existence of an ensemble of streamtubes that is suitable for averaging must be ensured by checking the validity of the SCR assumptions for the particular reaction system, initial, and boundary conditions. Second, the reaction system must be solved for the streamtube ensemble (or canonical streamtube, defined below, when it exists). Finally, the travel-time distribution function must be determined for the particular flow regime and initial and boundary conditions.

15.6.1 Streamtube Formulation

The concept of the streamtube as a path taken by solute parcels from a source to an observation is made under the requirement that flow in it (a subset of the total flow domain) is constant of both space and time. The assumptions under which such an ensemble of streamtubes may exist are as follows. First, the members of the ensemble are physically independent, with no mass transfer among streamtubes. This allows treatment of the solute paths as integral, averagable quantities. This restricts diffusive/dispersive mass transports lateral to streamtubes; e.g., this representation is incapable of incorporating local transverse dispersion. For simplicity we ignore also longitudinal mixing processes, although this is not a requisite assumption because a longitudinal mixing may be incorporated within the streamtube (Ginn et al., 1995). Second, the streamtube orientation in Eulerian space is arbitrary save that it must of course connect the source to the measurement location, and that the projection of its axis, s, onto the x-axis connecting the source and observation must be one-to-one (such that the x(s) is monotone increasing). This assumption allows us to associate with each streamtube a one-dimensional velocity function $v(x)$ ε {v(x)} that is positive, and so allows the definition of an associated travel-time function. Further, each streamtube obeys a known reaction function r. A streamtube in three-space (which may have variable areal cross section) is defined such that its average cross-sectional velocity $v_s(x')$ times the corresponding cross-sectional water content $n_s(x')$ is a constant, q. To represent a streamtube in three dimensions with

an equivalent one-dimensional streamtube with the same flow q we must honor the velocity in order to preserve the transport characteristics. Velocity is defined by $v_s(x') = q/n_s(x')$ in the streamtube and by $v(x') = q/n(x')$ in the equivalent one-dimensional system. Mass balance implies $v_s(x')n_s(x')A_s(x') = v(x')n(x')A_o$ where A_o is a (constant) reference area, and so we may define an effective porosity $n(x') = n_s(x')A_s(x')/A_o$ to give $v(x') = v_s(x')$ along the equivalent one-dimensional system. Under this construction the convective-reactive transport along a given streamtube is represented in the one-dimensional form

$$\frac{\partial(nC)}{\partial t} + \frac{\partial(nvC)}{\partial x} = nr \tag{21}$$

By construction $n(x)v(x) = q$ is a constant of x, and n is assumed constant of t, so Equation (21) simplifies to

$$\frac{\partial C}{\partial t} + v(x)\frac{\partial C}{\partial x} = r \tag{22}$$

A simple change of variables in Equation (22) allows expression of velocity in terms of solute travel-time. The travel-time corresponding to the streamtube velocity $v(x)$ is the time elapsed as a tracking particle is transported from x_o to x,

$$T(x; x_o) = \int_{x_o}^{x} dx'/v(x') = \int_0^x dx'/v(x') - \int_0^{x_o} dx'/v(x') = T(x) - T(x_o) \tag{23}$$

where $T(x)$ means $T(x',0)$. By positivity of velocity $T(x)$ is a monotone increasing function in x and the inverse function, $\Xi(T'(x)) = T^1(\tau) = x$, exists and is monotone increasing in variable τ. Note that mass loss over the ensemble corresponds to the case where ensemble maximum travel-time is infinity. In this case a finite T^* would be used which represents the largest travel-time of a tracer particle captured at a measurement location, and the mass lost would be accounted in the specification of the streamtube frequency distribution. Now invoking the chain rule with $dT(x) = dx/v(x)$ and incorporating the inverse function $\Xi(T)$ we may write for Equation (22)

$$\frac{\partial C}{\partial t} + \frac{\partial C}{\partial T} = r(\Xi(T), t, C) \tag{24}$$

This last form expresses the result that *C* can be obtained explicitly as a function of global travel-time. Equation (7) is the basic streamtube model. Note that if *r* is streamtube invariant (requiring that *r'* contain effective streamtube porosity as factor) the solution to Equation (22) depends only on *v*. Any set of streamtubes for which the velocity statistics are identical to those of {*v*} (determined from observation of conservative tracer breakthrough or from conductivity statistics), and for which r accurately represents the reactions, suffices. A working hypothesis of this approach is that the streamtube model more closely represents the actual conditions for the definition and calibration of the kinetic model r than does the conventional averaging approach. Closed-form solutions to Equation (22) may be found under general conditions on r, including nonlinear and dynamic forms (e.g., r depends on a parameter such as microbial mass that evolves according to its own differential equation). We will return to a few examples after introducing the travel-time distribution function.

Ensemble averaging of the solution to Equation (24), described next, is greatly simplified when the solution to Equation (24) may be written in terms of dimensionless travel-time, and rescaled to fit any particular travel-time. When the solution to the streamtube equation is rescalable over travel-time, the solution is termed *canonical*.

15.6.2 Stochastic-Convective Averaging

Here we examine averaging of the foregoing solutions according to the travel-time pdf, a measure of streamtube frequency. Once a canonical solution of Equation (7) is obtained as $C(T,t)$, the stream tube ensemble average solution, $\langle C \rangle(x,t)$, is found by averaging $C(T,t)$ over the travel-time pdf $p(t;x)$ for the streamtube travel-times to location x. This pdf is defined over the streamtube ensemble in the frequency sense as (Simmons, 1986; Dagan and Nguyen, 1989)

$$p\big(\tau;x\big) = \Big\langle \delta\big(\tau - T(x)\big)\Big\rangle \tag{25}$$

where δ is the Direc distribution function and the angle brackets indicate average over the streamtube ensemble. The expectation can be expressed as

$$\langle C \rangle \big(x,t\big) = \int_0^\infty C\big(\tau,t\big) p\big(\tau;x\big) d\tau \tag{26}$$

In Equation (26), $p(\tau;x)$ is the pdf for random travel-time τ to a plane normal to the major flow axis at x, measured along that axis (Simmons, 1982, 1986; Dagan and Nguyen, 1989; Wise and Charbeneau, 1994). The averaging in Equation (26) is over the realizations of (random) travel-time at a certain point in x space, and not over the deterministic measure of travel-time $T(x)$ itself. That is, the average on the right-hand side of Equation (26) is exactly the streamtube ensemble average, as is demonstrated on substitution of Equation (25):

$$\int_0^\infty C\big(\tau;t\big) p\big(\tau;x\big) d\tau \equiv \int_0^\infty C\big(\tau,t\big)\Big\langle \delta\big(\tau - T(x)\big)\Big\rangle d\tau = \left\langle \int_0^\infty C\big(\tau;t\big)\delta\big(\tau - T(x)\big) d\tau \right\rangle = \Big\langle C\big(T(x),t\big)\Big\rangle \tag{27}$$

The numerical calculation of Equation (27) via discrete integration requires first the enumeration of the surfaces $C_i(T,t)$ for each realization of travel-time function $T_i(x)$. When a canonical solution is available this is accomplished by simply rescaling the travel-time axis of the canonical solution $C(T,t)$ to the indicated $T_i(x)$. This approach is used by Wise and Charbeneau (1994), whose averaging is associated with the measured cumulative passive breakthrough (which they term the "fractional breakthrough," that is the integral of our travel-time pdf).

Finally, one may also introduce heterogeneity into the reaction term through either the space-time dependency and/or the solute component dependency in the right-hand side of Equation (24) [for instance when $r(x,t,C) = g(x,t)\,h(C)$, and g indicates concentration of reactive iron oxides or attached biomass]. Further, correlations between the flow properties and the reaction properties may be cast as a conditioning of the travel-time pdf upon the value of the reaction properties. When such heterogeneity is represented through pdfs of the random properties and when these pdfs are known, the expectation may be computed in many cases by simple sequential averaging of $C(T,t)$ according to each random property (with pdfs appropriately conditioned to reflect correlations). Suppose, for instance, randomness in the case of a factorable hyperbolic kinetic decay (e.g., biodegradation following Michaelis-Menten kinetics), such that we write Equation (24) as

$$r\big(x,t,C\big) = \big\{g\big\}\big(x,t\big)h\big(C\big) \tag{28}$$

where g represents time evolution of a random space function selected from the ensemble $\{g\}$. That is, Equation (28) is the reactive term written for one realization (streamtube) of the random pathways. In this case (Simmons et al., 1995) the streamtube solution is cast in terms of not only travel-time (cumu-

lative reciprocal velocity) but also cumulative "reaction property" G, defined as the integral of g along the streamtube coordinate. Then the travel-time averaging in Equation (26) is followed by averaging over the ensemble distribution of G, appropriately conditioned;

$$\left\langle C(x,t) \right\rangle_G = \int_0^\infty C(\tau,t) p(\tau; x|G) d\tau \tag{29a}$$

$$\left\langle C(x,t) \right\rangle = \int_{-\infty}^\infty \left\langle C(\tau,t) \right\rangle_G p(G; x) dG \tag{29b}$$

15.6.3 Travel-Time Distribution Function

The travel-time $T(x)$ determines the mapping of the random velocity along a streamtube onto a random velocity field for an equivalent one-dimensional system with distance coordinate x. Thus the travel-time pdf is experimentally observed as the conservative tracer breakthrough curve at the control plane at x in response to a unit Dirac-δ input function. Alternatively, the pdf may be deconvolved from the observed response to other known input functions (e.g., Ginn et al., 1995). In either case it should be noted that pore-scale mixing processes (such as diffusion and pore-scale dispersion) serve to smooth the observed breakthrough curve in an irreversible way. One may also hypothesize forms for $p(\tau;x)$, a priori, such as lognormal or inverse-Gaussian (Simmons, 1982).

Frequently, tracer test results are not available for characterizing transport between a source and an observation location. It is possible to estimate the *expected* travel-time distribution function that corresponds to the average arrival time distribution over an ensemble of flow fields. To derive the expected travel-time distribution for particles arriving at an observation point at distance x_1 from source centroid, $g(t;x_1)$, at first order, we follow Dagan (1989, ch. 5.8; see also Appendix A in Cvetkovic and Dagan, 1996). For simplicity it is assumed that all particles start within a finite plane source that is oriented normal to the average flow direction. A particle starts at coordinates $(0, a_2, a_3)$, at initial velocity parallel to x-axis $v_o = V_1.(0, a_2, a_3)$. The probability that a solute particle has crossed the regulatory control plane (assumed normal to the flow direction $\bar{x}$, located at x_1) at time $t = \tau$ is the same as the probability that the particle's displacement along the flow axis is larger than the distance to the control plane, i.e., that $X_1 > x_1$. Hence, with $\phi(X_1, v_0;t)$ as the joint pdf of X_1, v_o, at time t, one has for the cumulative arrival distribution

$$G(\tau; v_0, x_1) = \int_{x_1}^\infty \phi(X_1; v_0, \tau) dX_1 \tag{30}$$

It is important to note that the concentration probability function ϕ is a statistical construction that is not equivalent to any particular physical realization unless the transport is ergodic, e.g., that the lateral domain of the source is large compared to the heterogeneity scale. Various models for ϕ have been derived based on different constructs of the statistical description of the controlling (conductivity) field and boundary conditions (Dagan, 1991; Dagan et al., 1992). The expected travel-time pdf is obtained from the cdf G by averaging Equation (30) over possible initial velocities v_o (where v_o is assumed gaussian with mean U) and differentiating with respect to τ.

$$g(\tau; x_1) = \frac{1}{U} \int v_o g(\tau, v_o; x_1) dv_o = \frac{1}{U} \frac{\partial}{\partial \tau} \int_0^\infty \int_{x_1}^\infty v_o \phi(X_1, v_o; \tau) dX_1 \, dv_o \tag{31}$$

For instance, for asymptotic conditions where the impact of v_o and τ correlations vanish, g is given by

$$g(t; x_1) = -\frac{1}{\sqrt{2\pi}} \exp\left(\frac{-(x_1 - U\tau)^2}{2X_{11}}\right) \frac{\partial}{\partial \tau}\left(\frac{x_1 - U\tau}{\sqrt{X_{11}}}\right) \tag{32}$$

where X_{11} is the longitudinal particle covariance, a measure of the ensemble-averaged longitudinal second spatial moment, and U is the source-averaged initial velocity.

15.6.4 Some Reaction System Examples

Here we summarize some recent examples of reaction systems that have been constructed for particular biogeochemical systems and used in streamtube-ensemble reactive transport models. The examples include a generalized linear sorption model and a model of aerobic biodegradation with microbial transport.

Lassey (1988) provided a generalized formulation of a model combining equilibrium, kinetic-reversible, and kinetic-irreversible sorption processes and its solution for particular boundary conditions. The solution to the streamtube Equation (24) with r given by the kinetic reversible sorption model of Equations (1a) and (1b) [e.g., $r = -(\rho_b/n) \, \partial S/\partial t$], in the case of an instantaneous injection at the upstream boundary, is adopted from Lassey (1988; cf. also Cvetkovic and Dagan, 1996; Cushman et al., 1996)

$$C(\tau, t) = e^{-K_r K_d t}\delta(t - \tau) + K_r^2 K_d \tau e^{\left[-K_r K_d \tau - K_r(t - \tau)\right]} \tilde{I}_1\left[K_r^2 K_d \tau(t - \tau) H(t - \tau)\right] \tag{33}$$

where $\tilde{I}_1(z) = z^{-1/2} I_1(2z^{1/2})$, and I_1 is the modified Bessel function of the first kind of order one. This solution is appropriate for averaging as indicated in Equation (26). Note that as pointed out in Simmons et al. (1995) and Thiele (1995a), so-called "initial value problems", e.g., those involving the initial distribution of a reactant (such as attached biomass) in space, impose certain restrictive assumptions on the use of the streamtube approach. In the most restrictive case, uniformity of the distributed property over the streamtube ensemble is required. (This is only slightly less restrictive than the requirement of spatial biochemical homogeneity overall.) In the event of spatially nonuniform initial distribution of biomass, full averaging of the streamtube component evolution equations is feasible, but requires the joint probability of biomass occurrence and streamtube travel-time (as in Equation [29]).

Biogeochemical and biodegradation systems involving attached-phase biomass or biofilms typically formulate as initial value problems, and thus may be complicated by heterogeneity in not only initial microbial distributions, but also by the dynamic microbial heterogeneity associated with microbial growth. This is the case in the next example, involving nonlinear degradation of an organic solute by a microorganism that is initially uniformly attached to the solid phase but partitions to the aqueous phase according to a nonequilibrium and nonlinear rate equation. This example describes meter-scale laboratory experiments in aerobic biodegradation in physically heterogeneous media, and is greatly condensed from Murphy et al. (1997). Simplified evolution equations for the organic substrate benzoate, oxygen, mobile microbes, immobile microbes, and passive tracer bromide are shown here in one-dimensional (streamtube projection) format. Because we have retained one-dimensional longitudinal dispersivity d in the transport operation in Equation (34), in order to preserve the mixing between oxygenated and oxygen-depleted waters that is important to degradation, the equations are cast in terms of their streamtube velocity projected onto the x-axis, as opposed to travel-time. This aspect, and the possiblity of nonuniformities in the initial biomass concentrations, violate the requirements for a travel-time (i.e., constant-velocity) representation of streamtube velocities. Nevertheless for simplicity in calculating the SCR predictions, a constant-velocity ensemble, corresponding to an equivalent travel-time ensemble, is used as an approximation of the transport.

The reactions involve degradation of benzoate (C_b, Equation [34a]) as the electron-donor, with oxygen consumed as the terminal electron-acceptor (C_o, Equation [34b]). Degradation induces an increase in biomass, the rates of which are the same for aqueous (C_{mm}, Equation [34c]) and attached (C_{im}, Equation [34d]) biomass. The mass transformations associated with the degradation are governed by the bracketed dual-Monod terms in Equations (34a, b, c) that effectively shut down the transformation rate under limiting conditions, e.g., when C_o or C_b values drop to the order of K_o or K_b, respectively. Biomass partitions kinetically and reversibly, but with the added complexity that the equilibrium partition coefficient is dependent on the local ionic strength, changes in which are due to the presence of an inert bromide tracer (C_T, Equation [34e]) that effectively compresses the electrostatic repulsive layer of the quartzitic solid phase. For simplicity, attached biomass is treated as a stationary volumeless dissolved species. These transformations incur a strong nonlinear coupling among the four mobile components (organic solute, oxygen, aqueous biomass, bromide) and the one immobile component (attached biomass). More complex forms, involving microbial metabolic lag and dynamic endogenous respiration, are described in Murphy et al. (1997). The transport takes place at average pore-water velocity of 50 cm/d through a 1 meter flow cell that is $0.1\text{m} \times 0.2$ m in cross-section, in inoculated porous media consisting of high hydraulic conductivity sand containing a set of randomly located low conductivity inclusions. Details on the heterogeneity pattern and pre-experimental modeling are found in Ginn et al. (1995).

$$\frac{\partial C_c}{\partial t} + \frac{\partial v_x C_c}{\partial x} - d\frac{\partial^2 C_c}{\partial x^2} = -\frac{\mu}{Y}\left(C_{mm} + C_{im}\right) \cdot \left\{\frac{C_o}{\left(K_o + C_o\right)}\frac{C_c}{\left(K_c + C_c\right)}\right\} \tag{34a}$$

$$\frac{\partial C_c}{\partial t} + \frac{\partial v_x C_c}{\partial x} - d\frac{\partial^2 C_c}{\partial x^2} = -F\frac{\mu}{Y}\left(C_{mm} + C_{im}\right) \cdot \left\{\frac{C_o}{\left(K_o + C_o\right)}\frac{C_c}{\left(K_c + C_c\right)}\right\} \tag{34b}$$

$$\frac{\partial C_{mm}}{\partial t} + \frac{\partial v_x C_{mm}}{\partial x} - d_{mm}\frac{\partial^2 C_{mm}}{\partial x^2} = -\mu C_{mm} \cdot \left\{\frac{C_o}{\left(K_o + C_o\right)}\frac{C_c}{\left(K_c + C_c\right)}\right\} - K_r K_d\left(C_T\right)C_{mm} + K_r C_{im} \tag{34c}$$

$$\frac{\partial C_{im}}{\partial t} = -\mu C_{im} \cdot \left\{\frac{C_o}{\left(K_o + C_o\right)}\frac{C_c}{\left(K_c + C_c\right)}\right\} + K_r K_d\left(C_T\right)C_{mm} - K_r C_{im} \tag{34d}$$

$$\frac{\partial C_T}{\partial t} + \frac{\partial v_x C_T}{\partial x} - d\frac{\partial^2 C_T}{\partial x^2} = 0 \tag{34e}$$

The initial conditions in the flow cell are fully oxygenated water ($C_o = 8$ mg/l) with equilibrium distributed biomass at 5×10^6 cells/ml in the aqueous phase. The distribution coefficient is roughly 2/3 in the absence of bromide and doubles at 300 mg/l bromide. The boundary conditions are continuous injection of oxygenated (8 mg/l) water with a 1-day pulse of benzoate (300 mg/l) and bromide (392 mg/l) in solution injected to evaluate the resulting biodegradation. Data collected include breakthrough curves at the outlet of the flow cell for benzoate, oxygen, and aqueous microorganisms. The first step in obtaining the SCR solution is the determination of a travel-time distribution function, which was taken from the breakthrough curve of the bromide tracer, shown by the "+" marks in the top panel of Figure 15.3. Using the principle of superposition in reverse allows deconvolution of the breakthrough curve to determine the breakthrough curve expected to result from an instantaneous injection (Dirac pulse) at the boundary, i.e., the travel-time distribution function. This is shown in the bottom panel of Figure 15.3, and the reconvolved simulation of the bromide breakthrough is shown as the solid line in the top panel of Figure
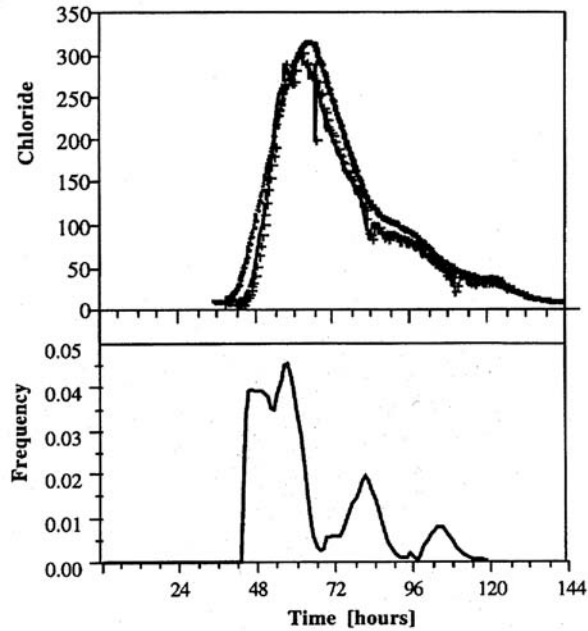
FIGURE 15.3 Top panel: Observed breakthrough of chloride (+) and reconstituted breakthrough curve (solid line) using convolution of the system travel-time distribution function with the 1-day square impulse function. Bottom panel: Travel-time distribution function.

15.3. (The deconvolution is not a trivial exercise because it involves solving a Volterra integral equation and is usually unstable numerically, although a unique solution exists.) In the second step, the SCR solution ensemble was generated by solving Equation (34) for 100 different but constant velocities corresponding to the travel-time domain shown in Figure 15.3. That is, effective velocities were selected according to $v = x/T$, where T is discretized on [40, 120] hours. These solutions were obtained using RAFT, a deterministic flow and reactive transport simulator capable of general reaction systems (Chilakapati, 1995).

The SCR predictions are obtained in the third step by averaging the ensemble of breakthrough curves per solute against the travel-time distribution function shown in Figure 15.3. This average was carried out by solving a discrete form of Equation (26) using an ensemble of 100. The resulting prediction is shown with observed data for benzoate in Figure 15.4.
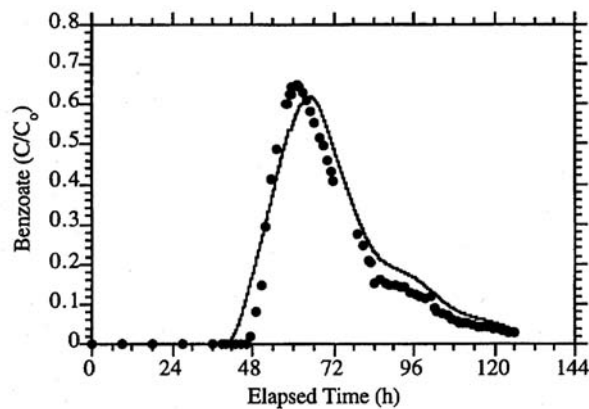
FIGURE 15.4 Observed (•) and simulated (solid line) breakthrough of benzoate.

Numerous reactions systems have been addressed with the streamtube approach, in part because of the facility in developing solutions in one transport dimension. In addition to the foregoing, generalized solution surfaces $C(\tau,t)$ have been derived for multicomponent nonlinear equilibrium sorption (Charbeneau, 1988; Dagan and Cvetkovic, 1996), immiscible two-phase displacement (the Buckley–Leverett problem, Thiele et al., 1995a), and simplified nonlinear decay, precipitation, or dissolution rate equations (Simmons et al., 1995; Dagan and Cvetkovic, 1996), to name a few.

15.7 Connection Between Approaches

The Eulerian methods described in Section 15.5 and the Lagrangian approach described above can be examined in light of one another through the use of a formal coordinate transformation that shows the particle (Lagrangian) paths in terms of the spatial (Eulerian) velocities (Dagan et al., 1992; Dagan and Cvetkovic 1993; Cvetkovic and Dagan, 1996; Cushman et al., 1996). In this section we present this coordinate transformation for reference. We begin conceptually again with Equation (1), taking $d_{ij} = 0$ and steady divergence-free velocity field $v(x)$, the mean of which is in the x_1 direction. Further, the magnitude of $v_1(x)$ is restricted to be positive. Now to trace the path of a solute parcel or particle in such a system, define a particle coordinate tracking function $X(t; t^0, x^0)$ as a solution to the Lagrangian equation $dX/dt = V(X)$, with the initial condition $X(t; t^0, x^0) = x^0$. Under steady velocity, we have also $X(t; t^0, x^0) = X(t - t^0; x^0)$. We restrict ourselves to the tracking of the particle path until it strikes an infinite plane positioned normal to the axis of mean flow, at x_1, and introduce the following coordinates to put transport in the x_1-direction in terms of travel-time and transport in the remaining two directions in terms of deviations from the starting point x^0.

$$T(x; x_o) = t - t_o \tag{35}$$

$$\eta(x; x_o) = X_2(\tau; x_o) \tag{36}$$

$$\xi(x; x_o) = X_3(\tau; x_o) \tag{37}$$

The travel-time function T is defined as the inverse of the particle position function X. The new coordinate system satisfies by construction $\tau = 0$, $\eta = x_2$, and $\xi = x_3$ for $x_1 = x_1^0$, and

$$\frac{d\tau}{dx_1} = \frac{1}{v_1(x_1, \eta, \xi)} \tag{38a}$$

$$\frac{d\eta}{dx_1} = \frac{d\eta}{d\tau}\frac{d\tau}{dx_1} = \frac{v_2(x_1, \eta, \xi)}{v_1(x_1, \eta, \xi)} \tag{38b}$$

$$\frac{d\xi}{dx_1} = \frac{d\xi}{d\tau}\frac{d\tau}{dx_1} = \frac{v_3(x_1, \eta, \xi)}{v_1(x_1, \eta, \xi)} \tag{38c}$$

Thus $x_2 = \eta(x_1, x_o)$ and $x_3 = \xi(x_1, x_o)$ are equations of streamlines (streamtube centerlines) for the particle path. In this coordinate system, along the streamlines, Equation (1a) takes the form of Equation (24), with r given by Equation (1b) [e.g., $r = -(\rho_b/n) \, \partial S/\partial t$]. Thus the transformation Equations (35) through (38) formally translates the governing equations into each other. Cushman et al. (1996) have used this transformation to compare the spatial moments of a plume evolving according to Equation (1) obtained

by both the Eulerian and Lagrangian (according to Dagan and Cvetkovic [1993], using Lassey's solution Equation [33]) approaches. In that case the zero and first spatial moments are equivalent between the two approaches, but the second moments of the Eulerian solution contain a linear dependence on time with coefficient in d_{ij}. In particular, the long-time transverse second moment is linear in time and otherwise constant, so the Eulerian approach predicts the plume width will grow linearly in time, while the Lagrangian approach predicts the plume width will reach a constant.

Note that this discrepancy is not inherent to any fundamental difference between the approaches themselves; rather, it arises from the assumption of zero dispersion d_{ij} in the Lagrangian approach. In the previous chapter the importance of pore-scale mixing (e.g., nonzero d_{ij}) to average plume dilution is illustrated and in the previous section the effect of mixing on scaling multicomponent nonlinear reactions is noted. This highlights a subtle aspect of mathematical scaling of processes in general. The total mass moved by pore-scale mixing is in general relatively small as the process is small scale in space and time (it is slow). However, it can play a significant role in the eventual large-scale and long-time behavior of the system.

Acknowledgment

For Further Information

Cushman (1997) provides an advanced treatment of stochastic models for reactive transport. Dagan (1989) provides an excellent discussion of stochastic models for nonreactive transport. Freeze and Cherry (1979) and Domenico and Schwartz (1990) provide excellent introductions to reactive chemical transport from a deterministic perspective.

References

Andricevic, R. 1995. Comment on "Stochastic analysis of the transport of kinetically sorbing solutes in aquifers with randomly heterogeneous hydraulic conductivity," by H. A. M. Quinodoz and A. J. Valocchi. *Water Resour. Res.* 31(1):237-243.

Berglund, S. and Cvetkovic, V. 1996. Contaminant displacement in aquifers: coupled effects of flow heterogeneiety and nonlinear sorption. *Water Resour. Res.* 32(1):23-32.

Bierkens, M. F. P. and Weerts, H. J. T. 1994. Block conductivity of cross-bedded fluvial sediments. *Water Resour. Res.* 30:2665-2678.

Charbeneau, R. J. 1988. Multicomponent exchange and subsurface solute transport: characteristics, coherence, and the Rieman problem. *Water Resour. Res.* 24, 57-64.

Chiang, C. Y., Dawson, C. N., and Wheeler, M. F. 1990. Modelling of *in situ* biorestoration of organic compounds in groundwater. *Tech. Rept. TR90-31*, Rice University. Houston, TX.

Chilakapati, A. 1995. A simulator for reactive flow and transport of groundwater contaminants. *PNNL-10636*, Pacific Northwest National Laboratory. Richland, WA.

Cushman, J. H. (ed.). 1990. *Dynamics of Fluids in Hierarchical Porous Media.* Academic Press, London.

Cushman, J. H. 1997. *The Physics of Fluids in Hierarchical Porous Media: Angstroms to Miles.* Kluwer Academic, New York.

Cushman, J. H., Hu, B. X., and Deng, F.-W. 1996. Comparison of Eulerian to Lagrangian expected spatial moments for transport in a heterogeneous porous medium with deterministic linear nonequilibrium sorption. *Chem. Eng. Comm.* Volumes 148-150:5-21.

Cvetkovic, V. and Dagan, G. 1996. Reactive transport and immiscible flow in geological media. II. Applications. *Proc. R. Soc. Lond.* A 452:303-328.

Dagan, G. 1989. *Flow and Transport in Porous Media.* Springer-Verlag, New York.

Dagan, G. 1991. Dispersion of a passive solute in non-ergodic transport by steady velocity fields in heterogeneous fomrations. *J. Fluid Mech.* 233:197-210.

Dagan, G. and Cvetkovic, V. 1993. Spatial moments of a kinetically sorbing solute plume in a heterogeneous aquifer. *Water Resour. Res.* 29(12):4053-4061.

Dagan, G., Cvetkovic, V., and Shapiro, A. 1992. A solute-flux approach to transport in heterogeneous formations. 1. The general framework. *Water Resour. Res.* 28:1369-1376.

Dagan, G. and Cvetkovic, V. 1996. Reactive transport and immiscible flow in geological media. II. General theory. *Proc. R. Soc. Lond.* A 452:285-301.

Dagan, G. and Nguyen, V. 1989. A comparison of travel-time and concentration approaches to modeling transport by groundwater. *J. Contam. Hydrol.* 4:79-91.

Domenico, P. A. and Schwartz, F. W. 1990. *Physical and Chemical Hydrogeology.* John Wiley & Sons, New York.

Dzombak, D. A. and Morel, F. M. M. 1990. *Surface Complex Modeling; Hydrous Ferric Oxide.* Wiley-Interscience, New York.

Fetter, C. W. 1993. *Contaminant Hydrology.* Macmillian Pub. Co., New York.

Freeze, R. A. and Cherry, J. A. 1979. *Groundwater.* Prentice-Hall, Englewood Cliffs, NJ.

Gear, C. W. 1988. Differential-algebraic equation index transformations. *SIAMJ. Sci. Comput.* 9(1):39-47.

Gear, C. W. and Petzold, L. R. 1984. ODE methods for the solution of differential/algebraic systems. *SIAM J. Numer. Anal.*, 21:716-728.

Gelhar, L. W. and Axness, C. L. 1983. Three-dimensional stochastic analysis of macrodispersion in aquifers. *Water Resour. Res.* 19:161-190.

Ginn, T. R., Simmons, C. S., and Wood, B. D. 1995. Stochastic-convective transport with nonlinear reaction: biodegradation and microbial growth. *Water Resour. Res.* 31:2689-2701.

Hassan, A. E., Cushman, J. H., and Delleur, J. W. 1997. The significance of porosity variability to flow and transport in random heterogeneous porous media. *Water Resour. Res.* (submitted).

Hu, B. X., Cushman, J. H., and Deng, F. W. 1997. Nonlocal theory of reactive transport with physical, chemical, and biological heterogeneity. *Adv. Water Resour.* 20:293-308.

Hu, B. S., Deng, F. W., and Cushman, J. H. 1995. A nonlocal theory of reactive transport with physical and chemical heterogeneity: linear nonequilibrium sorption with random K_d. *Water Resour. Res.* 31(9):2219-2237.

Karickhoff, S. 1983. Organic pollutant sorption in aquatic systems. *J. Hydraul Eng.* 110:707-733.

Lassey, R. R. 1988. Unidimensional solute transport incorporating equilibrium and rate-limited isotherms with first-order loss, 1, Model conceptualizations and analytic solutions, *Water Resour. Res.* 24:343-350.

Lichtner, P. C., Steefel, C. I., and Oelkers, E. H. (eds.). 1996. *Reactive Transport in Porous Media; Reviews in Mineralogy, Volume 34.* P. H. Ribbe, series editor. Mineralogical Society of America, Washington, D.C.

Miralles-Wilhelm, F. and Gelhar, L. W. 1996. Stochastic analysis of transport and decay of a solute in heterogeneous aquifers. *Water Resour. Res.* 32(12):3451-3459.

Murphy, E. M., Ginn, T. R., Chilakapati, A., Resch, C. T., Wietsma, T. W., and Phillips, J. L. 1997. The influence of physical heterogeneity on microbial degradation and distribution in porous media. *Water Resour. Res.* 33:1087-1104.

Novak, C. F. and Sevougian, S. D. 1993. Propagation of dissolution/precipitation waves in porous media, in *Migration and Fate of Pollutants in Soils and Subsoils*, D. Petruzzelli, F. G. Helffereich (eds.). NATO ASI Series, Series GL Ecological Sciences, Vol. 32. Springer-Verlag, Berlin.

Quinodoz, H. A. M. and Valocchi, A. J. 1993. Stochastic analysis of the transport of kinetically sorbing solutes in aquifers with randomly heterogeneous hydraulic conductivity. *Water Resour. Res.* 29(9):3227-3240.

Rubin, J. 1990. Solute transport with multisegment, equilibrium-controlled reactions: a feed forward simulation method. *Water Resour. Res.* 26(9):2029-2055.

Rubin, J. 1983. Transport of reacting solutes in porous media: relationship between mathematical nature of problem formation and chemical nature of reactions. *Water Resour. Res.* 19(5):1231-1252.

Scheibe, T. D. and Freyberg, D. L. 1995. Use of sedimentological information for geometric simulation of natural porous media structure. *Water Resour. Res.* 31:3259-3270.

Simmons, C. S. 1982. A stochastic-convective transport representation of dispersion in one-dimensional porous media systems. *Water Resour. Res.* 18:1193-1214.

Simmons, C. S. 1986. A generalization of one-dimensional solute transport: a stochastic-convective flow conceptualization, Proceedings 6th Annual AGU Hydrology Days, Fort Collins, CO, April 1986, eds. H. J. Morel-Seytoux and T. W. Warner, 139-151.

Simmons, C. S., Ginn, T. R., and Wood, B. D. 1995. Stochastic-convective transport with nonlinear reaction: mathematical framework. *Water Resour. Res.* 31:2674-2688.

Sposito, G. 1993. *Chemical Equilibria and Kinetics in Soils.* Oxford University Press, New York.

Steefel, C. I. and MacQuarrie, K. T. B. 1996. Approaches to modeling of reactive transport in porous media, in *Reactive Transport in Porous Media; Reviews in Mineralogy, Volume 34.* P. H. Ribbe, series editor. Mineralogical Society of America, Washington, D.C.

Sudicky, E. A., Schellenberg, S. L., and MacQuarrie, K. T. B. 1990. Assessment of the behavior of conservative and biodegradable solutes in heterogeneous porous media, in *Dynamics of Fluids in Hierarchical Porous Media,* J. H. Cushman, ed. Academic, New York, 429-462.

Thiele, M. R., Blunt, M. J., and Orr, F. M., Jr. 1995a. Modeling flow in heterogeneous media using streamtubes I. Miscible and immiscible displacements. *In Situ* 19:299-339.

Thiele, M. R., Blunt, M. J., and Orr, F. M., Jr. 1995b. Modeling flow in heterogeneous media using streamtubes II. Compositional displacements. *In Situ* 19:367-391.

Tien, C., Turian, R. M., Pandse, H. 1979. Simulation of the dynamics of deep bed filters, *AiChE J.* 25:385-395.

Valocchi, A. J. and Malmstead, M. 1992. Accuracy of operator splitting for advection-convection-reaction problems. *Water Resour. Res.* 28(5):1471-1476.

Widdowson, M. A., Molz, F. J., and Benefield, L. D. 1988. A numerical transport model for oxygen- and nitrate-based respiration linked to substrate and nutrient availability in porous media. *Water Resour. Res.* 24(9):1553-1565.

Wise, W. R. and Charbeneau, R. J. 1994. In situ estimation of transport parameters: a field demonstration. *Groundwater.* 32, 420-430 May-June 1994.

Wood, W. W., Kramer, T. P., and Hearn, P. P. Jr. 1990. Intergranular diffusion: an important mechanism influencing solute transport in clastic aquifers? *Science.* 247:1569-1572.

Yeh, G. T. and Tripathi, V. S. 1989. A critical evaluation of recent developments in hydrological transport models of reactive multichemical components. *Water Resour. Res.* 25(1):93-108.

Glossary

Heterogeneity   The occurrence of variations in the value of a field property at different points in space within that field; i.e., spatial variability, often represented in a nondeterministic fashion. Heterogeneity may occur in properties that are of a chemical, physical, or biological nature.

Nonequilibrium Reaction   Reaction of a chemical with its environment that is slow relative to other transport mechanisms such as convection.

Reactive Contaminant   A chemical that interacts with its environment in a nonconservative manner.

Scale of Observation   Window in space and time through which the environment or some property of the environment is viewed (or measured).

Sorption   Term used to indicate partitioning of a chemical between the mobile and immobile phases. Immobile phases include solids, dead-end pores, recirculation, etc.

Streamtube   Conceptual solid paths demarked by flow streamlines along which contaminants travel in solution. The solute flux is constant within the streamtube and no transverse mixing or other mass transfer occurs between streamtubes.

16

Contaminant Transport in the Unsaturated Zone

Stergios Dendrou and
Basile Dendrou

ZEi Engineering Inc.

16.1 Introduction

This chapter deals with the migration of chemicals (often contaminants) through the unsaturated zone. The unsaturated zone (also known as the vadose zone) is the soil horizon between the water table and the ground surface. In that space, water is below saturation levels, and the bulk of the porous medium voids is usually occupied by air. Other chemicals, for example organic contaminants, may appear in bulk mass (free product). Therefore, in general we are in the presence of three phases: the water phase, the air phase, and the chemical phase. These are in contact, in a state of dynamic equilibrium in response to external and internal stimuli: release of chemicals at the ground surface, dissolution and infiltration from runoff/precipitation, or creation of by-products from chemical reactions. Moreover, at the contact surface each phase dissolves into the other so that air, water, and chemicals also exist in the other phases in dissolved form. To complicate matters further, all phases interact with the solid matrix of the soil (minerals and organic matter), which itself may undergo deformation. All these phenomena take place at different time and space (geometric) scales. The dynamic equilibrium described above is quantified by a series of balance laws: for the mass of the different phases and species, the system energy, and the

chemical and biological potential. This general description in fact fits most groundwater conditions, including that of the saturated zone where the air phase is at zero saturation and the water phase is at saturation (after all, nature seldom recognizes nor follows man's classifications). What, then, is the specific interest in studying the unsaturated zone; and, if necessary, how does one deal with the encountered complexity? These are the questions that we address in this chapter.

Interest in the unsaturated zone is often indirect: you can draw water from a well in the saturated zone only, but the saturated zone is for the most part replenished from recharge through the vadose zone. Therefore, contamination reaches an aquifer by seepage through the unsaturated zone which acts as a contamination "source." Early engineering applications treated the vadose zone as a transmission zone. For example, the model PRZM (EPA, Pesticide Root Zone Model) predicts the vertical profile resulting from the movement of pesticides from the ground surface to the water table. This is accomplished by a one-dimensional (vertical) macroscale mass balance between runoff, erosion, plant uptake, leaching, decay, foliar washoff and volatilization. Bulk medium properties are used to describe the unsaturated medium, e.g., wilting point, without accounting for the wetting front dynamics or hysteretic wetting/drying soil conditions. The model EIS/PRZM extended the approach by integrating the vertical profiles over an areal distribution of agricultural land use and linking these pesticide loads to a saturated zone groundwater migration simulator. A suite of similar models termed "screening models" has been developed to address narrow environmental management problems (see for example, Bedient et. al., 1994, Chapter 9). However, more than a contamination source or a transmission zone, the unsaturated zone is the battle ground where the environmental pollution problem will be finally resolved. As *natural attenuation* and the *risk-based approach* (AFCEE, 1995, 1996) become legitimate parts of the environmental mitigation arsenal, it is becoming imperative to have a detailed understanding of the degradation processes that take place in the unsaturated zone so as to prescribe defensible levels of treatment of the contamination source. Therefore, we begin this chapter by describing the unsaturated zone from the soil, water, and air perspectives, and then succinctly developing the full multiphase, multispecies set of equations. This is followed by a presentation of the most important among the physical properties used to describe the unsaturated medium and a typical case study.

16.1.1 The Soil Framework from the Solid Phase Perspective

Not only is the unsaturated zone a highly active system of mineral particles and organic matter, water, and air, but a dynamic chemical and biochemical system as well. The soil constituents exist in solid, liquid, and gas states. The solid phase is represented by mineral particles, amorphous material, oxides, and hydrous, humic substances and carbonates, as illustrated in Figure 16.1. The variability of all these components, the chemistry of the pore water and specific interaction of the three phases, contribute significantly to the highly nonuniform composition of the soil material, its properties and characteristics, and the mechanisms which control transport of contaminants in the soil. On a smaller scale, the porous nature of soils determines the habitat for microorganisms that are the principal agents of decomposition of wastes in soils. On a still smaller scale are the various chemical processes such as the adsorption and desorption of molecules on clay particle surfaces and the control of acidity by buffering in the soil. All these mechanisms affect the migration of chemicals in the unsaturated zone which can take place in an aqueous (usually dissolved in water), nonaqueous (dense liquids), and gas state, in a variety of pathways that need to be clearly identified for an effective assessment of the risk of exposure. A systematic approach for quantifying the most relevant processes that contribute to the fate and transport of these chemical compounds in the macroscopic world must start with a detailed description of the soil composition and the various interacting processes quantifying the soil–water, soil–contaminant, contaminant–air, and contaminant–microorganism interactions.

Soil is a mixture of inorganic and organic material with variable amounts of water and air. The inorganic material is in the form of mineral particles derived from the soil's *parent* material, which is the rock or sediment on which the soil has formed. This may be the material underlying the soil, or may
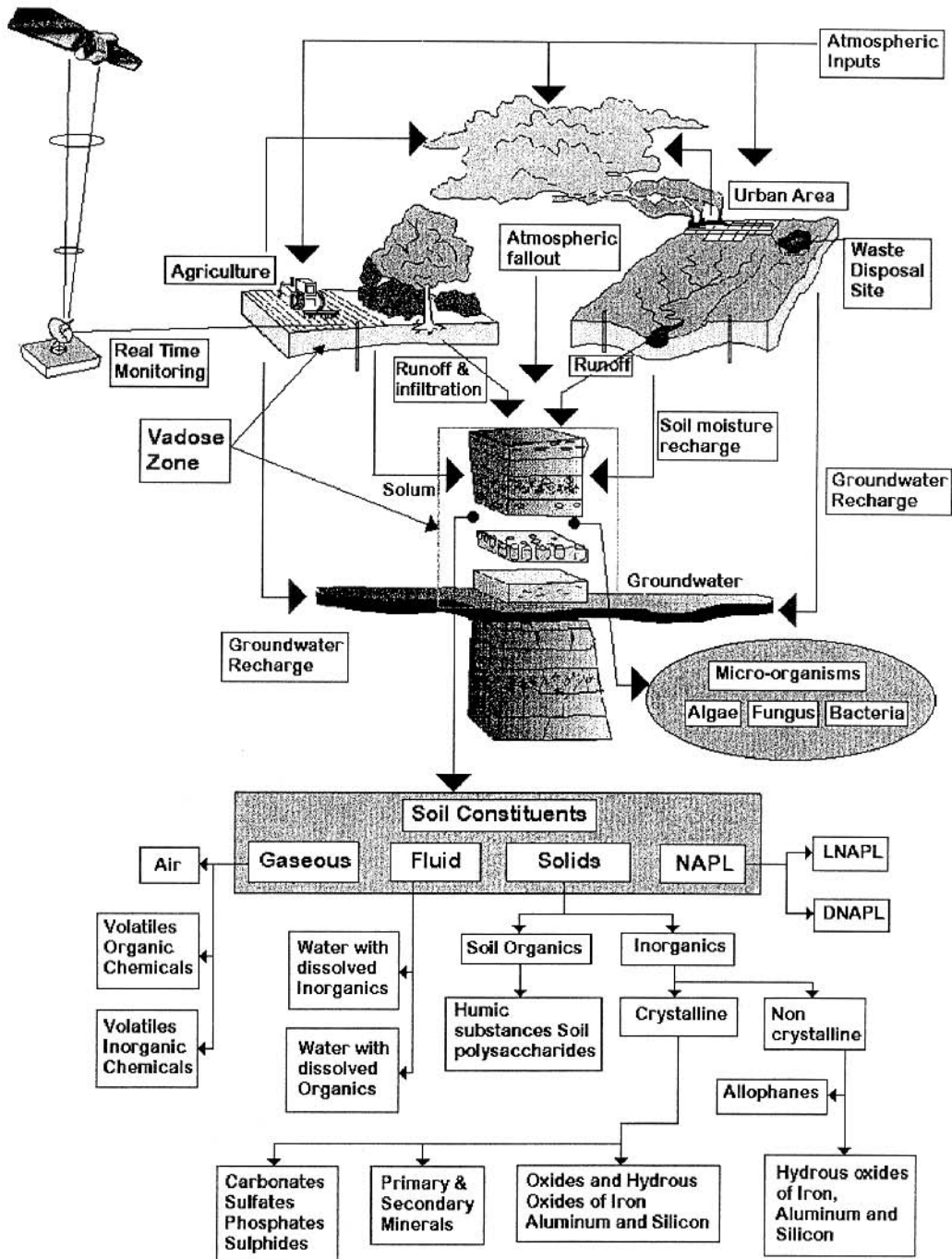
FIGURE 16.1 Principal contaminant sources and pathways in the soil chemical system.

have been brought from elsewhere by wind, water, ice, or man. The weathering of rocks and the processes of soil formation alter rock minerals so that soil minerals are partly inherited from the parent material and partly developed in the soil. Of particular significance is the development of altered or new minerals less than 2 μm in size, known as the *clay-sized fraction.* This together with associated organic materials known as soil *organic matter,* controls to a large extent the physical and chemical properties of soils.

Soil Particles. Without magnification the smallest particles that can be seen are about 0.5 mm in size. With a hand lens of 10× magnification, particles down to about 0.05 mm or 50 μm can be seen. These and smaller particles can be observed if soil is placed on a microscope slide, water added, and the soil broken up using a needle. When viewed using a microscope with 100× magnification, particles down to about 5 μm can be seen. Particles smaller than this give a cloudy appearance to the water. The larger particles have a crystalline appearance, and are usually the mineral quartz (silica) in soils of temperate regions. They are normally either angular or roughly rounded, but occasionally more perfect crystals can be seen. Many of the particles appear to be "dirty," having brown materials on their surfaces. This is an important characteristic of soils, where there is a close association between the small mineral particles, the organic matter and the surfaces of larger mineral particles. Note that the term *mineral* can be used in two ways. First, as in geology where it applies to a specific crystalline inorganic particle, and second for the inorganic nutrients in soils which are taken up by plants, such as nitrate, calcium, etc.

Soil Texture and Particle Size. The size of the mineral particles greatly affect the physical properties of soils such as drainage, the ability to hold water for plant use and the ease with which they can be cultivated. They also affect the chemical characteristics of soils because of the special properties of the surfaces of the very small particles. Thus we need a system for classifying particle size and for describing quantitatively the size distribution of these particles in soils.

No natural classification of particle sizes exists; the classifications are based on the contribution which particles of different sizes make to the physical and chemical properties of the soil. Table 16.1 gives the most commonly used system where names apply to particle-size classes and not to the types of minerals present in each class, although certain minerals may predominate in the size groups. For example, the sand-sized particles are commonly quartz and the clay-sized particles are often clay minerals. (The term *clay* also has a third more popular meaning, i.e., any fine-grained plastic material.) The analysis of particles in terms of size classes (the *particle-size distribution)* initially separates the stones from the oven-dry soil. The remaining *fine earth* composed of particles less than 2 mm in diameter is divided in terms of a percentage by mass of each size fraction in the fine earth.

TABLE 16.1   Particle-Size Classes

Soil Type	Particle Classes	Particle Size (mm)	Reference
Stones		>2	
Fine earth	Coarse sand	2–0.2 (2000–200 μm)	U.S. Department of Agriculture
	Fine sand	0.2–0.05 (200–50 μm)[1]	U.S. Department of Agriculture
	Silt	0.05–0.002 (50–2 μm)	U.S. Department of Agriculture
	Clay	<0.002 (<2 μm)	U.S. Department of Agriculture

[1] The U.S. Department of Agriculture uses 50 μm as the limit between fine sand and silt.

For convenience, soils are also classified according to their *texture,* each textural class having a given range of particle-size distribution. Thus there are small but generally unimportant variations in soil properties within each class, and significant differences between classes. The limits chosen do not indicate sudden changes in properties as composition changes.

16.1.2 The Soil Framework from the Air Phase Perspective (Effect of Pore Structure)

Particles in soil are in contact with one another. Spaces or pores which contain air and water are present between the particles. During soil formation large spaces are formed by physical and biological action, and so soils are even more porous than would be expected from a random packing of particles. This is an important soil characteristic: soils are described as being unconsolidated, which implies that they can be consolidated, i.e., pores may be reduced in volume when a force is applied. Soils are characteristically less consolidated than the parent materials from which they are formed. Soil *porosity,* or ratio of volume

of voids to total volume is characteristic of the *soil structure* property, which also includes the arrangement of particles into *aggregates* (groups of particles), and the size, shape, and distribution of pores both within and between the aggregates. It is of vital importance in the ability of soils to support plant, animal, and microbial life. The spaces hold water, allow drainage, allow entry of O_2 and removal of CO_2 from the soil, allow roots to penetrate, and are indirectly responsible for modifying the mechanical properties of soils.

Organic matter and the associated biological activity in soils are of major importance in maintaining soil porosity. Typical porosity values are given in Table 16.2. High porosities are also found in soils of the humid tropics containing appreciable contents of iron oxide or allophane because of the presence of pores within the mineral particles. Sandy soils, particularly fine sands, tend to contain smaller amounts of organic matter than heavier-textured soils, structure is less stable and, even though they are easily loosened by cultivation, natural consolidation rapidly follows when they are re-wetted.

TABLE 16.2 Typical Values of Porosity and Density of Soils

Soil Type	Porosity ($cm^3\ cm^{-3}$)	Particle Density ($g\ cm^{-3}$)	Dry Bulk Density ($g\ cm^{-3}$)
Grassland and woodland	0.67–0.50	2.40	0.8–1.2
Sub–soils	0.43–0.32	2.65	1.5–1.8
Light texture soil	0.46–0.35	2.60	1.4–1.7
Heavy texture–medium soil	0.69–0.46	2.60	0.8–1.4

Observation of soil structure in the field according to the appearance of aggregates and to the pores which are visible without magnification or with 10× magnification using a hand lens (larger than about 50 μm), allows us to detect the *macrostructure*. The corresponding pores are termed *macropores*. Total porosity minus the macroporosity gives the volume of smaller, nonvisible pores in the soil. These are *micropores*, less than about 50 μm in size. Large thin sections of soil, although difficult to make, allow more critical observation of macropores and can be combined with point-counting methods to give quantitative estimates of macroporosity.

Classification of Pores According to Function. The classification of pores in visible or nonvisible, often using 50 μm as the size limit, is convenient but it is also related to the drainage properties of the soil: the macropores are those which allow rapid drainage of water after heavy rainfall or irrigation, and once these pores are emptied, drainage becomes very slow. The soil is then said to be *at field* capacity. Further removal of water by evaporation from the soil surface or transpiration from leaves removes water from micropores, and a second critical state in relation to soil water is reached when plants permanently wilt, termed the *wilting point.* Here water is held only in pores less than about 0.2 μm in size. The amount of water held between these two states is termed the *available water capacity.* This leads to the classes of soil pores shown in Table 16.3.

TABLE 16.3 Soil Pores by Size and Function

Soil Pore Type	Pore Function	Pore Size (μm)
Macropore – Main transport pathway	Drainage after saturation aeration (movement of O_2 and CO_2) when the soil is at field capacity. Both aeration and drainage require pores to be prismatic breaking to continuous in the vertical blocky direction. Root penetration: for many arable crops pores >0.2 mm are required.	>50
Micropore – water storage pore	Store water available for plant use.	50–0.2
Micropore – water residual pore	Hold water so strongly that it is not available to plants. This water is mostly associated with clay-sized particles and controls the mechanical strength of the soils to a large extent.	<0.2

Pore size and pore function (macropore, etc.) allow the determination of the volume of macropores (and thus the volume of micropores) from the total porosity and the water content of the soil at field capacity. Similarly, the water content at the wilting point gives the volume of residual pores, which subtracted from the micropore volume gives the storage pore volume. Typical values for these three pore classes are shown in Table 16.4.

TABLE 16.4 Typical Volumes of Pores

| Pore Type | Texture | | |
	Heavy	Medium	Light
Macropores	0.05–0.15	0.10–0.15	0.20–0.30
Storage pores	0.15–0.20	0.20–0.25	0.05–0.15
Residual pores	0.25–0.35	0.15–0.20	0.05–0.10
Total porosity	0.50–0.70	0.45–0.55	0.35–0.45

There are three critical conditions: (1) if transmission pore volume is below about 0.1 cm^3 cm^{-3} then drainage problems may occur; (2) if storage pore volume is below about 0.15 cm^3 cm^{-3}, then water availability is likely to be restricted; and, (3) residual pore volumes above about 0.2 cm^3 cm^{-3} suggest that the soil may have difficult mechanical properties, being plastic and sticky when wet, and hard when dry.

The term *microstructure* is used for the arrangement of particles and pores that can be seen using a microscope. Micropores and the arrangement of clay-sized particles are important aspects of microstructure. Residual pores dominate the pore system in domains of clay-sized material. Storage and residual pores are developed by structure-forming processes. A sample of clay parent material taken from below the depth of drying in a soil profile may contain only fine storage and residual pores. On drying, shrinkage occurs which further decreases the size of these pores, but also introduces cracks which are large transmission pores. In contrast, the surface horizon may have a distribution of pore size which gives excellent properties in terms of physical conditions and biological activity, showing the effectiveness of the structure-forming processes.

16.1.3 The Soil Framework from the Chemical Perspective

The chemical composition of the soil is a significant factor characterizing the contaminant–soil interaction. An overview of the mineral composition of soils at different scales follows.

At the Macropore Scale (Sand and Silt). These fractions are dominated by resistant inherited minerals. In temperate regions much quartz may be present or, if the soil is formed on limestone, calcite and dolomite may dominate. In tropical regions iron and aluminum oxides and hydroxides dominate because more intense weathering may have dissolved away the quartz. In arid regions, inherited minerals dominate the soil because of the lack of water necessary for chemical and biological processes: the mineral particles are simply broken fragments of rock. Some pedogenic minerals such as gypsum and calcite are formed because, although they are relatively soluble, they crystallize on evaporation of solutions produced by the low intensity of chemical weathering.

At the Micropore Scale (Clays). The mineral components in this fraction are primarily clay minerals, sesquioxides, and amorphous minerals in association with humus. Because of their small particle size and their electrical charge they have physical and chemical characteristics described as *colloidal,* and the clay-sized fraction is sometimes termed the *colloidal fraction.* These characteristics include the ability to absorb water with resultant swelling and the development of plasticity. They may also disperse in water as a result of the development of repulsive forces between the particles. This fraction has electrical charge which results in the particle surfaces holding ions, including important plant nutrients. In soils formed on limestone parent materials, the clay-sized fraction may include calcite as a major component.
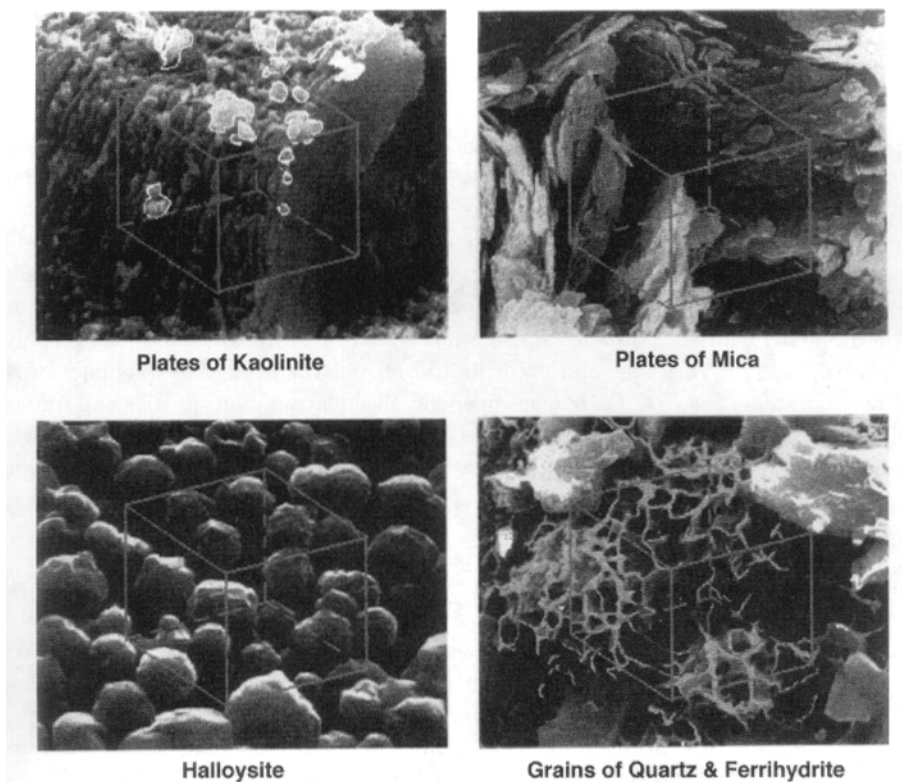
FIGURE 16.2 Geometric configuration of clay minerals and sesquioxides. Each cube represents $10 \times 10 \times 10$ µm.

The Clay Minerals. These are crystalline hydroxy-silicates containing aluminum, iron, and magnesium with other metallic elements present in small amounts. They can be either inherited minerals if the soil is formed on a sedimentary material rich in clay, or pedogenic minerals when produced by weathering. They are generally plate-shaped, ranging in size down from about 2 µm (Figure 16.2). Clay minerals vary in crystalline structure, electrical charge, surface area, and swelling characteristics. Several types are permanently negatively charged and hold *exchangeable cations* including potassium, calcium, and magnesium. Soil clay minerals are not perfect crystals as are those found in rocks, and are often present as mixtures of two or more clay minerals with sesquioxides on their surfaces. Thus although these minerals can normally be identified, the quantities present can only be determined approximately (+5% of the clay-sized fraction).

Allophane is an amorphous hydrous aluminum silicate, formed from volcanic ash and in other rapidly weathering systems, where it may be the dominating mineral. It has a very large surface area, carrying both positive and negative charge and is thus similar to the sesquioxides in some respects.

The Sesquioxides. These are also called iron and aluminum oxide clays, but this use of the term clay adds a fourth meaning to the word and for this reason sesquioxide or hydrous *oxide* is preferred. Various forms are produced by the weathering of silicates and other minerals, all being pedogenic minerals. They are mainly oxides or hydroxides of iron and aluminum including gibbsite, $Al(OH)_3$, goethite, $FeOOH$, and hematite, Fe_2O_3 with the poorly ordered ferrihydrite, $Fe_2O_3 \cdot nH_2O$, merging into a range of amorphous materials. They occur as discrete particles in the clay-sized fraction, and as very small particles (about 5 nm in diameter) and amorphous coatings on the surfaces of clay minerals (Figure 16.2). The iron oxides and hydroxides have red-brown colors, which combined with the black color of humus gives the dark-brown color of soil.

Sesquioxides act as binding agents between clay mineral particles and as a source of electrical charge. This can be either positive or negative depending on soil pH, and is known as *pH-dependent charge.* Their surfaces often adsorb anions, including phosphate, nitrate, and sulfate, which are plant nutrients. The clay mineral kaolinite also has this property but to a lesser extent. Sesquioxides and kaolinite dominate the clay-sized fractions of many soils of humid tropical regions, so that these soils have charge and anion adsorption characteristics very different from those of soils of temperate regions which are dominated by clay minerals having a permanent negative charge not dependent on pH.

Soil Carbonates. The most common minerals are calcite, $CaCO_3$, magnesian calcites with up to 20% $MgCO_3$, and dolomite, $CaMg(CO_3)_2$. Calcite occurs geologically both as soft white porous chalk and as hard limestone. Carbonates may be found in all the particle-size classes of soil, being either inherited minerals in the stone, sand- and silt-sized fractions, or a mixture of inherited and pedogenic minerals in the clay-sized fraction. They may also occur in arid regions cementing large volumes of soil into hardened material termed *calcrete.* Carbonates maintain alkaline conditions in soils and influence the growth of plants through the direct effect of dissolved bicarbonate and the indirect effect of high pH on the solubility and availability of nutrients, particularly phosphorus, copper, zinc, iron, and manganese.

16.1.4 Sources of Contaminants

What constitutes a waste material (refuse) for one industry may be a resource for another through recycling. Significant quantities of wastes have found their way into the natural environment, and we are called upon to mitigate against their deleterious effects. It is important to know the composition of the contaminant because its nature will dictate the best remedial action. Some ways of categorizing waste materials are: (1) by the medium to which they are released, i.e., air, water, or land; (2) according to their physical characteristics — whether they are gaseous, liquid, or solid; (3) by the type of risk or problem they create; (4) by the type of hazard they pose — ignitable, corrosive, reactive, or toxic; or (5) according to their origin, e.g., mine tailings, municipal waste, or industrial waste.

The first and second methods of classification provide ready estimates of quantities of wastes generated, although they overlook the primary processes that produce waste products or waste streams. This is a very important factor for implementation of technology designed to obtain reduction of waste materials at the source. Categories 3 and 4 provide statutory control which reflects the regulatory philosophy of governmental waste management programs. Only the regulated (hazardous and radioactive) wastes are tracked and managed in this way.

Often the first line of defense against the contaminant's spread is the unsaturated zone, which acts as a buffer between the source of pollution and the groundwater — a recognized resource that needs protection. In many cases special structures such as clay barriers are built in this zone to contain the contaminant. The chemical nature of the contaminant directly affects its fate and effect, and consequently its treatment. Some categories of contaminants are listed below.

Organic Liquids and Dissolved Solids. To experimentally determine or assess the effects of organic liquids on the integrity of clay barrier material, it is useful to classify the liquids according to their physical and chemical properties with respect to their interactions with clay liner materials. These properties include acidity, basicity, polarity (Debye, 1929), and solubility parameters of the organic components (Hildebrand and Scott, 1950). These groups are listed in Table 16.5.

Typical Waste Products. Typical waste products such as refuse and waste streams are shown in Table 16.6. For example, typical laboratories (educational, photographic, chemical, etc.) could be the source for acids, bases, heavy metals, inorganics, ignitable wastes, and solvents. Other kinds of wastes may also be produced, e.g., paper, cloth, etc. Many of the inorganics shown in Table 16.6 are composed of chemical compounds that do not contain carbon as the principal element. Most inorganic compounds are stable, soluble in water, and not combustible. Inorganics have rapid chemical reactions, large numbers of elements, and are less complex in their structure compared to organics. A few words for each of the main categories follow.

TABLE 16.5   Chemical Properties of Contaminants

Chemical Type	Description
Acids	This grouping includes liquid organic acids or organic compounds with acidic functional groups such as phenols and carboxylic acids. Proton-donating properties of Bronsted acids give these fluids the potential to react with and dissolve soil barrier components. The general sources of liquid organic acids in municipal solid waste landfills are derived as anaerobic decomposition byproducts. These include acetic, propionic, butyric, isobutyric, and lactic acids. Anaerobic decomposition yields carboxylic acid derivatives of other organic liquids placed in the landfill.
Bases	Organic bases (such as aniline) may be liquid and have been used as solvents in several industries. However, it is not immediately obvious if organic bases are strong enough proton acceptors to dissolve soil barrier components.
Neutral polar organic liquids	Neutral pear organic liquids do not exhibit a net charge but have an asymmetrical distribution of electron density resulting in an appreciable dipole moment, an indicator of polar character (Debye, 1929). Examples of such polar compounds are alcohols, aldehydes, ketones, glycols, and alkyl halides.
Neutral nonpolar organic liquids	Although neutral nonpolar organic liquid do not exhibit net charge, they may possess small dipole moments. The liquids have low water solubilities and little polarity. Examples of nonpo1ar organic liquids are aliphatic and aromatic hydrocarbons.

TABLE 16.6   Typical Composition of Wastes from Representative Industries

Industry Source	Contaminant Wastes
Laboratories	Acids, bases, heavy metals, inorganics, ignitable wastes, solvents
Printing industry	Acids, bases, heavy metals, inorganic wastes, solvents, ink sludges, spent plating
Pesticide user/services	Metals, inorganic, pesticides, solvents
Construction	Acids, bases, ignitable wastes, solvents
Metal manufacture	Acids, bases, cyanide wastes, reactives, heavy metals, ignitable wastes, solvents, spent plating wastes
Formulators	Acids, bases, ignitable wastes, heavy metals, inorganics, pesticides, reactives, solvents
Chemical manufacture	Same as metal manufacture, except no plating wastes
Laundry/dry cleaning	Dry clean filtration residue, solvents

Heavy Metals. Heavy metals are elements that lose electrons easily to form positive ions. Solubility, solution equilibrium, gas phase equilibrium, hydrogen bonding, charge transfer complexes, and metal atom and metal surface behavior are all described using the chemical classification of metals based on the principle of hard and soft acids and bases (Pearson, 1968). In a typical soil system, complexes can be envisioned to be made up of ligands (cations) and central species (exchange materials). Stable complexes are formed when cations become adsorbed to soil exchange surfaces. The ability of the cation to remain adsorbed is dependent upon further exchange reactions with other cations in equilibrium with the exchange surface. Heavy metals usually conduct electric current and constitute elements with high toxicity and large density ($\rho > 5$ mg/m^3). They are generally obtained as "waste stream material" from mining, ore refining, metal producing, and electroplating industries. The major elements included in this category are: arsenic (As), cadmium (Cd), chromium (Cr), copper (Cu), Lead (Pb), mercury (Hg), nickel (Ni), and zinc (Zn). The toxicity of heavy metals varies from one element to another, and the order of decreasing toxicity of these metals is: Hg > Cd > Ni > Pb > Cr > Li. The common characteristics of heavy metals are: (1) their tendency to accumulate in soil systems at high pH; and (2) their tendency *not* to migrate from the upper soil layer. They occur in the soil solution phase in association with organic molecules as metal-organic complexes, in the suspension phase, and in soluble solid forms as precipitates. Heavy metals can be removed from soils through precipitation with hydroxide and carbonates or by coprecipitation with iron and manganese hydrous oxides. Removal efficiency tends to improve with increased pH.

Acids and Bases. They are defined as follows: *An acid is a substance which can accept an electron pair from a base; a base is a substance which can donate an electron pair.* There have been several attempts to classify acids and bases into categories according to their mutual behavior, e.g., the classification of Lewis

acids and bases given by Pearson (1963) as *hard or soft* if they demonstrate some of the following properties:

- Hard acids are usually small in size with a high positive charge, high electronegativity, low polarity, and do not contain non-shared pairs of electrons in their valence shells.
- Soft acids are generally large in size with a low positive charge, low electronegativity, high polarity, and contain non-shared pairs of electrons in their valence shells.
- Hard bases usually have high electronegativity, low polarity, and are hard to oxidize.
- Soft bases usually have low electronegativity, high polarity, and are easy to oxidize.

The terms *hard* and *soft* are relative categories. Borderline or intermediate species exist which behave as relative hard or soft species depending upon the solution environment. A list of several hard, soft, and intermediate acids and bases can be found in Pearson (1963) and Lester (1987) for example. With a knowledge of the relative hardness or softness of various species in an aqueous solute-solvent system, the stability of complexes formed between species can be predicted by the hard-soft-acid-base (HSAB) principle. This principle essentially states that: *hard acids tend to associate with hard bases and soft acids with soft bases.*

Salts. Salts are substances that upon dissolution yield ions other than hydrogen ions (H^+) or hydroxyl ions (OH^-). They are formed by replacement of acidic hydrogen in an acid by a metal (e.g., producing sodium sulfate $Na_2SO_4^-$) or a basic group (e.g., producing ammonium nitrate, $NH_4 NO_3$). Most types of salts have an ionic structure. Since there is no ion that is common to all salts, it is difficult to establish common properties between them. For instance, some salts test sour, others bitter, salty, or metallic. Some salts are oxidizing agents (e.g., chlorides, nitrates), others are reducing agents (e.g., sulfates). Salts may vary in their action toward indicators, some being acid, some alkaline, and some neutral. Nevertheless, some salts may share their physical properties if they contain the same metal (e.g., NaCl and Na_2SO_4) or the same acid (e.g., NaCl and KCl). Salts occur naturally in many soils and are a common constituent in most contaminant streams. Salt inputs to the soil may occur from addition of waste, fertilizer applications, precipitation, and irrigation. The predominant salts in most soils include sodium, calcium, magnesium, chloride, and sulfate. Minor amounts of potassium, bicarbonate, carbonate, and nitrate may also be present.

Cyanides (CN). Cyanides can exist as an anion, CN^-, for pH > 8 or as an acid, HCN, for pH < 8. It is an extremely poisonous substance with high melting point. The CN^- (aqueous) ion is a powerful ligand. It can join to metal ions forming stable complexes (metallic cyanide). In many cases the ion is so solid that the properties of the free bonds cannot be detected in solution. Some metallic cyanides are ionic in the solid state, while others are covalent. The majority are insoluble in water but dissolve in CN^- (aqueous). Typical types of alkali metallic cyanides include sodium cyanide (NaCN), potassium cyanide (KCN), and iron cyanide ($Fe(CN)^{4/6}$). The cyanide ion, CN^-, combines with hydrogen ion, H+, forming a weak acid, HCN. The molecule of cyanide acid is linear and it is isoelectric with acetylene, possessing a triple bond: H–C=N. Unlike most other toxic anions, cyanides can be degraded by biological reactions. A soil organism degrades CN^-. This reaction occurs despite the toxicity established to most organisms.

16.1.5 Chemical Reactive Groups

In contaminants, the nature of the functional groups which form the compound will influence the characteristics of the compound, and its ability to bind with the soil constituents. For example, depending on how they are placed, the functional groups will influence the characteristics of organic compounds, and will thus contribute greatly in the determination of the mechanisms of accumulation, persistence, and fate of these compounds in soil.

The two main compound functional groups are: (1) functional groups having a C–O bond, e.g., carboxyl, carbonyl, hydroxyl, methoxyl, and ester groups, and (2) nitrogen-bonding functional groups, e.g., amine and nitrite groups. The acquisition of a positive or negative charge is through the process of dissociation of H+ from or onto the functional groups, dependent on the dissociation constant of each

functional group and the pH. The compounds can thus be viewed as fixed or variable charged compounds, or a mixture of both. More details about the various groups can be found under *Hydroxyl Group* in Sposito (1989) and Sawyer and McCarty (1978); under *Carbonyl Group* in Geissman (1977); under *Phenolic Group*, including pesticides such as 2,4-D and cyclic alcohols such as creosols, nahpthols, quionones, nitrophenol, and pentachloropehnols in Overcash and Pal (1979).

16.2 Conceptual Framework (Theory of Interactions)

Pollutants migrate in the subsurface environment primarily because of the presence and movement of groundwater. Their constituent species dissolve in water, and the dissolved substances are carried by groundwater flow. Groundwater flow is the primary driving mechanism for contaminant migration (advection). Other mechanisms (e.g., dispersion, degradation, soil deformation) also affect the flow process itself. The full *multiphase* description of flow accounts for all phases present in the underground system, namely, water, free-product, and gas, including the presence and influence of density gradients caused by temperature differences and different rates of dissolution. The multiphase flow description is accompanied by a *multispecies* migration description that accounts for the expansion and fate of the contaminant plume. Both descriptions are given in the following section. They form the foundation of the theoretical description of the fate and migration of contaminant substances in the underground environment. In addition to these sets of equations, the behavior of the soil matrix (solid phase) is also accounted for because some remediation technologies rely more on traditional geotechnical features (landfills, containment technologies). These require an evaluation of the soil behavior, stability, and deformation in the unsaturated zone. Also, the presence of contaminants tends to alter the flow properties of the medium (reduced hydraulic conductivities due to weakening of mechanical soil properties, chemical precipitation, flocculation, and other phenomena), thereby affecting the contaminant migration. Therefore, *soil deformation* equations are also included in the multiphase multispecies formulation.

Other mechanisms and processes affecting the fate and migration of contaminants especially with regard to innovative remediation technologies are osmotic, electrokinetic, and biological. *Biodegradation* (natural and man-induced) until recently was modeled as a simple decay mechanism. As a result of recent field and laboratory investigations, biodegradation is emerging as a series of complex, physicochemical, and biological phenomena. This new theory is representative of a more general procedure for complementing the multiphase multispecies descriptions with additional physicochemical processes.

16.2.1 Soil Properties, Definitions, and Notations

Geomedia characterization is based on numerous physical, chemical, mechanical, and other properties of soil. We list a series of the most pertinent properties of both the soil medium and the fluids with their definition and description. These definitions help develop a sense of the character of the unsaturated soil behavior as used in quantitative models.

Relative Saturation of a fluid is the ratio of the volume of that fluid, divided by the total pore volume:

$$S_r = \frac{\text{fluid volume}}{\text{total pore volume}} \tag{1}$$

In the presence of multiple fluids, usually the *wetting* fluid's saturation is determined. Relative saturations of all fluids sum up to 100%.

Wettability is characterized by the angle that is formed between a drop of fluid resting on a clean surface and the surface. If the angle is larger than 90°, the fluid is called nonwetting, (Figure 16.3). Wetting angle changes as a function of fluid velocity and direction of movement. In a porous medium, the wetting fluid occupies the finer pores, e.g., the case of water in the presence of air (vadose zone).

Wetting Fluid Content is the ratio between the volume of the wetting fluid and the corresponding total soil sample volume.
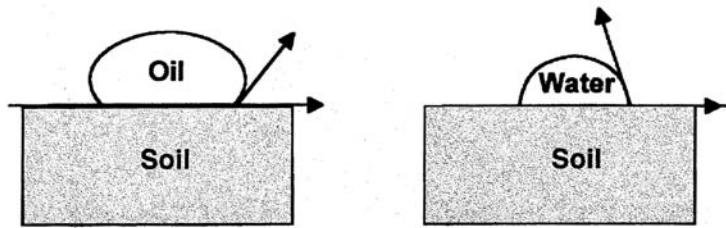
FIGURE 16.3 Fluid wettability.

$$\theta = 100\frac{V_m}{V} \tag{2}$$

or,

$$\theta = S_r \frac{n}{100} \tag{3}$$

where V_m is the wetting fluid volume
V is the soil sample volume
S_r is the relative saturation of the wetting fluid
n is the porosity

Fluid content varies from zero to the porosity. $\theta_s = n$ is called *saturation volumetric fluid content*. Not all fluid stored in the porous medium is available for extraction. Starting from a saturated state, i.e., at saturation fluid content, fluid would drain freely down to a content of retention capacity θ_r. This value is higher for fine-structured soils (clays) which have finer pores. This portion of fluid is called gravitational. Below the retention capacity, fluid can be extracted by applying a negative pressure. The finer the pore structure, the higher the suction that needs to be applied. The fluid extracted in this fashion comes from the capillary fringe. This operation stops at the irreducible fluid content θ_m. The remaining fluid content is quasi-immobile, adsorbed on the grain surface (Figure 16.4).

Capillary Pressure is the pressure difference at the interface between a nonwetting and a wetting fluid:

$$p_c = \psi - p_{nw} - p_w \tag{4}$$
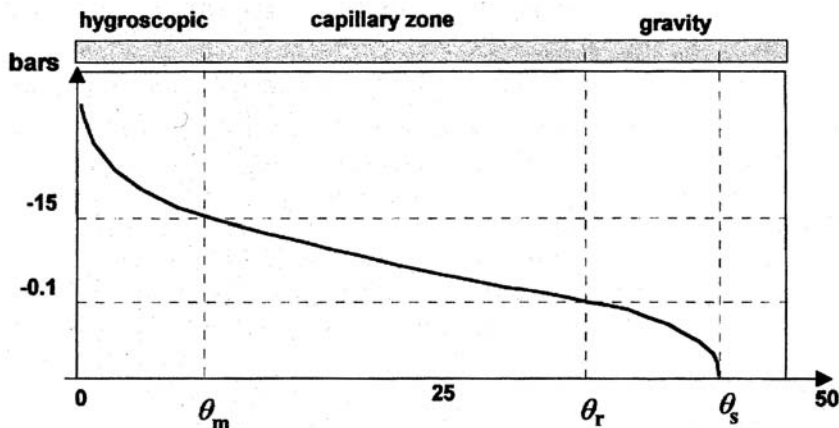
FIGURE 16.4 Fluid content, saturation, retention capacity, irreducible content.

where ψ is the capillary pressure

p_{nw} is the pressure of the nonwetting fluid

p_w is the pressure of the wetting fluid

With reference to the atmospheric pressure, the capillary pressure is negative (atmospheric pressure is applied both at the free surface of the capillary tube as well as at the surface of the saturated phase); and it is equal to the surface tension which is exercised by the surface of the tube, as follows for a tube of radius r (Figure 16.5):

$$\psi = \frac{2T\cos\alpha}{r} \qquad (5)$$

The *capillary rise* is inversely proportional to the radius and is given by:

$$h_c = \frac{\psi}{\gamma} \qquad (6)$$

where γ is the fluid specific gravity.
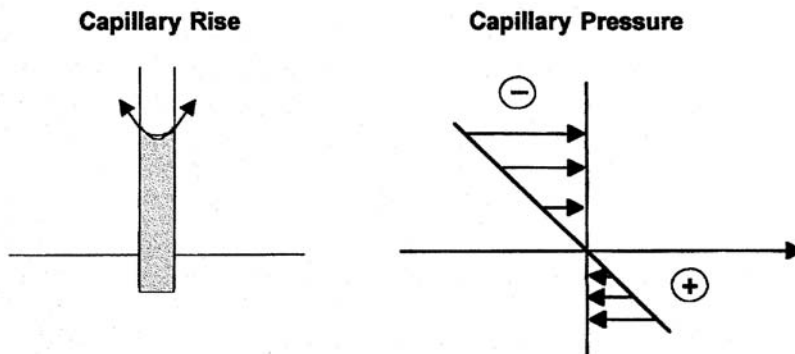
FIGURE 16.5 Capillary rise, capillary pressure.

There are four components to the *specific potential energy* of a fluid in a porous medium, namely: (1) gravity potential, (2) pressure potential, (3) matricial (capillary) potential, and (4) osmotic potential:

$$\phi_t = \phi_g + \phi_p + \phi_c + \phi_o \qquad (7)$$

where the gravity potential is given by:

$$\phi_g = \rho g \left(z_A - z_B \right) \qquad (8)$$

and is indicative of the relative position of a fluid particle with respect to a reference plane.

The pressure potential, for the saturated zone is:

$$\phi_p = p \qquad (9)$$

The capillary (or matric) potential, in the capillary fringe is:

$$\phi_c = \psi \qquad (10)$$

The osmotic potential, where density gradients of dissolved substances exist, is:

$$\phi_o = \pi \tag{11}$$

Density gradients are caused by chemical reactions, thermal, magnetic, or electric fields, and their combination. These potentials are written in *water height equivalent* by dividing by the term ρg where ρ is water density. Capillary potential and osmotic potential are both negative, designated as suction or tension.

Total Piezometric Head.

$$H = \frac{\phi_g + \phi_p + \phi_c + \phi_o}{\rho g} = z + \frac{p}{\rho g} + \frac{\psi}{\rho g} + \frac{\pi}{\rho g} \tag{12}$$

where z is elevation above datum at fluid particle
p is fluid pressure
ρ is fluid density
g is gravitational acceleration
ψ is capillary pressure
π is osmotic pressure

Typically, the relative saturation of the aeration zone in a soil profile varies from dry season to wet season to form the soil hydraulic profiles shown in Figure 16.6. These seasonal variations disappear below a certain depth. Dry–wet cycles permanently modify some soil properties as described by *hysteretic* models described later in the chapter.
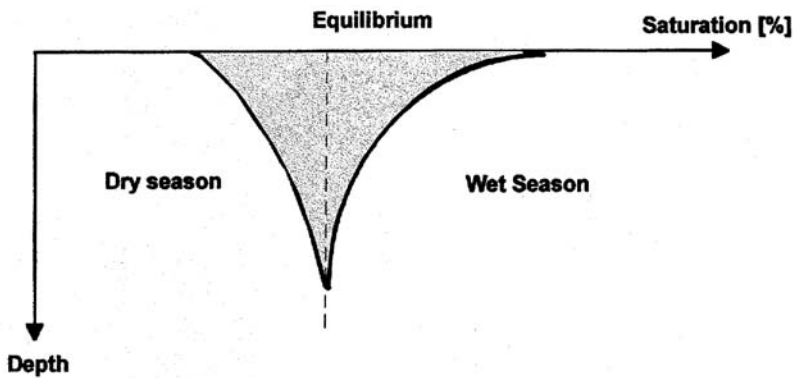
FIGURE 16.6 Soil hydraulic profile.

Permeability (Intrinsic), and Fluid Conductivity. Fluid conductivity is the coefficient of proportionality in Darcy's law, measured in units of velocity (length/time), and it depends both on the properties of the fluid (viscosity, density), and the porous medium (pore size distribution, tortuosity):

$$v = K \cdot \nabla h \text{ (Darcy's law)} \tag{13}$$

where K is fluid conductivity (length/time).

The effect of porous medium and fluid properties is functionally separated as follows (Bear, 1972):
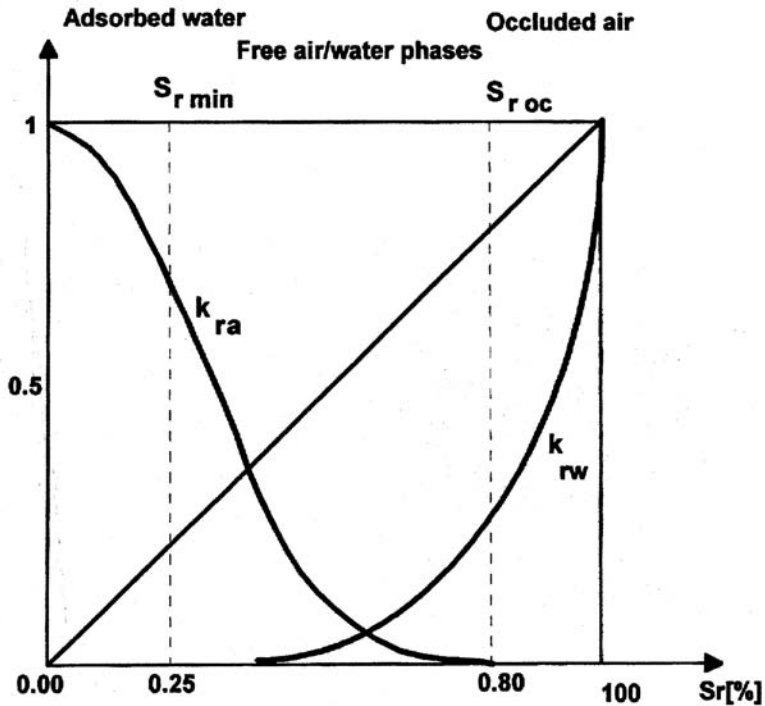
FIGURE 16.7 Relative permeabilities as a function of relative saturation.

$$K = f_1\left(\frac{\rho g}{\mu}\right) f_2\left(d^2\right)$$ (14)

where μ is the fluid viscosity

       d is a characteristic soil grain size, e.g., harmonic mean diameter, or

       d_{10} the diameter of particles below 10% (by weight) of the soil

       $f_2 = k = Cd^2$ is the *intrinsic permeability* (length^2)

Relative Permeability is the ratio of the unsaturated to the saturated permeability (or hydraulic conductivity).

$$k_r = \frac{k}{k_s}$$ (15)

Relative permeabilities of a wetting fluid and a nonwetting fluid in a multiphase system vary inversely as a function of relative saturation, as illustrated in Figure 16.7 for the case of water and air. Above the S_{roc} value (relative saturation of occlusion), all air is occluded in the water phase and there is no free flowing air. Below the S_{min} value, on the other hand, all water is adsorbed on the grain surfaces and the soil can be considered dry.

16.2.2 Multiphase Multispecies Flow and Migration Equations

The flow and migration equations consist of imposing the continuity conditions (mass balance) for each phase, water, free-product, and gas separately, where the mass flux is expressed in terms of pressure via

a generalized Darcy's law (macroscale representation). The mass content of each phase in the control volume is expressed in terms of the medium porosity and the relative saturation. Time changes in mass content are attributed to Darcian exchanges with the surrounding medium, density variations (gradients), and sinks or sources.

Continuity Equations (Macroscale Mass and Energy Balance).
Water Phase.

$$\frac{\partial(nS_w)}{\partial t} = \frac{\partial}{\partial x_i}\left[K_{w_{ij}}\left\{\frac{\partial h_w}{\partial x_j} + \frac{\partial(\rho_{rw}z)}{\partial x_j}\right\}\right] + \frac{R_w}{\rho_w} \tag{16}$$

Chemical Phase (Free-Product).

$$\frac{\partial(nS_O)}{\partial t} = \frac{\partial}{\partial x_i}\left[K_{O_{ij}}\left\{\frac{\partial h_O}{\partial x_j} + \frac{\partial(\rho_{rO}z)}{\partial x_j}\right\}\right] + \frac{R_O}{\rho_O} \tag{17}$$

Gas Phase.

$$\frac{\partial(n\rho_a S_a)}{\partial t} = \frac{\partial}{\partial x_i}\left[\rho_a K_{a_{ij}}\left\{\frac{\partial h_a}{\partial x_j} + \frac{\partial(\rho_{ra}z)}{\partial x_j}\right\}\right] + R_a \tag{18}$$

Alternatively, the gas phase continuity equation can be written to include the quantity of gas dissolved in the fluid phase according to Henry's law, as follows:

$$\frac{\partial(n\rho_a(S_a + HS_w))}{\partial t} = \frac{\partial}{\partial x_i}\left[\rho_a K_{a_{ij}}\left\{\frac{\partial h_a}{\partial x_j} + \frac{\partial(\rho_{ra}z)}{\partial x_j}\right\}\right] + \frac{\partial}{\partial x_i}\left[\rho_a H q_{wi}\right] + R_a \tag{19}$$

Solid Matrix.

$$\frac{\partial[\rho_s(1-n)]}{\partial t} = \frac{\partial[\rho_s q_{si}]}{\partial x_i} \tag{20}$$

where
n is porosity
S_k is k-phase saturation (related with each other and water height equivalent pressure via constitutive relations (e.g., van Genuchten, 1980)
x_j is spatial coordinate
R_k is mass flux (temperature or density *gradients*, or boundary conditions)
t is time
$K_{k_{ij}}$ is k-phase conductivity tensor
$h_k = P_k/g\rho_w^*$ is the k-phase water height equivalent pressure
P_k is k-phase pressure
z is elevation
ρ_w^* is pure water density
$\rho_{rk} = \rho_k/\rho_w^*$ is k-phase specific gravity
ρ_s is solid (grain) density
q_{si} is the Darcy solid matrix velocity in i-direction

$$q_{ki} \qquad \text{is k-phase Darcy velocity in i-direction}$$

q_{ki} is k-phase Darcy velocity in i-direction
H is Henry's constant

Generalized Darcy's Law.

$$q_{ki} = -K_{k_{ij}} \cdot grad\{h_k + \rho_{rk}z\} \tag{21}$$

Dissolved Species-a in k-Phase.

$$n\frac{\partial C_{\alpha k}S_k}{\partial t} = \frac{\partial}{\partial x_i}\left[nS_k D_{\alpha kij}\frac{\partial C_{\alpha k}}{\partial x_j}\right] - \frac{\partial\left(C_{\alpha k}q_{ki}\right)}{\partial x_i} + R_{\alpha_i\alpha_jk} - \mu_{\alpha k}C_{\alpha k} \tag{22}$$

where $C_{\alpha k}$ is the concentration of reactive *a-species* in *k-phase* (mass of *a-species* per phase volume)

    $D_{\alpha k_{ij}}$ is the dispersion tensor

    $R_{\alpha_i\alpha_jk}$ is the mass transfer rate between species a_i and a_j per medium volume (chemical reactions, biodegradation, etc.)

    $\mu_{\alpha k}$ is the first-order decay coefficient

The first term on the right-hand side is the *dispersion* term, the second is the *advection* term, and the other two terms are the *chemical reaction* term and the *decay* term.

Heat Transport (Conduction-Convection).

$$\left[n\rho_w c_w + \left(1-n\right)\rho_s c_s\right]\frac{\partial T}{\partial t} = \frac{\partial}{\partial x_i}\left[\kappa_e\frac{\partial T}{\partial x_j}\right] - n\rho_w c_w\frac{\partial\left(Tq_{wi}\right)}{\partial x_i} \tag{23}$$

where T is temperature

    c_w is the specific heat of water (per unit volume)

    c_s is the specific heat of solids (per unit volume)

    κ_e is the effective thermal conductivity

Phase pressures are written in terms of water height equivalent. They represent the state variables (unknown) of the system which is otherwise entirely controlled by the relative saturation distributions. In general, in a fully penetrating well the free-product draining from the medium would tend to segregate at the top (floating) or at the bottom (sinking) of the water column (e.g., light or dense hydrocarbons). The actual saturation profiles as illustrated in Figure 16.8 indicate a more disperse state of material distribution.

These profiles in fact bear *memory* and show *hysteresis* depending on whether they are on a wetting (flooding) cycle or on a draining cycle. The description of these saturation profiles represents the *constitutive* (empirical) relations characterizing the soil medium and the imbibing fluid, and take the general form shown below. Various empirical forms exist (for example, van Genuchten, 1980), relating water saturation to total liquid saturation, field capacity, capillary curve parameters, and other fluid-dependent scaling factors. Examples of hysteretic behavior between relative saturation and capillary pressure laws are given in the next section (Figure 16.9) (Morrow, 1970).

Constitutive Relations.

$$S_w = f\left(h_a, h_w, \text{hysteresis}\right) \tag{24}$$

MULTIPHASE FLOW IN GEOMEDIA
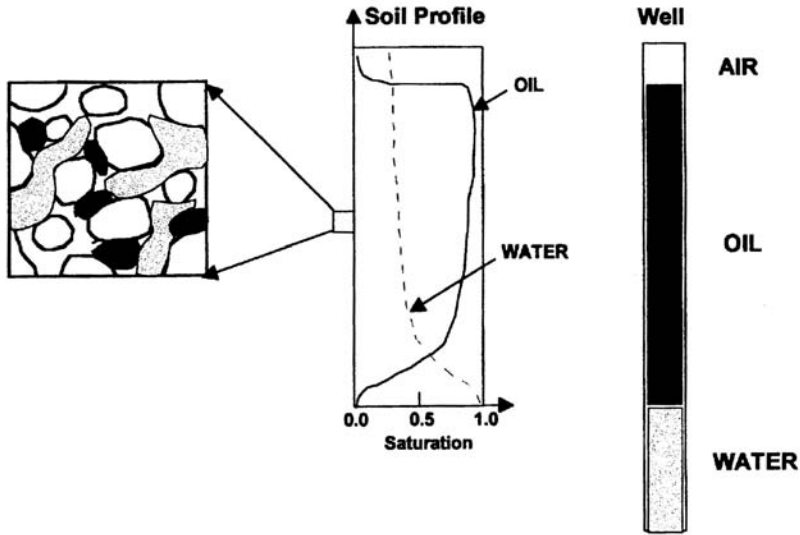
FIGURE 16.8 Relative saturation profiles in multiphase flow.
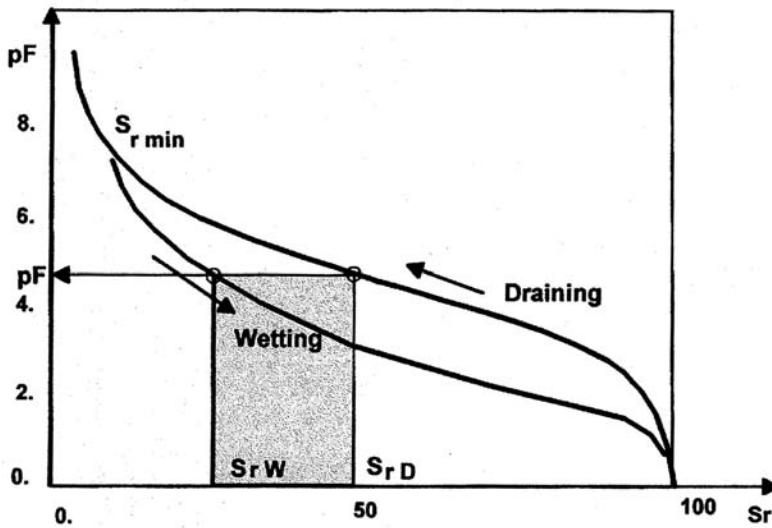
FIGURE 16.9 Hysteresis between relative saturation and capillary pressure.

The migration equations impose mass balance conditions on all dissolved species. The mass distribution of the dissolved species is described in terms of concentrations of each species for each separate phase. The net time variation of constituent mass is attributed to advection exchanges with the surrounding medium, dispersion, chemical/ biological reactions, and sink/sources.

16.2.3 Soil Matrix Deformation (Saturated-Unsaturated Consolidation)

The soil matrix in geomedia undergoes deformation (strain) in response to loads (from conventional or earth structures, landfills, excavations) and to changes in the flow and chemical regime in the pore structure (drained conditions, chemical precipitation, reactions). What is important from the point of

view of contaminant migration is that *this deformation will also have an effect on the flow and transport properties of the medium (conductivity, storage factor, and dispersivity), and hence on the ultimate fate of contaminants.*

Soil deformation occurs in a succession of stages — initially ("instantaneously") as the medium responds to a new load; then, over time as the fluid is compressed and drained away from the zone of surcharge; and lastly, after the fluid is drained, the soil matrix undergoes further deformation, called *creep*, due to the rearrangement of clay and sandy particles. The initial deformation is similar to the deformation of solids and for soils is mathematically modeled according to the plasticity theory (ZEi, ZSOIL, 1990). Initial (plastic) deformation and creep are geotechnical aspects of the consolidation phenomena (ZEi, ZSOIL, 1990). In the following we deal primarily with *soil deformation (consolidation) caused by fluid compressibility and groundwater flow in the unsaturated zone*, and its consequence on migration properties of the medium.

In this case, in Equations (16) through (21) water density, porosity, and relative saturation are considered as time variable. In addition, for the solid phase, no loss or gain to the solid mass is allowed, except as dictated by creep of solid particles across the boundary and changes of grain density due, for example, to *chemical precipitation*. It must also be noted that air and water pressure (or water height equivalent piezometric heads), and conductivities, all vary as a function of the relative saturation and porosity (or void ratio). These laws are written for a representative volume which contains a representative mix of the phases — solid, fluid, and gas — according to the macroscopic descriptors of porosity (n) and relative saturation (S_w) (Figure 16.10). The time rate of change of these properties is provided by rheologic laws.
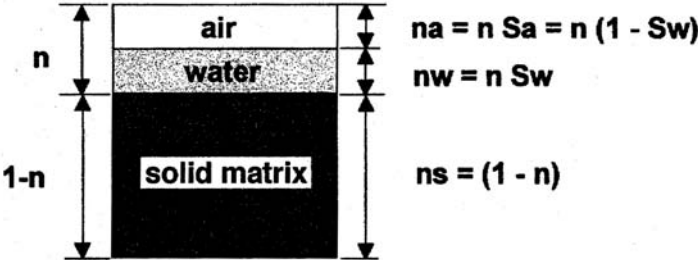
FIGURE 16.10 Macroscopic three-phase distribution: solid, fluid, and air.

Rheologic Laws of the Three Phases. These are laws relating the volumetric change of the three phases — solid, water, and gas — to empirical (experimental) properties.

Air Compressibility. The gaseous phase is comprised of air, and water and pollutant vapors. For purposes of linking the specific gravity of the gas to pressure, it can be considered to be an ideal gas governed by the following law:

$$\gamma_a = p_a / (RT) \tag{25}$$

where p_a is air pressure

γ_a is the specific gravity of air

R is the ideal gas constant

T is the absolute temperature (degrees Kelvin)

Therefore, for conditions of constant temperature, the specific gravity of gas varies with time proportionately to the time variation of pressure:

$$\partial \gamma_a \big/ \partial t = 1 \big/ \big(RT \big) \partial p_a \big/ \partial t \tag{26}$$

In terms of initial conditions, the constant (1/RT) can be written as follows:

$$1 \big/ \big(RT \big) = \gamma_{ao} \big/ p_{ao} = \gamma_{ao} m_{va} \tag{27}$$

where m_{va} is the coefficient of compressibility. Equation (26) is then written as follows:

$$\partial \gamma_a \big/ \partial t = \gamma_{ao} m_{va} \partial p_a \big/ \partial t \tag{28}$$

Water Compressibility. Water contracts or expands as pressure increases or decreases, respectively. Since mass M = $\rho_w V_w$ is conserved, this means that both the water density and volume change with pressure. The time rate of change of the water density is of interest and is defined in terms of compressibility as follows:

$$dM \big/ dt = \rho_w \partial V_w \big/ \partial t + V_w \partial \rho_w \big/ \partial t = 0 \tag{29}$$

$$\partial \rho_w \big/ \partial t = -\big(\rho_w \big/ V_w \big)\big(\partial V_w \big/ \partial t \big) \tag{30}$$

The rate of volume change at constant temperature is given in terms of the compressibility as follows:

$$\partial V_w \big/ \partial t = - V_w \beta_w \partial p \big/ \partial t \tag{31}$$

where $\beta_w = 1/K_w$ is the fluid compressibility (1/pressure)
K_W is the bulk modulus of compression (pressure)
ρ_w is water density
p is pressure
t is time

Equation (30) then becomes:

$$\partial \rho_w \big/ \partial t = \big(\beta_w \rho_w \big)\big(\partial p \big/ \partial t \big) = \big(\beta_w \rho_w^2 g \big)\big(\partial h \big/ \partial t \big) \tag{32}$$

where pressure is substituted by the hydraulic head: $h = z + p/\rho_w g$

Soil Matrix Deformation (Compressibility). The goal here is to determine the total soil volume variation as a function of effective stress, pore water pressure, and time. In its general form, the volumetric strain is written as follows (Hoshina, 1961):

$$\varepsilon_V = \Delta V \big/ V_o = \varepsilon_x + \varepsilon_y + \varepsilon_z = m_{VC} \sigma'_{oct} + m_{VG} \tau'_{oct} \tag{33}$$

$$= m_{VC} \sigma'_{oct} \left(1 + \frac{m_{VG}}{m_{VC}} \left(\frac{\tau'_{oct}}{\sigma'_{oct}} \right) \right) = m_{vc} m_G \sigma'_{oct} \tag{34}$$

where ε_V is the total volumetric strain
     σ'_{oct} is the spherical tensor
     τ'_{oct} is the deviatoric
     m_{VC} is the compressibility of the soil matrix
     m_{VG} is the deviatoric compressibility

(The effect of deformation due to *creep* is not included in this expression.)

Strain-stress relationships take the following tensorial form:

$$\{\varepsilon\} = [D]\{\sigma\} \tag{35}$$

where the coefficients of matrix D are written in terms of the modulus of elasticity and Poisson ratio v. For the case of symmetry about the vertical axis z, the coefficients of compressibility take the following form:

$$m_{VC} = \frac{1}{E_z}\left[(1 + 2\eta) - 6v_z\right] \tag{36}$$

$$m_{VG} = \frac{3\sqrt{2}\cos 2\theta}{2E_z}(1 - \eta) \tag{37}$$

where $\quad \eta = E_z/E_x$

$\qquad \theta$ is the angle between x-axis and principal direction

Effective Stress Concept. Soil deformation is caused by the portion of the stress that is carried by the soil matrix. The remaining stress is taken up by the pore fluid pressure (tensorial form):

$$\sigma' = \sigma - p \tag{38}$$

In the case of unsaturated flow, p is the equivalent pore pressure composed of the pressures carried by the water phase (p_w) and the gas phase (p_a), respectively, as follows (Bishop and Donald, 1961):

$$p = p_a - \chi(p_a - p_w) \tag{39}$$

where $\quad p \quad$ is the equivalent pore pressure

$\qquad p_a \quad$ is the gas pore pressure

$\qquad p_w \quad$ is the water pore pressure

$\qquad \chi \quad$ is the Bishop (partition) coefficient [0.-1.], depending on relative saturation

$(p_a - p_w) = \psi$ is the *capillary pressure*

Equilibrium Conditions (Stress-Strain Relationships). The soil medium is considered in quasistatic conditions (no dynamic effects, impulsive forces, explosion) and expressed in terms of the total stresses by the following partial differential equations (written for the x-direction):

$$\frac{\partial \sigma_x}{\partial x} + \frac{\partial \tau_{yx}}{\partial y} + \frac{\partial \tau_{zx}}{\partial z} = 0 \tag{40}$$

where σ and τ are the normal and shear stresses, γ_s is the solid matrix specific gravity, and z is the vertical axis.

Strain is related to displacement as follows (definitions):

$$\varepsilon_{x_i} = \partial u_i / \partial x_i$$

$$\varepsilon_{x_i x_j} = \frac{1}{2}\left(\frac{\partial u_j}{\partial x_i} + \frac{\partial u_i}{\partial x_j}\right) \tag{41}$$

where u_i is displacement in i-direction

ε_i is strain in i-direction

The equilibrium conditions integrated over the domain of interest take the following form:

$$\int_V \{\Delta\varepsilon\}^T \{\sigma\}\, dV - \int_V \{\Delta\delta\}^T \{\gamma\}\, dV - \int_S \{\Delta\delta\}^T \{F\}\, dS = 0 \tag{42}$$

where $\{\Delta\varepsilon\}$ is strain

$\{\sigma\}$ is stress

$\{\Delta\delta\}$ is displacement

$\{F\}$ is applied external forces

$\{\gamma\}$ is specific gravity

These conditions are solved for the unknown displacements $\{\Delta\delta\}$ (settlement). Resolution of the above set of laws requires that we introduce empirical relationships between relative saturation, compressibility, permeability, and specific storage, and special consideration of the scales of the various processes and their interaction.

16.3 Characterization of Soil–Water and Soil–Gas Interactions

To understand the physical behavior of water in soil we need to start from the nanoscale characterization of the electrical structure of water. The water molecule is composed of two hydrogen atoms and one oxygen. The water molecule is positively charged on one side and negatively on the other, and is thus a dipole. Spectral and X-ray data show that two hydrogen atoms share a pair of electrons with a single oxygen atom. This special electrical structure of the water molecules makes them group together in a special way: the negative lone-pair electrons of water molecules are attracted to a positively partially screened proton of another water molecule. Thus each corner of the water tetrahedron can be attached, by electrostatic attraction, to four other water tetrahedron molecules in solution. This type of bonding is called hydrogen bonding. However, a sufficiently strong positive system, as of cations, can distort the distribution of the lone-pair electrons enough so that these electrons will be shared between two systems, resulting in the so-called covalent bonding (for example hydroxyls in the surface of the clay). In addition to these bondings, lone pairs of electrons of the water molecule can be electrostatically attracted to such cations as sodium, which may be absorbed on the clay. Therefore, the dipolar structure of water influences such physical properties as diffusion, viscosity, freezing point, and surface tension (characterizing the capillary pressure).

The effect of ions on the physical properties of the soil are well known. The prime example is the ability of sodium to disperse soil colloids. The amount and kinds of ions in the soil solution can greatly influence the water permeability of the soil. It is in regard to water movement that we now examine the distribution of ions of the soil solution about clay surfaces. Because of the structure of the crystal lattice of clay, a clay particle has a net negative charge. This negative charge is compensated by an equivalent positive charge in the soil solution, an excess of cations over anions. The negative charges on a clay particle and the positive charges in solution about the clay particle comprise a double layer of electricity. The total effect of a double layer of charge associated with a clay particle is to make the clay–water system electrically neutral. If we assume the clay particle to be a semi-infinite flat surface, the electrical potential decays, due to the distribution of the cations, in an exponential manner from the clay surface to the bulk solution (Figure 16.11). The cations about the clay surface are divided into two layers: the Stern layer and the "diffuse" layer. The Stern layer consists of cations adsorbed on the surface of the clay, and its
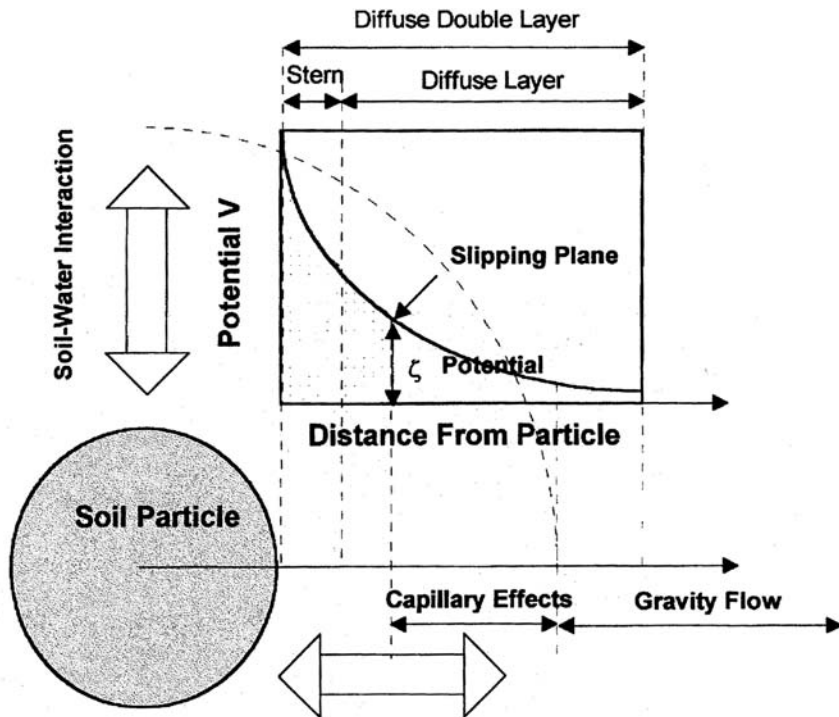
FIGURE 16.11 Potential distribution in clay double layer.

thickness is believed to be about the diameter of the adsorbed cation. The thickness of the diffuse layer, on the other hand, depends on the concentration and the type of ion in the soil solution.

If the clay particle with its ion atmosphere is placed in an electric field, the negative clay particle will tend to migrate in the positive field direction although the predominantly positive atmosphere will tend to migrate in the opposite direction. Some of the cations close to the clay surface will be held so tightly to the clay that they will migrate with the clay. However, other cations farther from the clay in the soil solution will shear away from the clay surface and migrate in the opposite direction of the applied electrical field. The plane that divides the ions into those migrating with the clay and those shearing away is called the "shearing plane" or the "slipping plane." The potential drop between the slipping plane and the bulk solution is called the zeta potential (see Figure 16.11).

If we decrease the cation density in the soil solution, the density of the cations near the clay surface will decrease, and a thicker layer of cations will be needed to screen the negative charge on the clay particle, and the double layer will be thicker. The water movement is influenced by the thickness of the double layer. The thicker the double layer (with the charge density of the clay remaining constant) the higher the zeta potential will be, and the less rapidly water will move through a column of soil.

Another important reference plane shown in Figure 16.11 is the plane between the gravity flow and the capillary effects. Knowing the reference point is essential to determine the retention and movement of fluids (water and contaminants) in the subsurface. The emphasis is in finding an optimum characterization of all the interactions that take place at the interface between the capillary zone and the flow. Indeed, most interaction phenomena (adsorption and desorption) take place at this microscale. For example, organic compounds have a great affinity for clay soils exhibiting a pronounced hysteretic effect between their adsorption and desorption phases. Clearly, the effectiveness of a given remedial technology depends on the accurate depiction of these interactions, influenced by the ambient soil conditions of temperature and pore fluid pressure. The simplest and more accurate parameters representing these interactions are precisely the internal contact angle between different constituents (provides the mag-

nitude of the capillary pressures) and the hysteretic curves of the capillary tension versus effective saturation.

16.3.1 Retention Curves

The fundamental properties of unsaturated flow are described by the retention curve which relates soil tension (water suction) to relative saturation. This relationship exhibits a scale invariance illustrated in Figure 16.12. A new model is presented here which offers the following advantages when compared with older parametric models (Brooks and Corey, 1964, 1966; van Genuchten, 1980): (1) it covers a wider range of capillary tensions; (2) it uses only two parameters (it has parsimony); (3) the two parameters have a well-established physical meaning, at the easily measurable 50% saturation level; and (4) these two parameters are also associated with physical descriptor soil properties, most importantly, not only with the *pore size distribution* (which is less descriptive) but also with the *Atterberg plasticity limits*, which afford a better description of the capillary properties of the soil medium.

The hysteretic behavior of the retention curve is also expressed by an extension of the above parametric model with the introduction of only one additional parameter. The retention curve model and its hysteretic extension are also used in describing the relative conductivity of the medium, all in the case of a deforming soil medium.

Capillary Pressure (Moisture Tension). Fluid is drawn above the saturated horizon by capillary forces. This zone is called the capillary fringe, and the fluid is in a state of tension (suction, negative pressure). The capillary effect is well understood for a capillary tube of fixed radius, but a porous medium is much more complex. A simplified conceptualization of the soil medium can help explain qualitatively the behavior of the medium: it can be thought of as formed of families of tubes of various radii. When the largest of the radii is full a small distance from saturation level, then all tubes are full, and the medium is fully saturated. As we move away from the saturation level, only the finer tubes are full and the relative saturation is reduced, and the corresponding moisture tension is higher. For a given state of the soil (after drainage or wetting), there is a one-to-one correspondence between moisture tension and water content or saturation.

This curve has a characteristic J-shape and can be determined for any type of soil from laboratory experiments. It is unique for every type of soil because not only does it include the effect of tortuosity of the soil pores, but it also includes the effect of irreducible water adsorbed on the soil particles. However, since the pore structure depends to a large extent on the soil particle-size distribution, an attempt has been made to associate all hydraulic soil properties to the particle-size distribution (porosimetry). Various semi-empirical relations have been advanced in the past 50 years expressing the capillary pressure as a function of saturation with the aid of empirical parameters. The most popular are presented here for reference.

J-Leverett Function. Leverett (1942) developed a nondimensional form as follows:

$$J(S_r) = \frac{\psi}{T\cos\alpha}\sqrt{\frac{k}{n}} \tag{43}$$

where ψ is capillary tension (force/area)

J is the J-Leverett function (-)

T is surface tension (force/length)

α is the wetting angle

k is the intrinsic permeability (length^2)

n is porosity (-)

Its relationship with saturation S_r is established from data and this allows us to relate capillary tension to saturation. Separate curves must be developed for the draining or wetting conditions.
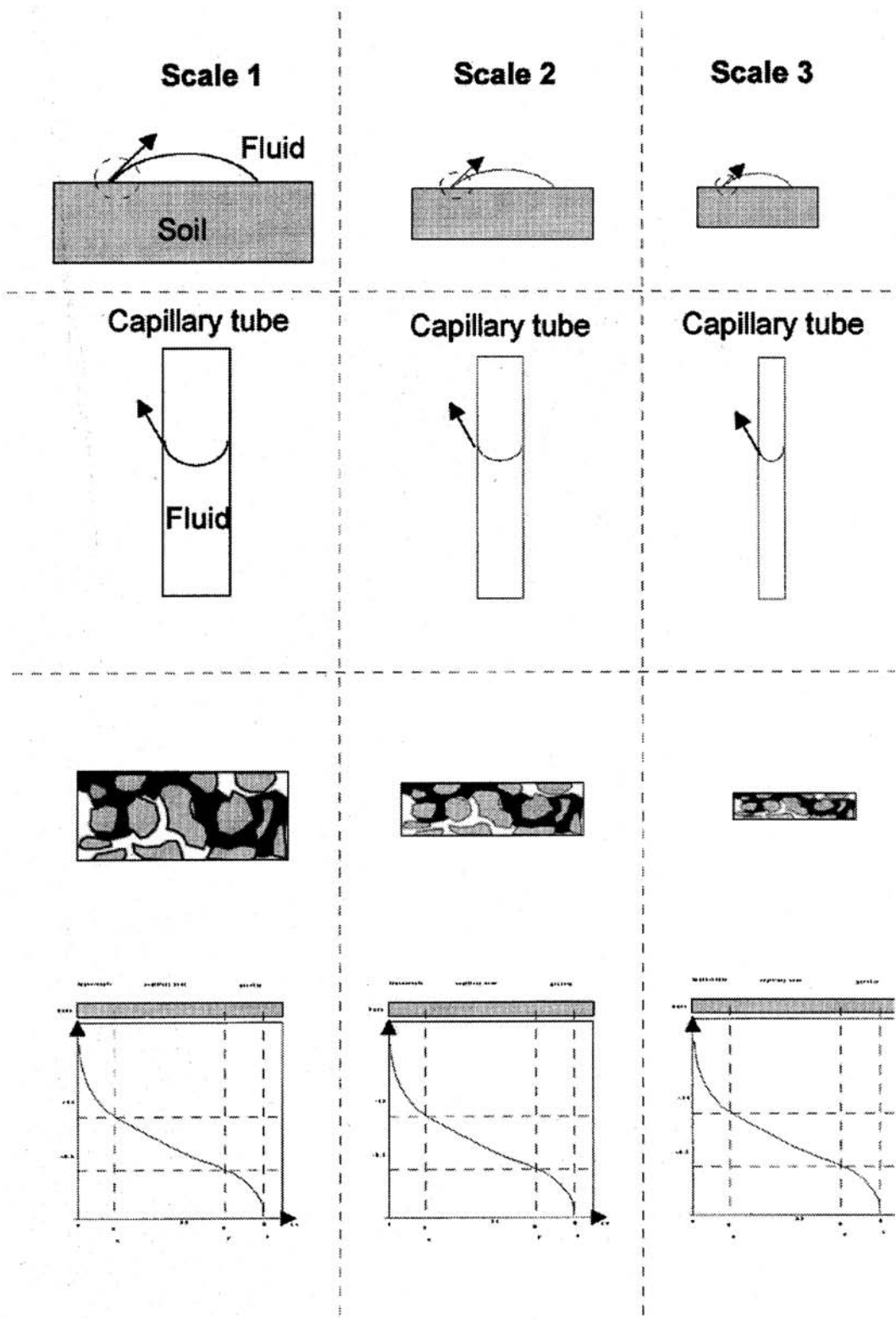
FIGURE 16.12 Scale invariants in unsaturated flow conditions.

The Brooks-Corey Model is an empirical model where capillary tension (or capillary rise) is plotted against the *effective saturation S_e* which is normalized over the effective saturation range, i.e., above the *irreducible saturation S_{ro}*, as follows:

$$S_e = \frac{S_r - S_{ro}}{1 - S_{ro}} = \frac{\theta - \theta_o}{\theta_s - \theta_o} \tag{44}$$

where S_e is effective saturation
   S_r is relative saturation
   S_{ro} is irreducible saturation
   θ_s is water content at zero capillary head
   θ_o is irreducible water content
   θ is water content at saturation level

The Brooks-Corey model is the following:

$$S_e = \left(\frac{p_b}{p_c}\right)^{\lambda} \quad \text{for } p_c \geq p_b$$

$$S_e = 1 \quad \text{for } p_c < p_b \tag{45}$$

where λ is the *pore size distribution index*, representing the medium pore structure (sand 2.30, clay 0.40)
   p_b is the *bubbling pressure* (when gas first appears when dewatering, or capillary tension at $S_e = 1$)

These two parameters are characteristic of the medium. Different sets of values are given for draining and wetting conditions.

The van Genuchten Mode is similar to the Brooks-Corey retention curve, but it requires three parameters, as follows:

$$S_e = \left(1 + \left(\alpha p_c\right)^{n_1}\right)^{-m} \tag{46}$$

where, α, n_1, and m are model parameters.

There is a one-to-one correspondence between the Brooks-Corey and van Genuchten model parameters as follows:

$$\alpha = 1/p_b$$
$$n_1 = \lambda + 1$$
$$m = 1 - 1/n_1 \tag{47}$$

The hysteretic properties of the medium must also be considered separately in the van Genuchten model, with separate curves for wetting and draining conditions.

New Model. Among the limitations to the above models are the following: no explicit treatment of the hysteretic behavior; inconsistent treatment of the very high range of values that capillary tension can take; no explicit physical meaning to the empirical parameters used (α, λ, n_1, and m). The model presented below overcomes these shortcomings (Seker, 1983). First we deal with the question of the range of capillary tension. Following Schofield (1935), we use the variable *pF* which is the logarithm of the capillary tension expressed in water height equivalent, in (cm) units:

$$pF = \log_{10} h_c \qquad (48)$$

where h_c is given in centimeters.

Second, in relation to the capillary-tubes analogy of the soil medium, we express the pore size distribution (porosimetry), in terms of equivalent capillary radii, as a function of saturation. Porosimetric data from Sridharam et al. (1971) show that the following expression describes well a wide range of porosimetries:

$$\log_{10}\left(\frac{r}{r_{max}}\right) = -\psi_o\left(\frac{1 - S_r}{S_r}\right)^{\psi_1} \qquad (49)$$

Both parameters ψ_o and ψ_1 characterize the pore structure and the particle adsorption properties. A precise physical/geometric meaning of these parameters is also given below. Relating capillary rise h_c to capillary tension (matric potential) ψ, to capillary radius r results in the following expression:

$$pF = \psi_o\left(\frac{1 - S_r}{S_r}\right)^{\psi_1} \qquad (50)$$

where ψ_o typically varies in the range (1.00 to 5.00)

ψ_1 typically varies in the range (0.05 to 0.30)

Physical Meaning of Coefficients ψ_o and ψ_1; Exact Definition of Capillary Fringe. The capillary fringe is customarily defined as that zone of the underground horizon which is quasi-saturated ($S_e \cong 100\%$) but which is under capillary suction. For a more exact definition of the capillary fringe, we idealize the $pF(S_r)$ curve as described by Equation (50) and illustrated in Figure 16.13. The slope of this curve at the 50% saturation level intercepts the 100% saturation level at a capillary rise corresponding to pF_c. That capillary rise defines the capillary fringe in terms of the ψ coefficients:

$$pF_c = \psi_o\left(1 - 2\psi_1\right) \qquad (51)$$

$$H_c = 10^{pF_c} \quad \text{in [cm]} \qquad (52)$$

Thus the physical meaning of the coefficients is as follows:

ψ_o is the pF value at 50% saturation (i.e., logarithm of capillary tension in water

height equivalent, in [cm], at 50% saturation).

$-\psi_o\psi_1$ is the slope of the retention curve at 50% saturation.

The advantages of this model in relation to the previous models are the following: (1) it covers the entire range on the capillary tension scale; (2) it does not require experimentally difficult and ambiguous parameters to be determined, such as the "apparent irreducible saturation," and the "bubbling pressure"; (3) because of the double exponentiation, it fits the extreme ranges of the retention curve as well as the middle range. Finally, it requires only two parameters, ψ_o ψ_1, which are defined unambiguously at the easily accessible 50% saturation level.

Correspondence Between $\psi_{o,1}$ Coefficients and Brooks-Corey Coefficients: For comparison purposes, we can establish the following relationship between the new model and the Brooks-Corey model, by equating the two expressions (Equations [45] and [50]) at 50% saturation (where $pF = \log_{10} h_c = \psi_o$):
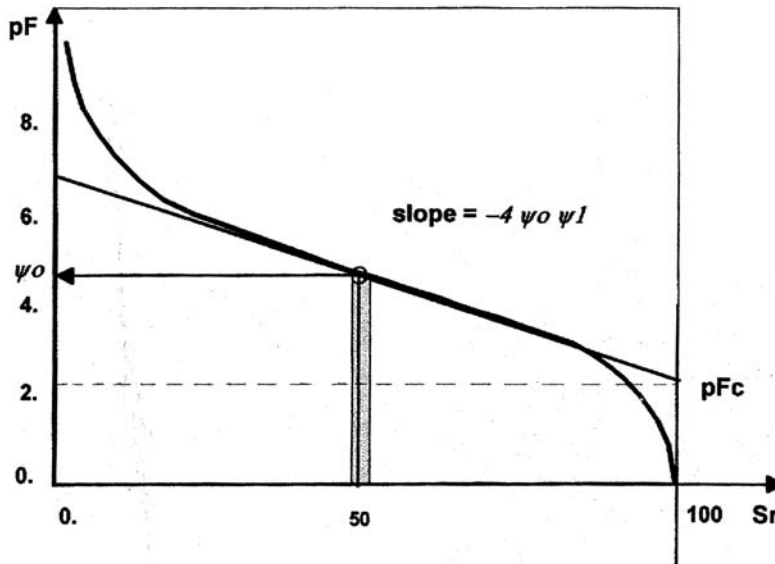
FIGURE 16.13 Analytical determination of capillary fringe.

$$\left(\frac{0.5 - S_{ro}}{1.0 - S_{ro}}\right) = \left(\frac{p_b}{\psi_o \gamma}\right)^{\lambda} \tag{53}$$

When establishing these types of correspondences we must keep in mind that *the earlier models are valid only for smaller ranges of the h_c scale.* Examples of $\psi_{o,1}$ values, pF_c values, corresponding capillary tension, and capillary fringe values are given in Table 16.7 for three types of soil.

TABLE 16.7 Capillary Tension Parameters and Capillary Fringe for Typical Soil Types

Soil Type	ψ_0	ψ_1	pF_c	Capillary Fringe (cm)
Silty sand	1.5–2.0	0.0–0.1	1.5–1.6	32–40
Silts	2.0–3.0	0.1–0.2	1.6–1.8	40–63
Clay	3.0–5.0	0.2–0.25	1.8–2.5	63–316

16.3.2 Hysteretic Models

There are many phenomenological explanations for the hysteretic behavior of soil moisture retention. Essentially, for a given saturation the capillary tension is higher under draining conditions than under wetting conditions. The most prominent explanations are as follows: (1) the pore structure is more complex than the constant radius family of capillaries; in particular, there is a "link-bottle" effect, requiring higher tension during draining; (2) there is a smaller wetting angle during draining; and (3) there is presence of trapped air, which tends to reduce the saturation level, corresponding to a lower capillary tension during the wetting phase.

Mathematically, the hysteretic behavior can be expressed in terms of effective saturations as follows (Mualem, 1976, 1977, 1978) (see also Figure 16.9):

$$S_{eD} = \left(S_{eW}\right)^{\alpha}\left\{2 - \left(S_{eW}\right)^{\alpha}\right\} \tag{54}$$
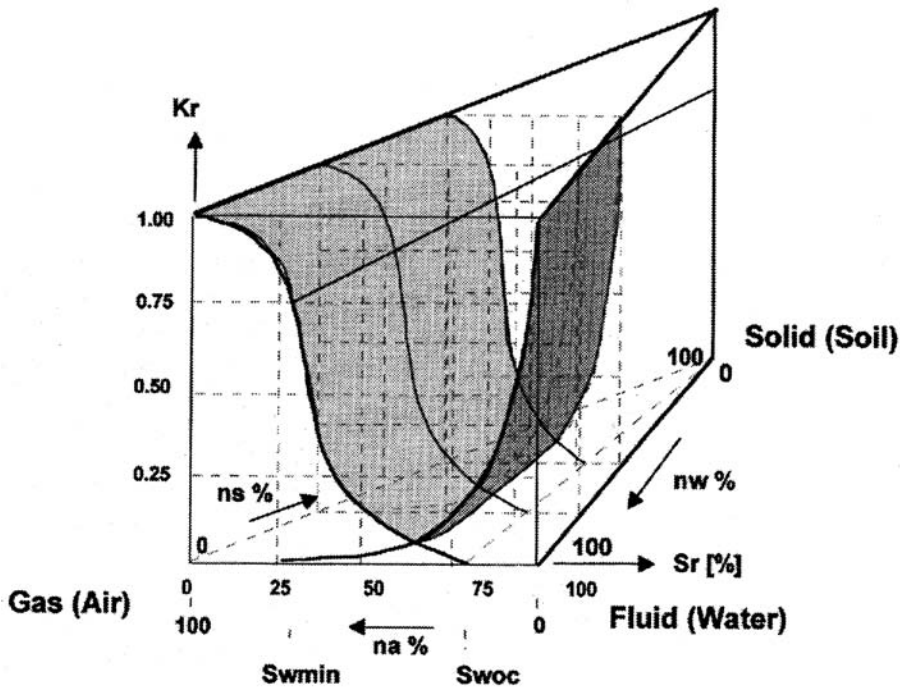
FIGURE 16.14 Relation between relative saturation, permeability, residual saturation, and occlusion condition.

where S_{eD} is effective saturation during draining

S_{eW} is effective saturation during wetting

α is soil medium characteristic exponent (0.5 to 1.0)

This hysteretic model can also be expanded to include the effect of soil compressibility and the hysteretic behavior of permeabilities (ZEi, 1994). Also, an interesting relationship is illustrated in Figure 16.14, namely the inverse variation of the relative permeability of air and water as a function of relative saturation (Barden, 1965): as relative saturation increases, so does the relative permeability of water, while the air permeability decreases. This is because at higher saturation, the air is not free to flow (funicular form) but instead is trapped on the surface grains in the form of bubbles, while water flows at maximum capacity.

Of interest are two particular degrees of relative saturation, namely the *residual saturation* (S_{wmin}) level below which all water is attached (adsorbed) to solid particles and is not free to flow through the pores (zero water permeability); and the *saturation of occluded air* (S_{woc}) level above which all air is entrapped as bubbles in the water phase and there is no air flow (zero air permeability). The general case is when the relative saturation is between the two critical levels, residual and occlusion. In this case, both water and air are continuous (funicular) and are free to flow. For saturations above the occlusion limit only the water phase is mobile, while at saturations below the residual saturation the soil can be considered dry (soil deformation occurs quasi-instantaneously).

16.3.3 Processes and Scales of Interaction

The complexity and the large number of physicochemical processes that take place in the unsaturated zone necessitate a new framework to keep track of their interdependency. The macroengineering framework as implemented in the EIS/GWM platform provide this approach where every process that is selected for analysis is evaluated in terms of its potential interaction and impact on the soil environment. The basic concepts behind this macroengineering framework are explained in Chapter 19, *Holistic Macroengineering Approach for Environmental Information Systems*. An informational network keeps track of the

simulation requirements, adapting itself to the degree of interaction between a given process and the characteristics of the geoenvironment and other existing processes. Time and space synchronization is maintained under the control of a *process scale operator*. This operator controls the integration at the microscale and macroscale levels for each natural process retained for the simulation within the domain of the study. The problem of scale transitions and how to account for them has very important implications for the prediction of the contaminant migration and fate. Essentially, in order to determine the retention and movement of fluids (water and contaminants) in the subsurface, one needs to know the relationship between fluid pressure, degree of saturation, effective stress, void ratio, and permeability. Moreover, depending on whether it is a wetting episode or a draining episode, these relationships show hysteretic behavior. That is, they are not monovalent. Therefore, the history of saturation and pressure must also be known at a given point. But what is the representative scale at this point? The key element in the overall approach is to determine parameters that remain invariant through these different scales. This is where simplicity emerges through some very complex soil interactions.

16.4 Laboratory and Field Measurements of Capillary Tension and Relative Permeability

Laboratory Measurement of Capillary Retention Curve. The general methodology consists of subjecting a sample to successive values of negative pressure and measuring the water content after equilibrium has been reached at every step. For pressures less than 1 atm a simple suction plate can be used. The sample is placed in contact with a porous membrane which is placed under pressure h_f. After equilibrium has been reached, the water content of the sample is measured and this gives a point in the capillary tension–water content curve. For tensions higher than 1 atm, a pressure chamber is used. Rather than subjecting the porous membrane to suction, the sample is put under external pressure and allowed to drain until equilibrium has been reached.

In Situ Measurement of Capillary Retention Curve. In situ measurements are made with the help of a tensiometer. This is a tube filled with water. One end is mounted with a porous head which is pushed into the ground. At the other end, a manometer controls the pressure of the water in the tube. The water in the tube comes in contact with the water in the soil medium and comes to equilibrium. At equilibrium, the manometer reads the water tension prevailing at that depth. This device can read maximum tension in the range of 80 to 100 kN/m². Additional examples of field tests are given in the case study, which follows.

16.5 Description of a Case Study

A typical case study is presented to illustrate some of the concepts advanced in this chapter. Site ST01 is located in the northeastern part of an airport, near the east end of the north–south runway (see Figure 16.15). There are three storage tanks containing aviation gasoline (AVGAS) which were installed at the site in the 1960s. The storage tanks, in addition to numerous above-ground spills, have contributed to a widespread contamination at the site. Several thousand m³ of JP-4 jet fuel were spilled in this area between 1975 and 1984. All tanks and piping have since been taken out of service. A site investigation shows that BTEX is the primary contaminant emanating from the vicinity of the north Tank P100 as shown in Figure 16.16. Limited soil and water quality data are available from which to: (1) quantify the effects of the hydrologic cycle as well as temperature variations on the rate of biodegradation; and (2) evaluate the potential for contamination of the groundwater based on the kinetics of the prevailing natural attenuation mechanisms.

A 20-ton CPT rig was mobilized and, beginning at the leaking tank and stepping outward, CPTu soundings with concurrent vapor sampling at 5-foot intervals were conducted. In the first sounding, sand, silty sand, and clayey sand were identified to the total depth of 59 feet below grade (fbg) (Figure 16.17). Groundwater was identified at a depth of 19 fbg as indicated first by the no-flow conditions in
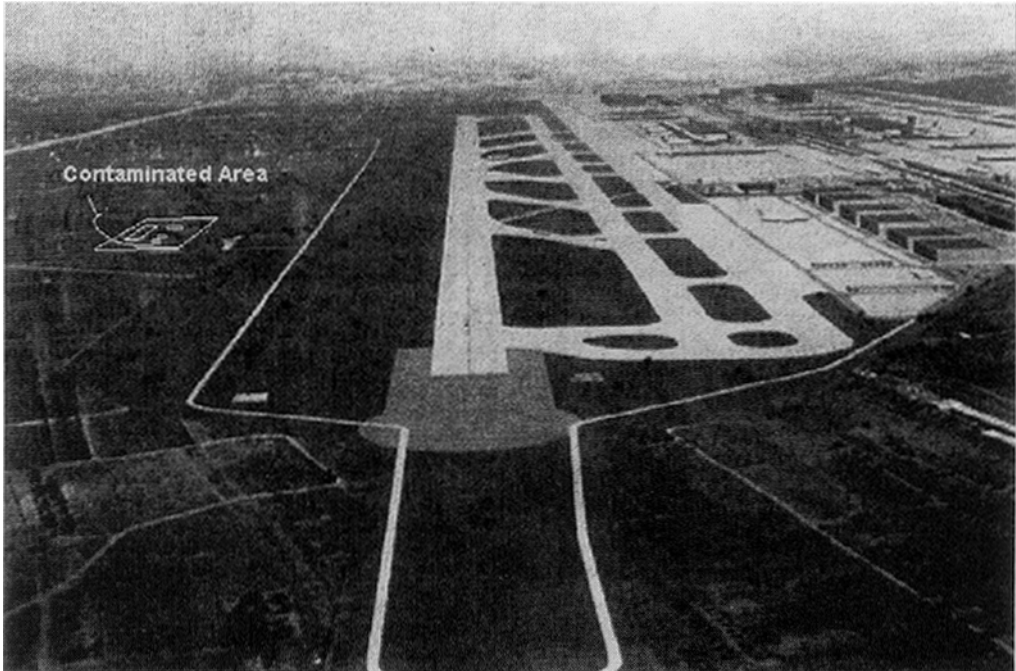
FIGURE 16.15 Aerial view of contaminated airport site.
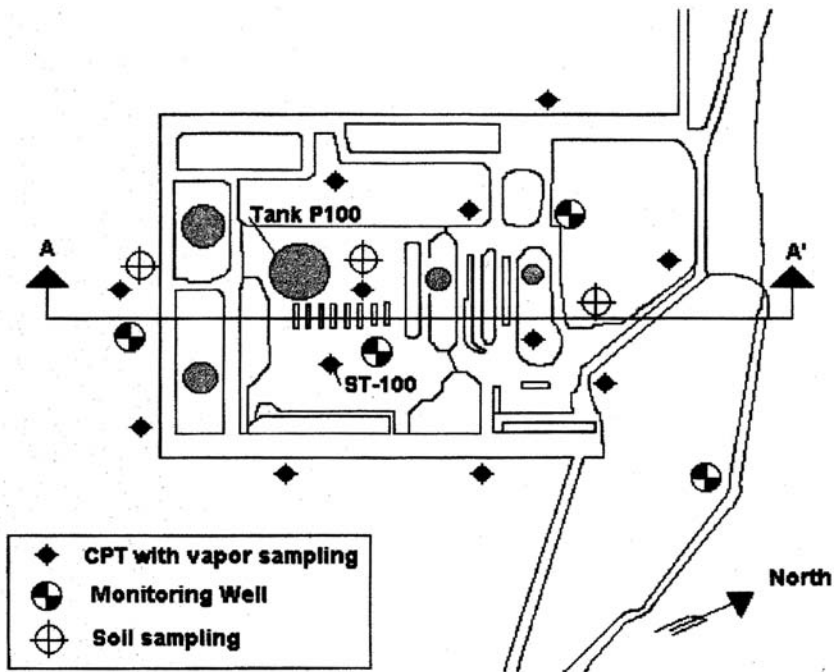
FIGURE 16.16 Plane view of contaminated site.

the vapor sampler, and by a dissipation test at a depth of 32 fbg that showed a pressure head of 3 feet of water (potentiometric surface depth of 19 feet). The vapor sampler began to show flow at a depth of 32 fbg, and a dissipation test at the same depth showed a negative pressure head indicative of unsaturated
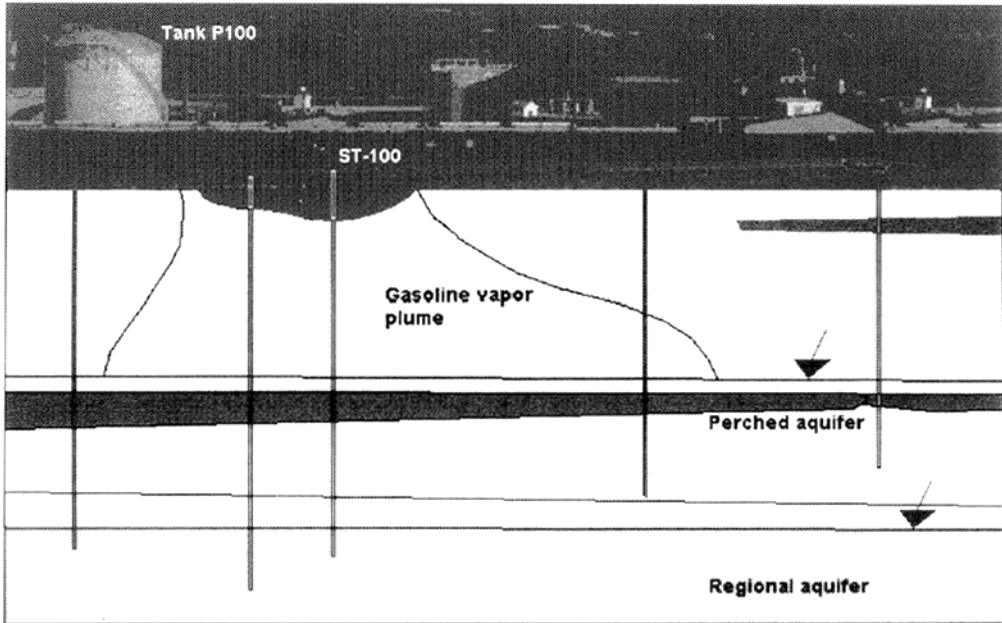
FIGURE 16.17 BTEX plume in the unsaturated zone (cross-section A-A′).

conditions. Water was encountered again at a depth of 38 fbg, as indicated by the dissipation tests at depths of 35 and 50 fbg. The 50-foot and 60-foot dissipation test results showed similar potentiometric surface elevations, indicating no vertical gradients within the deeper aquifer.

The potentiometric surface at 19 fbg was interpreted as a perched aquifer associated with clayey sand at a depth of 20 to 27 fbg (Figure 16.17). The deeper aquifer was interpreted as the regional water table. Vapor sampling indicated high concentrations of TPH in the form of gasoline above a depth of 19 fbg and nondetect TPH below a depth of 19 fbg, even in the unsaturated zone below the perched aquifer. Similar soundings were conducted outward until the lateral limits of the soil plume were defined in all directions. The vertical vapor concentration profiles showed that the plume was restricted to the soils above the perched aquifer, which had thus far stopped downward migration to the regional water table. Regulations required that impacted groundwater be defined by sampling from conventional 4-inch-diameter monitoring wells. This assessment followed a two-step procedure.

Beginning at the known leaking tank, a survey using CPT with dynamic pore pressure measurement (CPTu) outfitted with a soil vapor sampler was performed. A total of 12 soundings to a depth of approximately 60 fbg was performed in 3 days. This survey defined the horizontal and vertical extent of the gasoline vapor plume (see Figure 16.17). All pore dissipation test locations indicated that the horizontal groundwater flow gradient is 0.01 with flow toward the north. Hydropunch™ samples were collected at about half of the CPT locations, which delineated the center and limits of the BTEX plume. This concluded Step 1 of the survey and resulted in detailed delineation of site geology, site hydrogeology, and the extent of soil and groundwater plumes. Step 2 began with collection of soil samples with the CPT rig at two locations within the soil vapor plume and at one location around the vapor plume, and resulted in quantification of soil concentrations and confirmed the definition of the limits of the soil plume (see Figure 16.18). The data of the site investigation were sorted and analyzed in the EIS platform resulting in the determination of the spatial distribution of key biodegrading parameters.
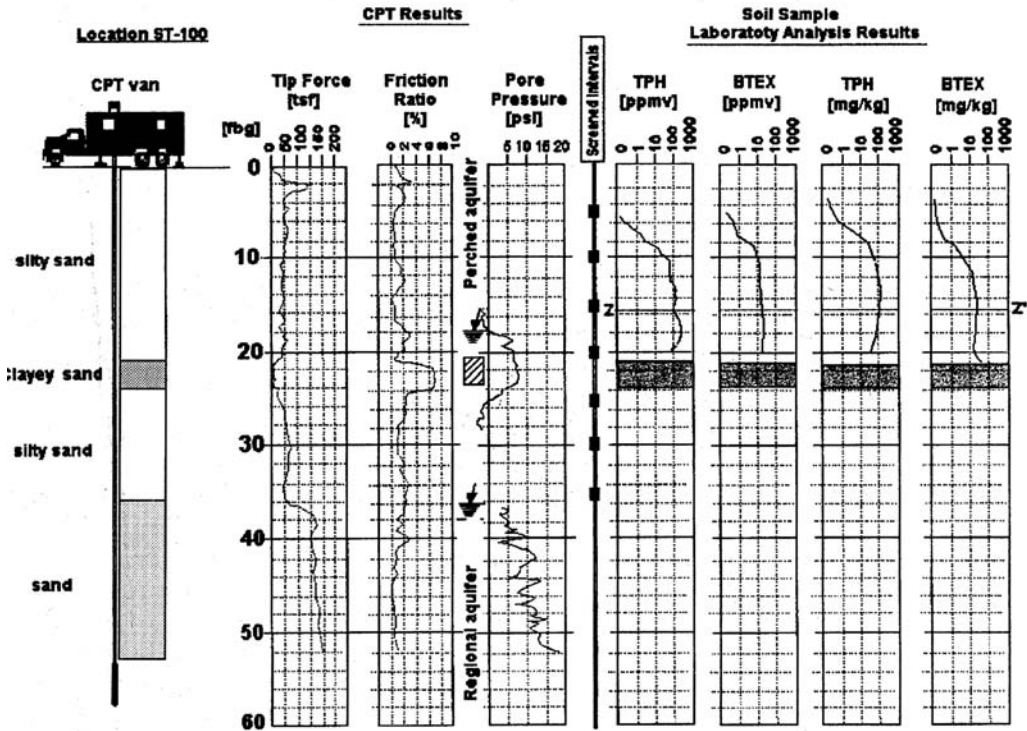
FIGURE 16.18 Typical CPT data.

16.5.1 Parameters Affecting Microbial Processes in the Unsaturated Zone

Biodegradation is influenced by the bioavailability of the contaminants, driven at least in part by solubilization. Since microorganisms exist in pore water, contaminants must partition into the pore water to be available to be degraded. Although high soil contaminant concentrations may be present, the actual concentration of hydrocarbon dissolved in the pore water and available to the microorganisms may be low. In practice, oxygen utilization rates tend to decline slowly with time during remediation. At many sites, this trend may be difficult to follow over periods of less than 1 to 3 years because of other variables affecting the rate, such as temperature and soil moisture. This decline may not be indicative of true first-order kinetics, but simply may be the result of selective early removal of more degradable compounds such as benzene. Therefore it is of primary importance to characterize as accurately as possible the spatial distribution of the environmental parameters influencing the microbial life. Biodegradation in the unsaturated zone depends upon providing microorganisms optimal conditions for active growth. Several factors may affect a microorganism's ability to degrade contaminants, including availability and type of electron acceptors, moisture content, soil pH, soil temperature, nutrient availability, and contaminant concentration.

Each of these parameters was measured at the studied site. The actual effect of individual parameters on microbial activity is difficult to assess in the field due to interference and interactions among these parameters. To complete the missing data, an automated kriging inference scheme was used (EIS platform). The *in situ* respiration test was used as a measurement that integrates all factors to assess whether the microorganisms are metabolizing the fuel. All site data obtained from CPTu and laboratory tests were analyzed to obtain their *spatial distribution*. In particular the following features were used:

1. An editor for dynamically assigning attributes to log points of the site investigation. This tool allows the user to work freely with any set of field data, by scripting the properties of each sensor (measuring device).
2. An automatic procedure allowing the user to delineate the medium stratification on the basis of all profile data (CPT, CPTu tests).
3. An automatic procedure for developing spatial distributions of the measured parameters (properties) as integrated over the user-specified stratification. Spatial distributions are also associated with interpolation error estimates.
4. An automatic procedure to predict the anticipated probe signal at points which have not been sampled. This feature (fractal kriging) allows the site manager to complete the QA/QC loop on the spot, and develop a sense of reliability on his/her interpretation of the site geostratigraphy (e.g., look for missing clay lenses).

The outcome of these operations is summarized below.

Recorded Site Conditions. Biodegradability of a compound depends on the availability of electron acceptors. While hydrocarbons may undergo limited biodegradation under anaerobic conditions, in general, aerobic conditions are most suitable for relatively rapid remediation of petroleum hydrocarbons. Therefore, oxygen supply is critical to the success of biodegradation. Oxygen depletion is an indication that biodegradation is taking place as shown in Figure 16.19. Note that the largest depletion takes place around the core of the contaminant plume. Other parameters are also contoured in Figure 16.20. In particular, the high concentrations of iron underneath the contaminated area show anaerobic biodegradation at work. Since only one snapshot of the data exists (no time variation), the EIS platform is needed to assess the validity of the biodegradation hypothesis. This modelization requires a closer look at the spatial distribution of the moisture content of the site because both aerobic and anaerobic processes are very sensitive to the water content of the soil medium.

Moisture Content. Soil moisture affects the biodegradation process in the unsaturated zone by its effect on microorganisms and soil gas permeability. Microorganisms require moisture for metabolic processes and for solubilization of electron acceptors and nutrient supplies. Conversely, soil moisture content directly affects the soil permeability, with high moisture contents resulting in poor distribution of oxygen. Also, excess moisture has led to significant reductions in soil gas permeability. The lowest soil moisture content measured at the site was 2% by weight, and microbial activity still was observed in these soils. To better characterize the water release characteristics between saturation and field capacity an undisturbed core sample was analyzed *in situ*, using a modified Haines test in which the soil is placed in a Buchner funnel made with a porous sintered glass plate the base of which is connected via a water-filled polyethylene tube to a graduated glass tube. The soil is initially saturated. Then by lowering the graduated tube, a suction is applied to the soil water. Soil water suction is measured by the equivalent hydraulic head, and changes in soil water content by changes in water level in the graduated tube. The site data of this test are shown in Table 16.8.

The mass of water present in 50.32 g of sand is 1.43 g. Assuming that 1 g of water occupies 1 cm^3 the final value of water content is 1.43 from which the other values of the column of the water content are calculated. The conversion to volumetric water contents shown in the last column of the table is obtained as follows: The particle density of sand is 2.65 gcm^{-3} and thus its specific volume is $1/2.65 = 0.377$ cm^{3}g^{-1}. Thus saturated sand with a gravimetric water content of 0.235 g H_2Og^{-1} sand has 0.235 cm^3 H_2O $+$ 0.377 cm^3 sand $= 0.612$ cm^3 of total volume and its volumetric water content $= 0.235/0.612 = 0.384$ cm^3 cm^{-3}. Assuming that the volume of sand does not change, the volumetric water content is evaluated dividing the gravimetric water content by 0.612.

The water release characteristics are illustrated in Figure 16.21. A small amount of water is withdrawn initially which is "free" on the surface of the soil bed. Most of the water is released between 10 and 30 cm H_2O of suction. An interesting observation is that the volumes of the water released can be correlated to the pore size distribution (Arya and Paris, 1981; Burdine, 1952). This correlation complements the results of the CPT tests and is automatically handled by the EIS platform as follows. The volumes of
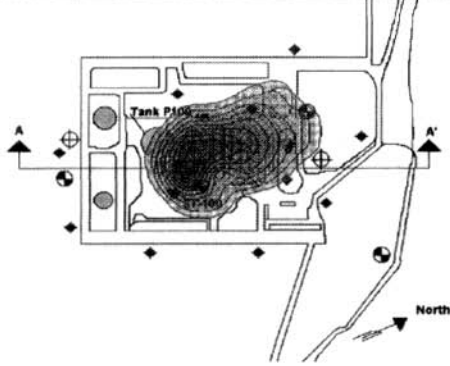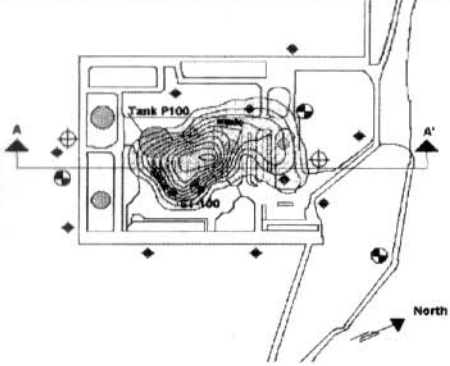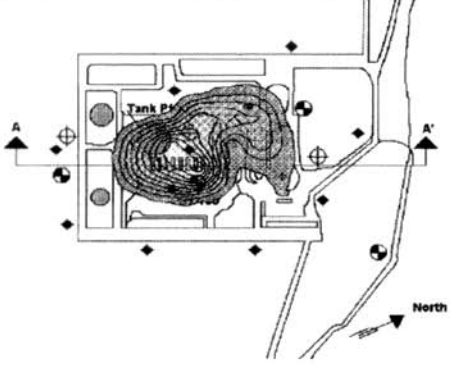
Contours at Level z = 19.00 fbg	Description
	BTEX in the (Soil + Moisture) **Max**= 1000 [ppmv] **Min**= 10 [ppmv] **Remarks:** The contours are obtained just above the perched aquifer. The plume seems to expand in the north direction.
	BTEX in the Air **Max**= 100 [ppmv] **Min**= 10 [ppmv] **Remarks:** The contours are obtained just above the perched aquifer. The plume seems to expand in the north direction.
	Oxygen **Max**= 10 [ppmv] **Min**= 1 [ppmv] **Remarks:** The contours are obtained just above the perched aquifer. The plume seems to expand in the north direction.

FIGURE 16.19 Contours of the most important parameters recorded at the site.

water released for each 5 cm increase in suction are also illustrated in Figure 16.21. The x-axis is labeled as pore size calculated by the following relation: d ~ 3000/h where d is in m and h is in cm. As it can be observed the pores which empty between 10 and 30 cm suction have diameters between 300 and 100 m, and occupy 0.288 cm^3 cm^{-3}, which is 75% of the total pore volume, and 88% of the pore volume is in transmission pores (>50 μm). A sand of this particle size in the field would drain to a water content at field capacity of only about 0.05 cm^3 cm^{-3}, indicating the small amounts of available water held by coarse sandy soils. These results are essential to quantify infiltration rates and groundwater recharge.
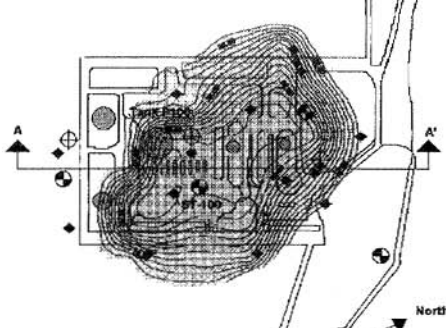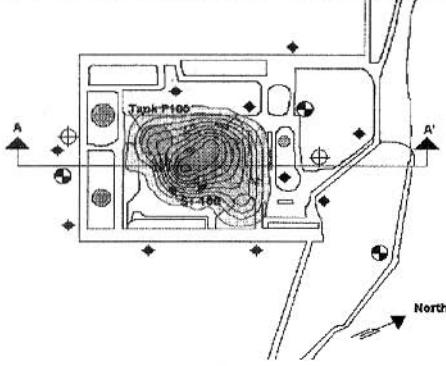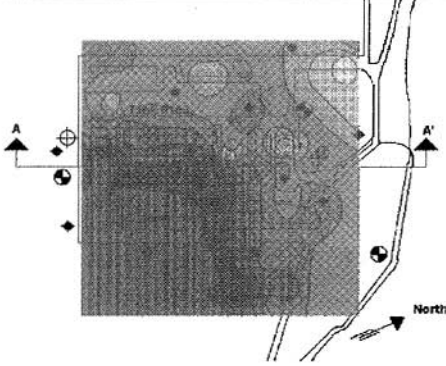
Contours at Level z = 19.00 fbg	Description
	CO₂ in the Air

Max=

Min=

Remarks: The contours are obtained just above the perched aquifer. The plume seems to expand in the north direction. |
| | **Iron in the (Soil + Moisture)**

Max= 12,000 [mg/kg]

Min= 2,000 [mg/kg]

Remarks: The contours are obtained just above the perched aquifer. The plume seems to expand in the north direction. |
| | **Moisture**

Max=35 % wt

Min= 5 % wt

Remarks: The contours are obtained just above the perched aquifer. It seems that the contours follow the gradient of the groundwater flow. |

FIGURE 16.20 Contours of other important parameters recorded at the site.

16.5.2 Modeling Objectives and Results

The modeling objectives were the following: (1) quantify the impact of rainfall/infiltration on the bio-degradation rates of the BTEX contaminant; and (2) quantify the contaminant depletion at various phases after a two-year period. The model for the unsaturated zone was developed as follows. The domain of interest was delineated to the shallow unsaturated zone, which was conceptualized as a nonhomogeneous geologic formation composed of silty sand with variable thickness and hydraulic conductivity. The source of contamination was assumed to be Tank P100, and it was further assumed that the remaining free-phase and residual-phase contamination would act as a discontinuous source of contamination. Twelve

TABLE 16.8   Data of the Haines Test

Hydraulic Heads (cm)	Volume Reading (cm³)	Volume Withdrawn (cm³)	Water Volume in the Sand (cm³)	Gravimetric Water Content (g/g oven-dry)	Volumetric Water Content (cm³/cm⁻³)
0	$V_0 = 33.05$	0	11.83	0.235	0.384
7.40	32.30	0.75	11.08	0.220	0.359
14.60	31.40	1.65	10.18	0.202	0.330
18.40	26.90	6.15	5.68	0.113	0.185
26.75	23.80	9.25	2.58	0.051	0.083
38.90	22.90	10.15	1.68	0.033	0.054
65.60	22.65	10.40	1.43	0.028	0.046
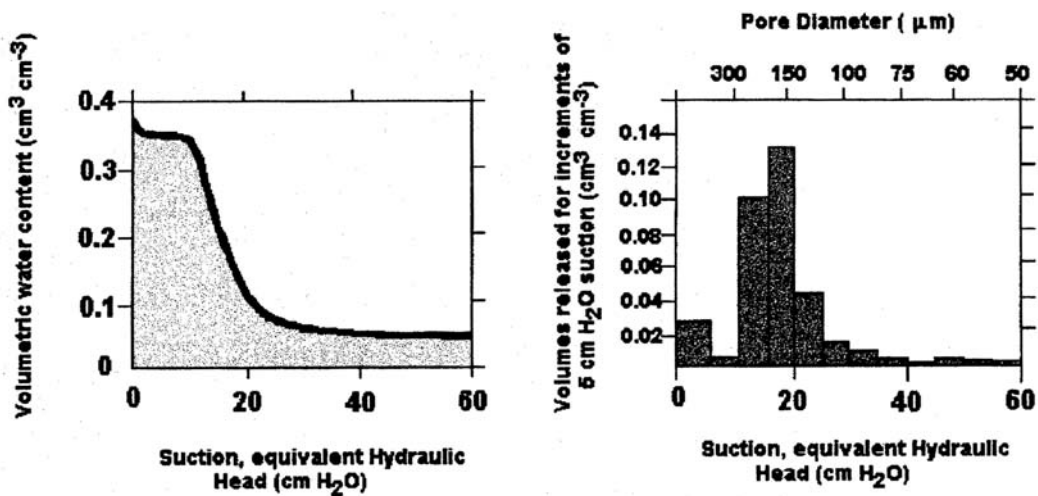
FIGURE 16.21 The water release characteristics and the deduced pore size distribution.

numerical layers above the perched aquifer were retained for an accurate overall simulation of the effects of infiltration.

Soil moisture content ranged from 2% to 4% by weight and, although the site was contaminated with jet fuel, significant oxygen limitation was not observed. A water infiltration was applied at the site in order to investigate its effects on biodegradation. The site was irrigated for more than a week. *In situ* oxygen depletion rates after infiltration were significantly higher than those prior to water infiltration (Figure 16.22). In addition, prior to infiltration, oxygen was not depleted below approximately 17% before microbial activity stopped. After infiltration, activity continued until oxygen was completely consumed to less than 1%. These simulation results demonstrate that infiltration may improve the performance of biodegradation through enhanced microbial activity.

For Further Information

Bear (1972) is an original reference text on all theoretical aspects of groundwater flow and transport in porous media. Bedient et al. (1994) in Chapter 9 provide a compilation of most existing screening models for transport in the unsaturated zone. Maidment (1993) in Chapter 15 of the *Handbook of Hydrology*, addresses unsaturated transport and also provides a compilation of computer models. Brady (1974) provides a good physicochemical description of soils. Hanks and Ashcroft (1980) and Rowell (1994) provide a good compendium of applied soil physics information. Domenico and Schwartz (1990) is an
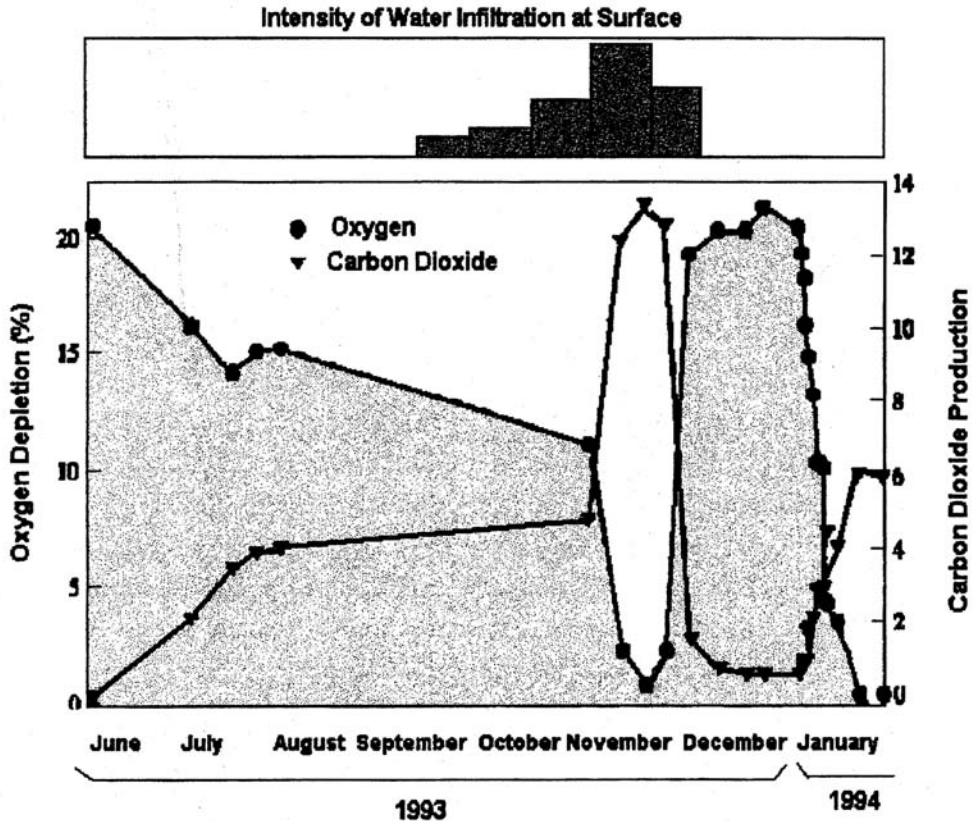
FIGURE 16.22 Oxygen depletion and carbon dioxide production prior to and after infiltration.

excellent reference of physicochemical processes from a hydrogeologic viewpoint. Dullien (1992) provides an original series of theories on pore structure as it affects flow and transport. Scheidegger (1974) is also a classic reference on theoretical concepts about the physics (modeling) of flow through porous media.

References

AFCEE. 1995, 1996. *Technical Protocol for Evaluating Natural Attenuation of BTEX and Chlorinated Solvents in Groundwater.* Air Force Center for Environmental Excellence, Brooks AFB, San Antonio, Texas.

Arya, L. M. and Paris, J. F. 1981. A physico-empirical model to predict soil moisture characteristics from particle size distribution and bulk density data. *Soil Sci. Soc. Am. J.* 45, 1023-1030.

Barden, L. 1965. Consolidation of compacted and unsaturated clays. *Geotechnique.* 15, 267-286.

Bear, J. 1972. *Dynamics of Fluids in Porous Media.* Elsevier, New York.

Bear, J. 1979. *Hydraulics of Groundwater.* McGraw-Hill, New York.

Bedient, P. B., Rifai, H. S., and Newell, C. J. 1994. *Ground Water Contamination, Transport and Remediation.* Prentice Hall, Englewood Cliffs, NJ.

Bishop, A. W. and Donald, I. B. 1961. The experimental study of partly saturated soil in the triaxial apparatus. *Proc. 5th Int. Conf. Soil Mech. and Found. Eng.*, Vol. 1, Paris, 13-21.

Brady, N. C. 1974. *The Nature and Properties of Soils,* 8th edition. Macmillan, New York.

Brooks, R. H. and Corey, A. T. 1964. Hydraulic properties of porous media, Hydrology paper No. 3, Colorado State University, Fort Collins, CO.

Brooks, R. H. and Corey, A. T. 1966. Properties of porous media affecting fluid flow, *Proc. ASCE, J. Irrigation Drainage Div.*, no. IR2, paper 4855, 61-88.

Burdine, N. T. 1952. Relative permeability calculations from pore-size distribution data, *Petrol. Trans.*, Am. Inst. Mining, Met. Petrol. Engrs., vol. 198, 71-77.

Debye, P. 1929. *Polar Molecules.* Reinhold, New York.

Domenico, P. A. and Schwartz, F. W., 1990. *Physical and Chemical Hydrogeology.* John Wiley & Sons, New York.

Dullien, F. A. L. 1992. *Porous Media, Fluid Transport and Pore Structure*, Second Edition. Academic Press, San Diego.

Geissman, T. 1977. *Principles of Organic Chemistry.* W.H. Freeman and Company, San Francisco.

Hanks, R. J. and Ashcroff, G.L. 1980. *Applied Soil Physics.* Springer-Verlag, Berlin.

Hildebrand, J. H. and Scott, R. 1950. *The Solubility of Nonelectrolytes*, 3rd. ed. Rheinhold Publishing Corp., New York.

Hoshina, K. 1961. An analysis of the volume change, distortional deformation induced pore pressure of soil under triaxial loading, *Proc. 5th Int. Conf. Soil Mech. and Found. Eng.*, Vol. 1, Paris, 155-157.

Lester, J. N. 1987. *Heavy Metals in Wastewater and Sludge Treatment Processes.* CRC Press, Boca Raton, FL.

Leverett, M. D., Lewis, W. B., and True, M. E. 1942. Dimensional model studies of oil-field behavior. *Petroleum Technology.* T. P. 1413, Jan., 175-193.

Maidment, D. R. (Ed.). 1993. *Handbook of Hydrology.* McGraw-Hill Book Company, New York.

Morrow, N. R. 1970. *Ind. Eng. Chem.* 62, 32. Reprinted as a chapter in *Flow through Porous Media.* American Chemical Society, Washington, DC.

Mualem, Y. 1976. Hysteretical model for prediction of the hydraulic conductivity of unsaturated porous media. *Water Resour. Res.* 12, 1248-1254.

Mualem, Y. 1977. Extension of the similarity hypothesis used for modeling the soil water characteristics. *Water Resour. Res.* 13, 773-780.

Mualem, Y. 1978. Hydraulic conductivity of unsaturated porous media: generalized macroscopic approach. *Water Resour. Res.* 14, 325-334.

Overcash, M. and Pal, D. 1979. Design of Land Treatment Systems for Industrial Wastes-Theory and Practice. Methuen, London.

Pearson, R. G. 1963. Hard and soft acids and bases. *J. Am. Chem. Soc.*, 85.

Pearson, R. G. 1968. Hard and soft acids and bases, Parts I and II. *J. Chem. Educ.* 45.

Rowell, D. L., 1994. *Soil Science: Methods and Applications.* Addison Wesley Longman Limited, London.

Sawyer, C. and McCarty, P. 1978. *Chemistry for Environmental Engineering.* McGraw-Hill Company, New York.

Scheidegger, A. E. 1974. *The Physics of Flow Through Porous Media,* Third Edition. University of Toronto Press, Toronto.

Schofield, R. K. 1935. The interpenetration of the diffuse double layers surrounding soil particles. *Int. Congr. Soil Sci.*, 3rd. 1, 30-33.

Seker, E. 1983. *Etude de la Deformation d'un Massif de Sol Non-Sature.* Thesis No. 492, Federal Institute of Technology, Lausanne, Switzerland.

Sposito, G. 1989. *The Chemistry of Soils.* Oxford University Press, New York.

Sridharam, A., Altschaeffl, A. G., and Diamond, S. 1971. Pore size distribution studies, *J. Soil Foundation Div.*, ASCE SM5 97. 771-787.

van Genuchten, M. Th. 1980. A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Sci. Soc. Am. J.* 44, 892-899.

van Genuchten, M. Th. and Alves, W. J. 1982. Analytical solutions of the one-dimensional convective-dispersive solute transport equation. U.S. Department of Agriculture Technical Bulletin 1661, 151 p.

ZEi, ZSOIL. 1990. *Soil Mechanics Using Plasticity Theory.* ZEi Engineering, Annandale, VA.

ZEi. 1994. *Theoretical Foundation of EIS Scientific Engines*, Report No. Mi-94-M003. ZEi Engineering, Annandale, VA.

Glossary

Capillary Pressure   is the pressure difference at the interface between a nonwetting and a wetting fluid.

Permeability (intrinsic), and Fluid Conductivity   Fluid conductivity is the coefficient of proportionality in Darcy's law, measured in units of velocity (length/time), and it depends both on the properties of the fluid (viscosity, density), and the porous medium (pore size distribution, tortuosity). The effect of porous medium and fluid properties is functionally separated as

$$K = f_1\left(\frac{\rho g}{\mu}\right) f_2\left(d^2\right)$$

where μ is the fluid viscosity, d is a characteristic soil grain size, and $f_2 = k\,Cd^2$ is the *intrinsic permeability* [length^2].

Relative Permeability   is the ratio of the unsaturated to the saturated permeability (or hydraulic conductivity).

Relative Saturation   of a fluid is the ratio of the volume of that fluid, divided by the total pore volume.

Wettability   is characterized by the angle that is formed between a drop of fluid resting on a clean surface and the surface.

Wetting Fluid Content   is the ratio between the volume of the wetting fluid and the corresponding total soil sample volume.

17
Groundwater Flow and Solute Transport in Fractured Media

P. A. Lapcevic and
K. S. Novakowski

National Water Research Institute
Burlington, Ontario

E.A. Sudicky

University of Waterloo

17.1 Introduction

Many of the world's largest aquifers reside in fractured consolidated media. Studies have shown that even media traditionally considered to be of low permeability, such as shales and granites, are fractured to the extent that significant groundwater flow may occur. In North America, there are many locales at which bedrock of either sedimentary or crystalline origin is used for either domestic or commercial water supply. In many geological settings, the lack of a protective layer of low-permeability clays or clay tills makes these water supplies particularly vulnerable to groundwater contamination.

Understanding groundwater flow in fractured consolidated media has long been important when undertaking engineering tasks such as dam construction, mine development, the abstraction of petroleum, slope stabilization, and the construction of foundations. To study groundwater flow in support of these tasks, often only relatively simple measurements of permeability conducted *in situ* are required. More recently, the discovery of groundwater contamination in fractured bedrock materials has led to the

advent of many new investigative techniques that are quite different from those previously used in the geotechnical and petroleum industries. In particular, the focus of most hydrogeological investigations has been on the characterization of the transport properties of the higher-permeability fractures in the rock mass.

On many occasions, groundwater investigations benefit from the development of a numerical model which is used to simulate the flow system. However, because groundwater flow is primarily governed by discrete fractures at the local scale, traditional mathematical methods, in which it is assumed that the hydrogeological properties vary in a smooth and continuous fashion (a single continuum), provide a poor approximation. Only at large, regional scales of groundwater flow or where fracturing is very dense might continuum methods be applied in fractured media. Further, the often complex and heterogeneous arrangement of discrete fractures makes for difficulties in the representation of these features on an individual basis. There are several modeling approaches that can be used to circumvent this problem, including the use of stochastic or mixed deterministic-stochastic techniques (e.g., Smith and Schwartz, 1984) and the use of scale-dependent multicontinua (e.g., Bear, 1993). However, numerical models which can represent two-dimensional groundwater flow and solute transport in each discrete fracture within a three-dimensional randomly fractured domain of practical size have yet to be applied. Because it is impossible to characterize a field site to the degree necessary to support this type of model, a compromise such as the alternate approaches suggested above will almost always be required.

In this chapter, we will try to provide basic theory and practical advice for guiding investigations of the hydrogeology of fractured rock sites. Our aim has been to focus this discussion on the hydrogeological representation of fractured rock systems at scales below which an equivalent porous media (EPM) approach is adequate. To this end, the modeling and field sections will outline techniques to characterize discrete fractures and fracture networks in the subsurface. A basic introduction to structural geology will be provided to assist in determining how the features that govern groundwater flow in fractured media are formed and how to go about finding them. A brief outline of the principal processes involved in the flow of groundwater and transport of contaminants in individual fractures is presented. In addition, some methods commonly used in the field characterization of fractured rock will be described. These elements become the building blocks for the development of a conceptual model of groundwater flow. A discussion of conceptual models for both single fractures and fracture networks is included. Finally, an overview of analytical and numerical modeling techniques for fractured rock will be presented.

17.2 Structural Geology of Fractured Rock

In this chapter, the term fracture is used to refer to any discontinuity in a solid material. In geological terms, a fracture is any planar or curviplanar discontinuity that has formed as a result of a process of brittle deformation in the earth's crust (Engelder et al., 1993).

Planes of weakness in rock respond to changing stresses in the earth's crust by fracturing in one or more different ways depending on the direction of the maximum stress and the rock type. Three broad types of fractures have been identified based on the motion involved in the formation of the fracture. Opening mode (mode I) cracks form by the separation of the crack walls, under the action of tensile stresses normal to the crack plane (Figure 17.1a). Mode I fractures are thus extension fractures in which the fracture walls have displaced normal to the crack plane and no appreciable displacement parallel to the fracture plane has occurred. Sliding mode (mode II) cracks propagate by mutual shearing of the crack walls with the shearing oriented in the direction normal to the crack front (Figure 17.1b). Tearing mode (mode III) cracks advance when the crack walls are subject to a shearing aligned parallel to the crack front (Figure 17.1c). Mode II and III fractures are defined as shear fractures.

Joints (mode I fractures) are the most common type of fracture structure and are found in all types of rocks as well as in partly consolidated to unconsolidated sediments. Joint sets are characterized as systematic or nonsystematic. Joints belonging to a set maintaining consistent characteristics such as orientation, morphology, mineralization, and distribution between outcrops in a local and regional scale
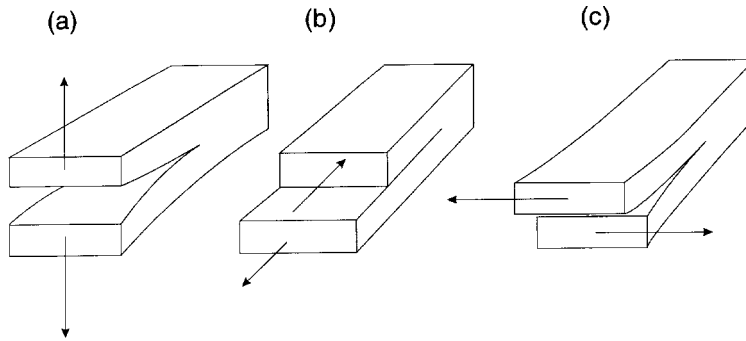
FIGURE 17.1 Schematic diagram showing mechanism of fracture formation: (a) Mode I (opening), (b) Mode II (sliding), (c) Mode III (tearing). (From Atkinson, B. K. 1987. *Fracture Mechanics of Rock*. Academic Press, Orlando, FL. With permission.)

are defined as systematic (Engelder et al., 1993). Nonsystematic joints have highly curved and irregular fracture surfaces and are not easily related to a recognizable stress field.

Groundwater movement in fractured rock depends on discontinuities at a variety of scales ranging from microcracks (micrometers, μm, in length and width) to faults (kms in length and meters in width). Fracturing in a large-scale rock mass can be loosely classified on the basis of three scales of discontinuity: (1) microscopic, (2) mesoscopic, and (3) megascopic.

At the microscopic scale, isolated or continuous disc-shaped microcracks ranging in length from 100 to 1000 μm and 1 to 2 μm in width are common to both crystalline and sedimentary rocks. Microcracks are normally mode I fractures, and often form in a disc-shape. In some cases, particularly adjacent to shear fractures, microcracks may form under mode II conditions.

In crystalline rocks, the total porosity of the rock matrix (i.e., the rock bounded by macroscale fractures) is due to microcracks and may range from 0.01% to 1.0%. In sedimentary rock, in addition to microcracks, significant porosity in the matrix may also occur due to incomplete cementation, the weathering of vugs, or imperfect lattice development. Matrix porosity in sedimentary rock may range from less than 1% to 50%.

At the mesoscopic scale, fractures (or joints) ranging between <0.5 m and >1000 m in length with widths ranging from <10 μm to >10 mm predominate in the groundwater flow system. These fractures are often grouped in sets having similar orientation which are related to the mode of fracture genesis. Fracture zones at this scale are defined as sets of closely spaced fractures contained within a narrow band. Shear fractures are also observed at this scale. A common and extremely important class of fractures at this scale is known as sheeting structure (Holzhauser, 1989). Sheeting structure forms due to the vertical expansion of the rock resulting from erosional unloading. Sheeting structures are present in all types of rock, often found parallel to the earth's surface, and are common conduits for groundwater flow at the regional scale. In plutonic and metamorphic rocks, they are often independent of the original macro-structure and fabric of the rock and are found at decreasing frequency with increasing depth (Trainer, 1988).

At the megascopic scale, faults and shear zones dominate the fracture morphology. A fault is defined as a planar discontinuity between two blocks of uniform rock where displacement has taken place parallel to the discontinuity. These are large-scale fractures with many associated features and may be either of high or low permeability depending on infilling and local stress fields. Faults may be of any orientation, although it is the sub-vertical features that are most easily identified through air-photo or map interpretation. Shear zones are similar to faults (i.e., a large-scale fracture zone which has undergone shearing) and are typically found in metamorphic terrain. These features are often of high permeability. In many fractured rock environments, the frequency and spacing of faults and shear zones are sparse. For example, Figure 17.2 illustrates in plane view the possible surface expression of a hypothetical distribution of
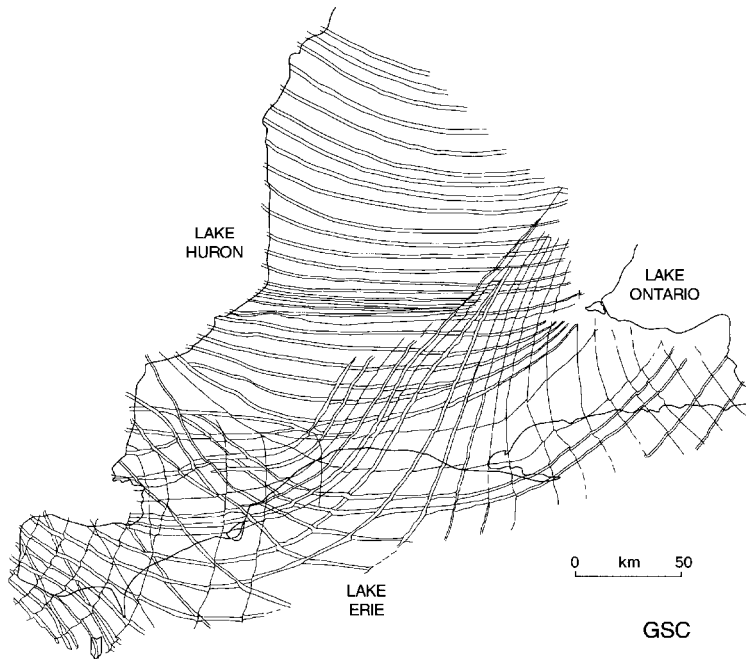
FIGURE 17.2 Distribution of regional faults in southern Ontario. (From Sanford, B. V., Thompson, F. J., and McFall, G. H. 1985. Phanerozoic and recent tectonics of the Canadian Craton, in *The Geoscience Program — Proceedings of the Seventeenth Information Meeting of the Nuclear Fuel Waste Management Program*, Atomic Energy of Canada Limited, TR-299:334-352. With permission.)

vertical faults in the southern Ontario region of central Canada. These faults are postulated to vertically penetrate the entire thickness of the sedimentary strata which ranges from ~200 m to ~2 km, in this area (Sanford et al., 1985). Note that the spacing between faults is several to tens of km. Thus, at the local to regional scale, any one of these faults may dominate the groundwater flow system.

Karst aquifer systems develop due to the extensive dissolution of fractured carbonate rock by flowing groundwater. The hydrogeology of this type of rock terrain is discussed in the following chapter.

17.3 Fundamentals of Groundwater Flow and Solute Transport in a Single Fracture

17.3.1 Groundwater Flow in a Single Fracture

Groundwater flow in a single fracture is often interpreted by assuming the fracture walls are analogous to parallel plates separated by a constant aperture. Using this analogy, solution of the Navier–Stokes equations for laminar flow of a viscous, incompressible fluid bounded by two smooth plates leads to an expression referred to as the cubic law. Figure 17.3 shows the parabolic distribution of velocity predicted by the solution to the Navier–Stokes equations in a cross section of a fracture having parallel walls. A rigorous mathematical development of the cubic law can be found in Bear (1972). Expressed as volumetric flow rate, Q, per unit drop in hydraulic head, Δh, the cubic law can be written as:

$$\frac{Q}{\Delta h} = C(2b)^3 \tag{1}$$
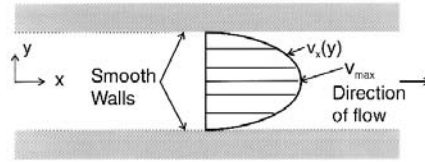
FIGURE 17.3 Schematic diagram showing parabolic distribution of velocity in a fracture having parallel walls.

where $2b$ is fracture aperture and C is a constant related to the properties of the fluid and the geometry of the flow domain. Thus, the flow rate is proportional to the cube of the aperture, hence the term cubic law. For straight uniform flow, the constant of proportionality is:

$$C = \frac{\rho g}{12\mu} \frac{W}{L} \tag{2}$$

where ρ is fluid density, g is acceleration of gravity, μ is kinematic viscosity, W is the width of the fracture, and L is the length of the fracture. For flow in radial domain, C is given as:

$$C = \frac{\rho g}{12\mu} \frac{2\pi}{\ln\left(r_e / r_w\right)} \tag{3}$$

where r_w is well radius and r_e is radius of influence. By substitution of the cubic law into Darcy's law, the transmissivity (T_{fr}) and hydraulic conductivity (K_{fr}) of a fracture can be defined:

$$T_{fr} = \frac{\rho g}{12\mu}\left(2b\right)^3 = K_{fr} 2b \tag{4}$$

Thus, using Equation (4) we can relate terms typically used to describe the permeability of a porous medium to that for discrete fractures.

Figure 17.4 illustrates for a range of hydraulic conductivities, the thickness of porous medium which would be equivalent to a single fracture of given aperture. For example, using Figure 17.4, we see that the volume of groundwater passing through a 10-m section of porous media having $K = 10^{-4}$ m/s is the same as can pass through a fracture less than 1 mm in width under the same driving force. Thus, even in sparsely fractured rock, considerable volumes of groundwater are moved about under typical hydraulic gradients.

The average groundwater velocity in a fracture can be found by combining Equation (4) and Darcy's law:

$$\bar{v} = -\frac{\rho g}{12\mu}\left(2b\right)^2 \frac{dh}{dx} \tag{5}$$

Groundwater velocities predicted using the cubic law over a range of hydraulic gradients are presented in Figure 17.5. Note that even for very small hydraulic gradients, groundwater velocity in discrete fractures is very high in comparison to velocities typical of porous media.

Where large hydraulic gradients are present, turbulent flow may occur. The transition from laminar to turbulent flow is estimated using the Reynolds number, R_e. For a planar fracture R_e is defined as:

$$R_e = \frac{\bar{v} D_h \rho}{\mu} \tag{6}$$
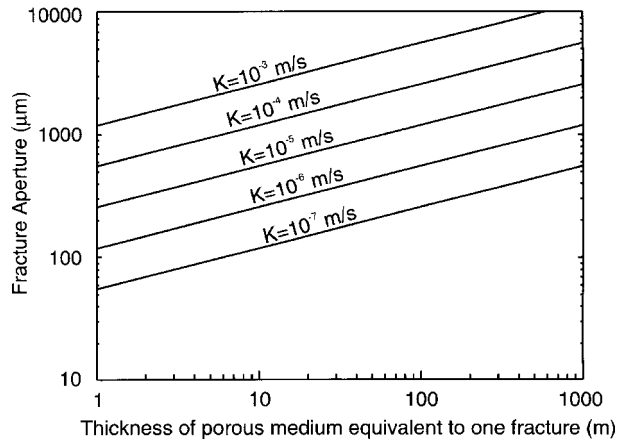
FIGURE 17.4 Comparison of aperture of a single fracture to equivalent thickness of porous media. (From de Marsily, G. 1986. *Quantitative Hydrogeology.* Academic Press, Orlando, FL. With permission.)
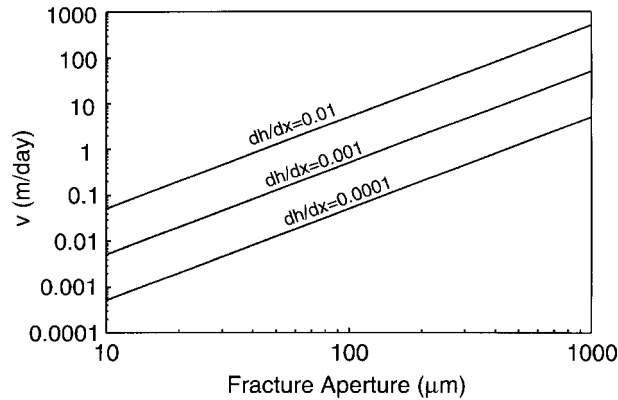
FIGURE 17.5 Groundwater velocity in a single fracture as predicted by the cubic law.

where $\bar{v}$ is mean velocity, and D_h is the hydraulic diameter. For planar fractures, D_h is defined as:

$$D_h = 2 \cdot 2b \tag{7}$$

In addition, the relative roughness of a fracture can be defined as:

$$R_r = \frac{\varepsilon}{D_h} \tag{8}$$

where ε is the mean height of the asperities (the peaks in the topographic surface of the fracture walls). The transition from smooth to rough flow occurs at an R_r of 0.033 and the transition from laminar to turbulent flow at a R_e of 2300 (de Marsily, 1986). Using R_r, the K_{fr} in a rough-laminar fracture can be expressed as:

$$K_{fr} = \frac{\rho g 2b^2}{12\mu\left(1 + 8.8 R_r^{1.5}\right)} \tag{9}$$
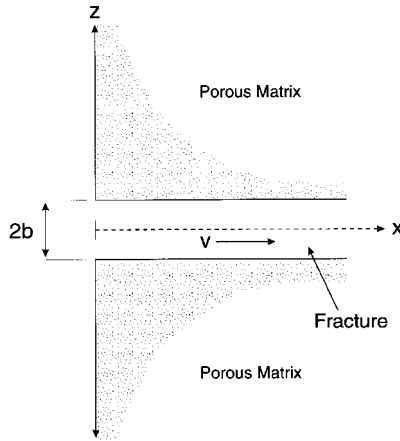
FIGURE 17.6 Schematic diagram of a single fracture in a porous rock.

Alternatively, the influence of fracture surface roughness can be incorporated directly into the cubic law by introducing a factor f that accounts for deviations from the ideal conditions that were assumed in Equation (1). In experimental studies using both radial and straight flow geometries and in fractures of various rock types having apertures ranging from 4 to 250 μm, f was observed to vary from 1.04 to 1.65 (Witherspoon et al., 1980). These results lead to a more generalized form of the cubic law:

$$\frac{Q}{\Delta h} = \frac{C}{f}(2b)^3 \tag{10}$$

where $f = 1$ for smooth walls and f is slightly greater than 1 for rough surfaces. Thus, predictions of groundwater flow based on the cubic law where $f = 1$ are generally adequate for most conditions.

17.3.2 Solute Transport in a Single Fracture

The transport of both conservative and reactive aqueous constituents in discrete fractures is governed by (1) hydrodynamic dispersion, (2) advection, (3) radioactive or biological decay, and (4) diffusion into the rock matrix. In addition, the transport of reactive compounds may be influenced by geochemical processes such as acid/base reactions, oxidation/reduction reactions, precipitation-dissolution, or adsorption-desorption.

The governing equation for solute transport in a single fracture (Figure 17.6) which incorporates advective transport, longitudinal dispersion, molecular diffusion from the fracture into the matrix, adsorption onto the face of the matrix, adsorption within the matrix, and radioactive decay is written as (Tang et al., 1981):

$$2b\left[R_a\frac{\partial c}{\partial t} + \bar{v}\frac{\partial c}{\partial x} - D_L\frac{\partial^2 c}{\partial x^2} + \lambda R_a c\right] + 2q = 0 \qquad 0 \le x \le \infty \tag{11}$$

where the x-coordinate is in the direction of the fracture axis, C is the concentration of solute in the fracture, λ is a decay constant, and q is the diffusive flux perpendicular to the fracture axis. The retardation factor, R_a, is defined by:

$$R_a = 1 + \frac{K_{df}}{b} \tag{12}$$

where K_{df} is the distribution coefficient for the fracture walls defined as mass of solute adsorbed per unit area of surface divided by the concentration of solute in solution. To relate this to the distribution coefficient for the rock matrix, K_d, which is based on the bulk density of the material, an estimate of internal specific surface area of the medium, γ, is required. The specific surface area is defined as the area of the pore walls exposed to the fluid and has units of area per unit mass. Thus, the relation is given as:

$$K_d = \gamma K_{df} \tag{13}$$

The hydrodynamic dispersion coefficient, D_L, is defined following the standard formulation for porous media:

$$D_L = \alpha_L \bar{v} + D' \tag{14}$$

where α_L is the longitudinal dispersivity in the direction of the fracture axis and D' is the effective diffusion coefficient which incorporates the geometric effect of the pathways in the pore space. The governing equation for the matrix as derived by Tang et al. (1981) is:

$$\frac{\partial c'}{\partial t} - \frac{D'}{R'} \frac{\partial^2 c'}{\partial x^2} + \lambda c' = 0 \qquad b \le z \le \infty \tag{15}$$

assuming a single fracture penetrating an infinite medium where c' is concentration in the matrix, and the matrix retardation coefficient R' is defined as:

$$R' = 1 + \frac{\rho_b}{\theta_m} K_d \tag{16}$$

where ρ_b is the bulk density of the matrix, and θ_m is matrix porosity. The diffusive loss term q in Equation (11) represents diffusive mass flux crossing the fracture–matrix interface. Using Fick's first law, this can be expressed as:

$$q = -\theta_m D' \frac{\partial c'}{\partial z} \bigg|_{z=b} \tag{17}$$

Thus, the final coupled equation for the fracture becomes:

$$2b \left[R_a \frac{\partial c}{\partial t} + \bar{v} \frac{\partial c}{\partial x} - D_L \frac{\partial^2 c}{\partial x^2} + \lambda R_a c \right] - 2\theta_m D' \frac{\partial c'}{\partial z} \bigg|_{z=b} = 0 \qquad 0 \le x \le \infty \tag{18}$$

In porous and fractured rock, the loss of solute from the fractures into the rock matrix through the process of diffusion can be significant. To illustrate the effect of matrix diffusion, Figure 17.7 uses the solution to Equation (18) to show the relative concentration of solute in a single 500-μm fracture pervading a medium having a porosity of 1 to 10%. The velocity in the fracture ranges from 1 to 10 m/day and dispersivity is varied from 0.05 m to 0.5 m. Concentration input is continuous and the point of measurement is 50 m from the source. It is immediately evident that the effect of solute loss to the
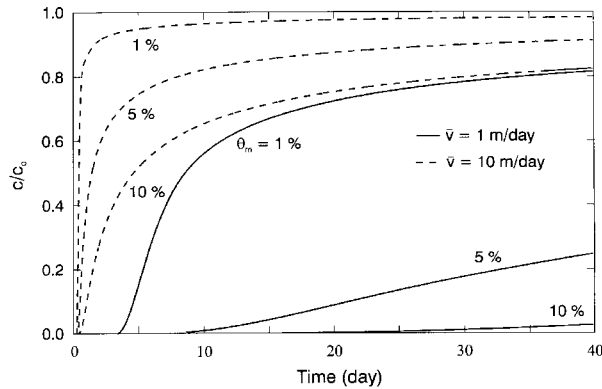
FIGURE 17.7 Concentration profiles in a 500-μm fracture, 50 m from a continuous source, illustrating the influence of matrix porosity on solute transport in a single fracture.

matrix is profound. The effect is so significant that the influence of hydrodynamic dispersion in the longitudinal spreading of solute is completely overwhelmed (i.e., different α_L results in little change in the shape of the curve).

17.4 Field Techniques

In order to develop a defensible conceptual model for a given flow system in fractured rock, some characterization of the medium is necessary. The degree of characterization is dependent on a variety of factors, including the inherent complexity of the fracture system (i.e., sedimentary versus crystalline rock), the type of conceptual model desired, and the amount of funding available. Because of the inherent heterogeneity of fractured rock masses (e.g., presence of sheeting structures), some characterization of the subsurface is essential in supporting conceptual models based only on surface measurements (i.e., fracture maps).

In the following, methods for the measurement of the surface and subsurface expression of fractures will be discussed. In addition, techniques for the measurement of the hydraulic parameters of both the fractures and unfractured matrix are discussed, and an overview of tracer methods used to measure solute transport parameters in fractured rock is presented. Finally, permanent multilevel borehole instrumentation specifically designed for fractured rock investigations is described. An essential component of comprehensive investigations of groundwater flow and solute transport in fractured media is the geochemistry of the groundwater. A discussion of groundwater geochemistry in fractured media, however, is beyond the scope of this chapter.

17.4.1 Fracture Mapping

To determine the distribution of systematic fractures, outcrops, quarry walls, pavements, or mine tunnels are mapped to estimate the strike, dip, and spacing of individual fractures. In practice, reference lines are established on the rock surface using a measuring tape oriented from north. The orientation, spacing, and trace length of all fractures intersecting these scanlines are measured. The bias introduced by the intersection of the fractures with an oriented scanline in this type of sampling must be considered during post-processing (La Pointe and Hudson, 1985). Other useful observations of the fractures on traces include the nature of the termination, mineralization, staining, and surface roughness. Qualitative observations of water or fluid seepage along fractures in road cuts or gorge faces can also be useful indicators of groundwater flow. It is important to note that in many cases surface fracture measurements will be reflective of near-surface stress conditions or may be influenced by the removal of the overlying rock in quarrying operations. Consequently, physical estimation of fracture aperture cannot reliably be con-
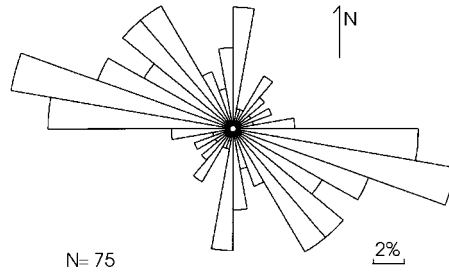
FIGURE 17.8 Rose diagram of orientations of vertical fractures intersecting seven boreholes to depths of 60 m in a flat-lying Silurian dolostone.

Contours of Pole Density (% of total poles)	
0-1	
1-2	
2-3	
3-4	
4-6.8	
Total Poles = 1489	

POLES TO PLANES PROJECTED
ON LOWER HEMISPHERE OF
EQUAL AREA DIAGRAM
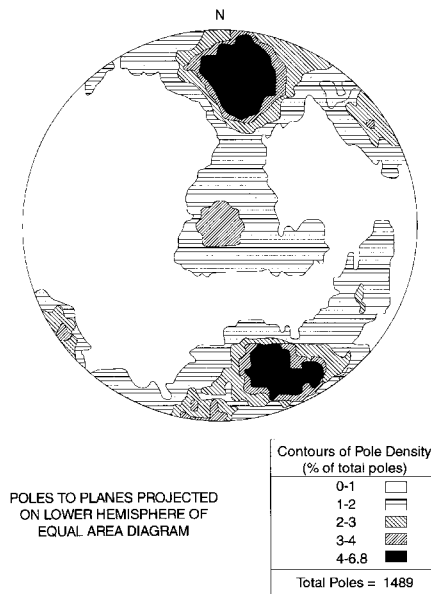
FIGURE 17.9 Equal area stereonet of fracture orientations plotted as poles to planes. (From Raven, K. G. 1986. Hydraulic characterization of a small ground-water flow system in fractured monzononitic gneiss. *Nat. Hyd. Res. Inst. Scientific Series No. 149*, No. 30, 133 pp. Government of Canada, Ottawa. With permission.)

ducted from surface measurements. In addition, due to the presence of sheeting structure, the fracture measurements obtained from surface exposures may not agree with those obtained *in situ*.

Measurements of fracture orientation are usually presented in one of two ways: (1) rose diagrams or (2) stereographic plots. Rose diagrams are histograms in which the frequency (usually as a percentage of total measurements) of joint orientations within 5 to 10° intervals are plotted by azimuth (Figure 17.8). Rose diagrams are commonly used in flat-lying sedimentary rock where most nonhorizontal fractures are vertical in orientation. In crystalline rock or deformed sedimentary rock with fracture sets of more random orientation, data are often presented as the projection of the poles of the planes (fractures) plotted (Figure 17.9) on equal-area stereonets (Ragan, 1973).

17.4.2 Core Analysis

Boreholes installed for groundwater investigations in fractured rock, using either percussion, tri-cone, or diamond drilling techniques, generally range in diameter from 48 mm to 152 mm. In North America, 76-mm or 96-mm diameter boreholes are generally drilled as test or monitoring wells, while 152 mm or

greater diameter boreholes are used for water supply. Rock core obtained from diamond-drilled boreholes can be used to determine the subsurface pattern of fractures. While it is common to install only vertical boreholes, for most hydrogeological site investigations, installing boreholes in the inclined orientation will maximize the number of nonhorizontal and vertical fractures intersected. There is only nominal cost differential between inclined and vertical boreholes, and working in inclined boreholes is no more difficult than in vertical boreholes. In flat-lying sedimentary rock, inclined boreholes may be of better quality and less prone to blockages due to spalling rock.

The physical examination of rock core can be used to identify potentially open fractures through which groundwater flows. Natural fractures can be distinguished from mechanical breaks (breaks resulting from drill vibration, etc.) using the following criteria: (1) roughness of the fracture surface (wavy irregular fracture traces in natural fractures contrast to sharp, clean mechanically induced breaks), (2) presence of infilling (normally, calcite or sulfate mineralization), and (3) in crystalline or deformed sedimentary rock, evidence of shear or slickenside structure. Evidence of weathering or iron staining is indicative of groundwater flow. Examination of the core in detail for lithological changes is important, because discontinuities at geological contacts are frequently water-bearing fractures.

It is important to note that the aperture or width of naturally open fractures cannot be determined from physical measurements on the core. This is because removing core from the subsurface disrupts the natural stress conditions such that aperture values become arbitrary.

If the orientation of bedding is known or the core is oriented using an orientation device, it is a relatively simple procedure to obtain the orientation of nonhorizontal fractures by measuring the arc length of the fracture ellipse in the core (Lau, 1983). Orientation data may be displayed using the same methods as used for surface fractures. Fracture spacing from core can be determined as the spacing between naturally open fractures (eliminating all other fractures). Further details on the analysis of core, particularly in sedimentary rock, are discussed by Kulander et al. (1990).

17.4.3 Borehole Geophysical Methods

Borehole televiewer (BHTV) is one of the more useful borehole geophysical tools for determining the location and orientation of fractures in an open borehole (Zemanek et al., 1969). As the televiewer probe travels the borehole length, a rotating electronic transducer emits a pulsed ultrasonic beam. The reflection of this acoustic energy from the borehole wall produces an image which can be used to identify the strike, dip, and relative size of the fractures intersecting the borehole. The system characterizes fractures only in the immediate vicinity of the borehole. Additionally, it may be difficult to distinguish open fractures from the effects of drilling damage. BHTV surveys are carried out by relatively few groups worldwide and are usually expensive.

A borehole video camera (BVC) survey is a cost-effective alternative to BHTV, particularly for shallow boreholes (<100 m depth). BVCs film the borehole walls and produce either a color or black-and-white image of the borehole. Many of the cameras designed to inspect sewers or pipelines are easy to adapt to a borehole application. Automated depth counters allow for the location of specific features such as fractures, vugs, precipitates, mineralization, changes in lithology or the presence of flowing fractures. Mirrors can be attached to permit a side view of the borehole walls. A BVC survey conducted prior to hydraulic testing or the installation of permanent packer strings will reduce the chances of packer damage or zone leakage due to the inflation of packers over an enlarged portion of the borehole.

Recently, borehole temperature probes and borehole flow meters have been used to detect the presence of large, open fractures using either ambient flow in the borehole (Paillet and Kapucu, 1989) or pumping-induced vertical flow (Hess and Paillet, 1990). In both cases, these methods are sensitive only to the largest one or two fractures and provide little quantitative information about these features. Other geophysical logging tools suitable for fracture detection in open boreholes include: (1) two- or three-arm caliper, (2) formation microscanner, and (3) azimuthal laterlog. Conventional and imaging logs are compared for their ability to detect fractures in Jouanna (1993).

17.4.4 Hydraulic Testing Methods

The characterization of groundwater flow systems in fractured rock is generally dependent on measurements of the permeability of the rock (hydraulic testing). The principal aim of hydraulic testing is to determine the permeability of the fractures and of the rock mass to assist in the estimate of groundwater velocity through the fractures and to appraise groundwater resources. The key to successful field characterization of fractured rock systems lies in the design of hydraulic testing programs. This design should be developed from a preliminary conceptual model of the fracture framework and groundwater flow system. The choice of hydraulic test will depend on several factors including the expected permeability, fracture frequency, matrix properties, type of rock, depth of the borehole, inter-borehole spacing, and purpose of investigation. In this section, constant-head injection tests, and pulse interference testing will be described. In addition, some of the concepts important in the interpretation of hydraulic tests (i.e., wellbore storage and skin effects) will be introduced. A general description of various testing methods applicable to hydrogeological investigations in fractured rock can be found in Doe et al. (1987), and a comprehensive case study is provided by Raven (1986).

17.4.4.1 Hydraulic Testing Equipment

Essential equipment for testing the hydraulic properties of open boreholes in rock includes packers of some type and devices to measure pressure (i.e., pressure transducers). Packers are used to temporarily seal portions of an open borehole to allow for testing in specific zones. In some instances, hydraulic testing is carried out after the installation of permanent borehole instrumentation (discussed later in this section). Typically, for studies conducted at shallow depth (i.e., <100 m) in open boreholes, test zones may range from 2 to 10 m in length with packer seal lengths ranging from 1 to 2 m. In shallow systems, packers are normally inflated with compressed air or nitrogen. In deeper systems or in low-permeability rock, water can be used to provide a stiffer seal. Commercially available packers generally consist of a gland of rubber, steel-reinforced rubber, or neoprene over steel or PVC pipe with either a "fixed-head" or "sliding-head" design. Fixed-head packers rely on the stretch of the gland material during inflation to provide the seal. Conversely, with sliding-head packers, o-rings in the bottom head of the packer permit the gland to slide upward during inflation. This latter design allows for the use of stiffer (i.e., reinforced) materials to inflate against the borehole wall. The fixed-head design is adequate for shallow rock systems with competent borehole walls, while sliding-head packers are more suitable for use in deeper flow systems.

A two-packer system having a single isolated zone is usually adequate for hydraulic testing programs conducted in moderate to sparsely fractured rock. For conditions where fracture frequency is high and packer leakage or short circuiting effects around the main packers are anticipated, additional packers may be installed above and below the test section to monitor for these effects.

For multiwell testing, multiple-packer strings are normally used in the observation boreholes. In some cases, it may be convenient and cost-effective to install permanent packer strings in the observation boreholes (described later in this section).

The pressure response in the isolated source zone or in the observation zones to perturbations generated by injecting or withdrawing water from the source zone is usually measured using pressure transducers or transmitters. The most commonly used types of pressure transducers are based on either strain gauge or vibrating wire membranes. Descriptions of transducer electronics and other geotechnical instrumentation useful to hydrogeological studies can be found in Dunnicliff (1988).

17.4.5 Single-Well Hydraulic Tests

17.4.5.1 Constant-Head Injection Tests (CHITs)

Because open fractures may intersect a borehole at any depth, a continuous measurement of permeability with depth is necessary to locate these features. The most widely used method for testing individual boreholes is the constant-head injection technique. The principle behind a constant-head injection test (CHIT) is to inject or withdraw water at a constant hydraulic head into an isolated portion of the borehole
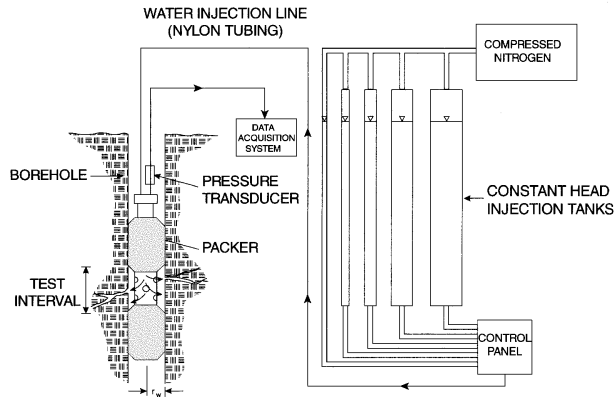
FIGURE 17.10 Schematic diagram of a typical constant-head injection testing system. A series of five tanks is shown on the right while the left shows the downhole instrumentation consisting of a two packer system and pressure transducer.

and measure the flow rate at steady-state conditions (Ziegler, 1976). CHITs may also be referred to as "Lugeon tests," or "constant pressure tests." While a pump and water supply could be used to inject water into the test zone, a series of different diameter tanks fitted with manometers is more commonly used and extends the limits of the testing apparatus to generate a wide range of injection flow rates (Figure 17.10). A range of 10^{-10} to 10^{-3} m²/s in transmissivity, T, can typically be determined using a series of three or more tanks ranging in diameter from 0.01 to 0.3 m. Flow rate is measured by timing the change in water level as viewed through the tank manometers. Compressed nitrogen gas added above the water in each tank ensures a constant flow rate as the tank empties. Flow meters can be used to measure flow rate; however, most flow meters have a very limited measurement range and thus a complex series of flow meters connected by valves would be required to achieve a range in measurement of T similar to the tank system.

Using measurements of flow rate, Q, and the change in pressure expressed as a change in hydraulic head, ΔH, at steady conditions, the T of the tested zone is calculated using the Theim equation:

$$T = \frac{Q}{\Delta H 2\pi} \cdot \ln\left(\frac{r_e}{r_w}\right) \tag{19}$$

The radius of influence, r_e, or outer flow boundary, can generally be assumed to be between 10 and 15 m for moderate values of T (Bliss and Rushton, 1984). While the actual radius of influence is unknown in most field situations, because it appears as a logarithmic term in Equation (19), large errors in estimation of r_e will result in only small errors in the calculation of T. Further discussion of this issue can be found in Doe and Remer (1980).

An equivalent single fracture aperture, $2b_{eq}$, can be determined from the test results by using the cubic law:

$$2b_{eq} = \left(T \cdot \frac{12\mu}{\rho g}\right)^{\frac{1}{3}} \tag{20}$$

The vertical distribution of T or $2b_{eq}$ is determined by systematically testing the length of the borehole in sections using a two-packer system. If possible, two to three increasing increments of pressure should be used during each test. In most studies, injection pressures leading to $\Delta H \leq 10$ m are recommended. Care must be taken not to use pressures that generate fracture dilation or hydro-fracturing. Figure 17.11
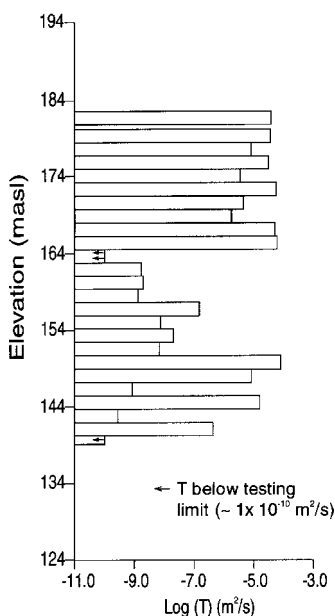
FIGURE 17.11 Typical profile of T with respect to depth obtained from a borehole drilled in Silurian dolostone and tested with constant-head injection tests using a 2-m packer test interval.

shows a typical profile of T with respect to depth obtained from a borehole drilled in Silurian dolostone. T varies by over seven orders of magnitude over 50 m illustrating the hetereogeneity commonly observed in fractured rock.

It is important to note that a single permeable fracture within a test zone will predominate the permeability measurement. Thus, the choice of test zone length will depend on the expected fracture spacing with depth. A too-short test interval will lead to an unnecessarily large number of tests, while a too-long test interval will not capture any of the variability in T and will lead to general overestimates of the permeability of the rock mass. To properly characterize a borehole in detail, an initial survey using a large packer separation length will identify the more permeable zones, while a subsequent second survey using a shorter packer separation length can be used to characterize the properties of individual permeable features and fracture zones (Figure 17.12).

17.4.5.2 Slug Tests

The slug (pulse or bail) test is a transient single-well method that is also frequently used in hydrogeological studies of fractured rock to obtain estimates of the hydraulic properties of a given length of borehole. Type curves based on radial flow in a confined porous aquifer are usually employed to obtain estimates of T and S (Cooper et al., 1967). Although, in principle, this method should provide estimates of T that are more accurate than CHITs, slug tests conducted in fractured rock frequently exhibit results which deviate from the ideal "Cooper et al., 1967" response. This has led to the development of numerous conceptual models which consider the effects of alternate boundary geometries, composite zones, and the other physical conditions more typically found in fractured rock systems. For example, in the presence of a damaged zone around the well, known as wellbore skin, T and S determined from a slug test may reflect only the skin zone (e.g., Faust and Mercer, 1984; Sageev, 1986). Wellbore skin can refer to a zone of enhanced permeability (negative skin) or of reduced permeability (positive skin). Double porosity effects (e.g., Doughtery and Babu, 1984), where both the fractures and matrix contribute to the flow system, can also be incorporated. Composite models are used to account for separate regions surrounding the well which have properties that are different from the formation properties as a whole (e.g., Moench and Hsieh, 1985; Karasaki et al., 1988). Open-hole slug tests are generally limited to zones of $T \geq 10^{-8}$
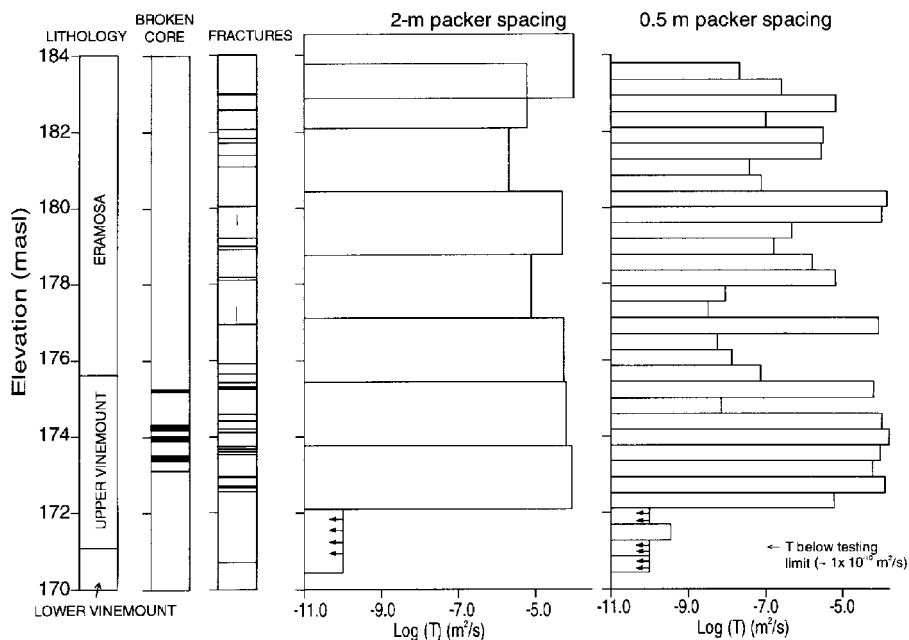
FIGURE 17.12 Results of two testing surveys (2-m and 0.5-m packer intervals) compared to fractures and broken core zones identified by core examination.

m²/s due to the time required to achieve sufficient recovery for interpretation. In lower-permeability material ($T < 10^{-10}$ m²/s) slug tests are conducted under shut-in conditions, reducing the effects of storage in the borehole (Bredehoeft and Papadopolus, 1980; Neuzil, 1982).

17.4.6 Multiwell Hydraulic Tests

17.4.6.1 Interference tests

In the following, the term "interference testing" is used to refer to all types of multiwell hydraulic tests. Interference testing between specific zones in two or more boreholes serves two main objectives: (1) establish the presence of hydraulic connection between the zones and (2) obtain estimates of interwell T and S. In the case of isolated discrete fractures, interwell aperture can also be calculated. In hydrogeological investigations, the choice of interference test method will be influenced by the diameter of the borehole, ability or desire to pump or inject large quantities of water into the formation, and the presence of a water table. Pumping tests conducted in fractured rock are done in a similar manner to those conducted in porous media. The difference lies in the factors influencing the response to pumping. Similar to the case for slug tests, pumping interference tests are influenced by: (1) wellbore storage, (2) skin effects, (3) double porosity, (4) boundary effects (e.g., a linear feature), and (5) vertical or subvertical fractures. Interpretation of interference tests is accomplished by matching the pressure or hydraulic head response in both the source and observation zones to theoretically derived response curves. This is accomplished by either manually matching field data to generated theoretical type curves (e.g., Earlougher, 1977) or using automated algorithms (e.g., Piggott et al., 1996).

17.4.6.2 Wellbore Storage

Fracture apertures usually range from a few microns to a few millimeters. Consequently, the volume of water stored in an open borehole will be several orders of magnitude greater than the volume of water in the fractures intersecting the borehole. This will influence the pressure response measured in the well

during pulse interference and pumping tests. The wellbore storage factor, C_s, for an open well, is defined as πr_c^2 where r_c is the radius of the casing. The dimensionless wellbore storage coefficient, C_D, is defined as:

$$C_D = \frac{C_s}{2\pi r_w S} \tag{21}$$

where S is storativity of the rock formation. In low storativity media (most fractured rock systems), wellbore storage will significantly influence the early time response to a pressure disturbance in either a source or observation well. To accurately measure the hydraulic properties of low storativity media, wellbore storage must either be eliminated or accounted for in the interpretation of transient multiwell tests.
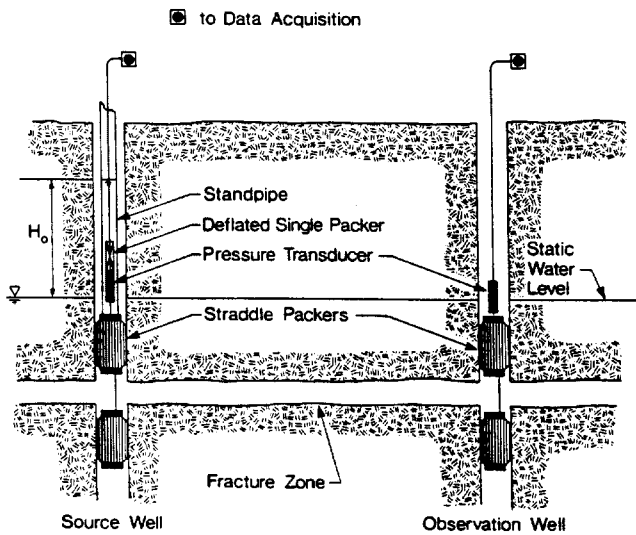
FIGURE 17.13 Schematic diagram showing field set-up for pulse interference tests. (From Novakowski, K. S. 1989. Analysis of pulse interference tests. *Water Resour. Res.* 25(11):2377-2387. With permission.)

17.4.6.3 Pulse Interference Tests

Single-pulse interference tests are an easy yet powerful extension to single-well slug tests. In this case, the pressure response in one or more observation zones due to the introduction of an instantaneous slug of water in a nearby source well is measured (Figure 17.13). A pulse interference test can be analyzed quickly using a graphical method (Novakowski, 1989). To do so, the magnitude of peak response and time lag in the observation zone must be measured. Peak response is defined as the ratio of the maximum rise in hydraulic head at the observation well (Δh) divided by the initial rise in hydraulic head in the source well (H_o). Time lag (t_l) is the elapsed time between the start of the slug test in the source well and the maximum peak response in the observation well. The remaining required dimensionless parameters (r_D, t_D), and dimensionless wellbore storage coefficients in both the observation and source well (C_{DO}, C_{DS}) are defined as follows:

$$r_D = \frac{r}{r_w} \qquad C_{DS} = \frac{C_s}{2\pi r_w^2 S} \qquad C_{DO} = \frac{C_o}{2\pi r_w^2 S} \qquad t_D = \frac{Tt_L}{r_w^2 S} \qquad \Delta h_{DO} = \frac{\Delta h}{H_o} \tag{22}$$

where r is the distance between the source and observation boreholes, and C_s and C_o are the storage coefficients in the source and observation boreholes, respectively. One of two sets of graphs is used to

determine T and S using the dimensionless variables. Set 1 (Figure 17.14) is used when there is no observation well storage ($C_{DO} = 0$), whereas Set 2 (Figure 17.15) is used when the storage capacity in the source well is equal to the storage capacity in the observation well ($C_{DO} = C_{DS}$). First, using h_{DO} and r_D, a value for C_{DS} is determined from the graphs in part a, and S is calculated from the definition of C_{DS} in Equation (22). Second, using h_{DO} and r_D, a value for t_D is determined from the graphs in part b, and T is calculated from the definition of t_D in Equation (22). Figure 17.16 shows an example of both source and observation zone responses to a typical pulse interference test where $r = 22$ m, $T = 5 \times 10^{-4}$ m²/s, and $S = 4 \times 10^{-5}$.

17.4.7 Point Dilution Method

Measuring groundwater velocity in discrete fractures is often difficult because of the uncertainty associated with estimates of low hydraulic gradient (i.e., hydraulic head may vary by only centimeters over distances of many meters). The point dilution method may be employed to determine direct measurements of groundwater velocity. This method is based on the decay in concentration with time of a mixed tracer in a single well due to dilution caused by the natural groundwater flow through the borehole (Drost et al., 1968; Grisak et al., 1977). To conduct a point dilution experiment, a small section of the borehole having one or more fractures is isolated using a set of two pneumatic packers. To minimize the duration of the experiment, the mixing volume must be reduced by using a short spacing between the two packers. The experiment is initiated by instantaneous injection of a small volume of conservative tracer into the test zone, and mixing is continued throughout the duration of the experiment. Figure 17.17 shows schematically a typical arrangement for a point dilution experiment. Suitable conservative tracers include bromide or chloride, some fluorescent dyes (e.g., Lissamine FF), radioactive isotopes (e.g., tritium), or stable isotopes (e.g., deuterium).

The decay in concentration is interpreted using (Drost et al., 1968):

$$\frac{dc}{dt} = -Av_a \frac{c}{V} \tag{23}$$

where A is the cross sectional area available to flow, v_a is the apparent velocity of groundwater flowing through the wellbore, and V is the volume of the sealed-off portion of the borehole in which the dilution occurs. The solution to Equation (23) for the case of instantaneous injection and relating the apparent velocity in the test section to the true formation velocity, v_f, yields (Drost et al., 1968):

$$v_f = \frac{V}{\xi At} \ln \frac{c}{c_o} \tag{24}$$

where c_o is the initial concentration at $t = 0$, c is the concentration at time t after the tracer was injected, and ξ is a dimensionless correction factor accounting for additional flow captured by the open well due to the convergence of flow lines in the neighborhood of the wellbore. Groundwater velocity in a single fracture intersecting the test zone (v_f) is determined from the measurements of dilution of tracer in the sealed off portion of the borehole by plotting the results in the form of ln c vs. t and fitting a linear regression line to the data (Figure 17.18). Using the time (t_m) at $c/c_o = 0.5$ from the regression line, Equation (24) reduces to:

$$v_f = \frac{0.693V}{\xi At_m} \tag{25}$$
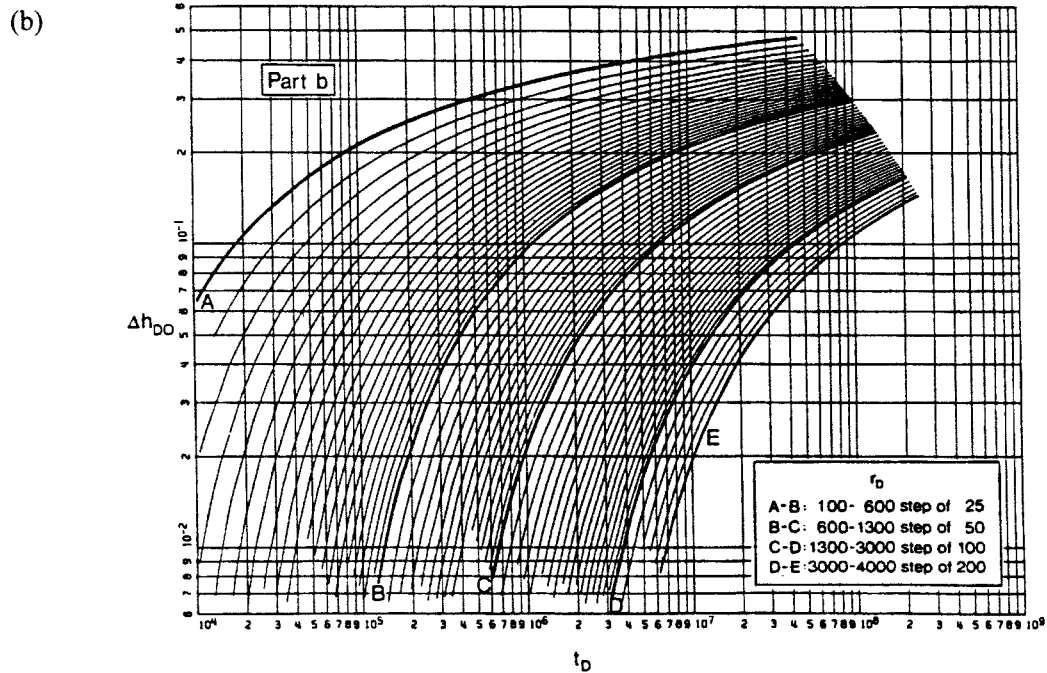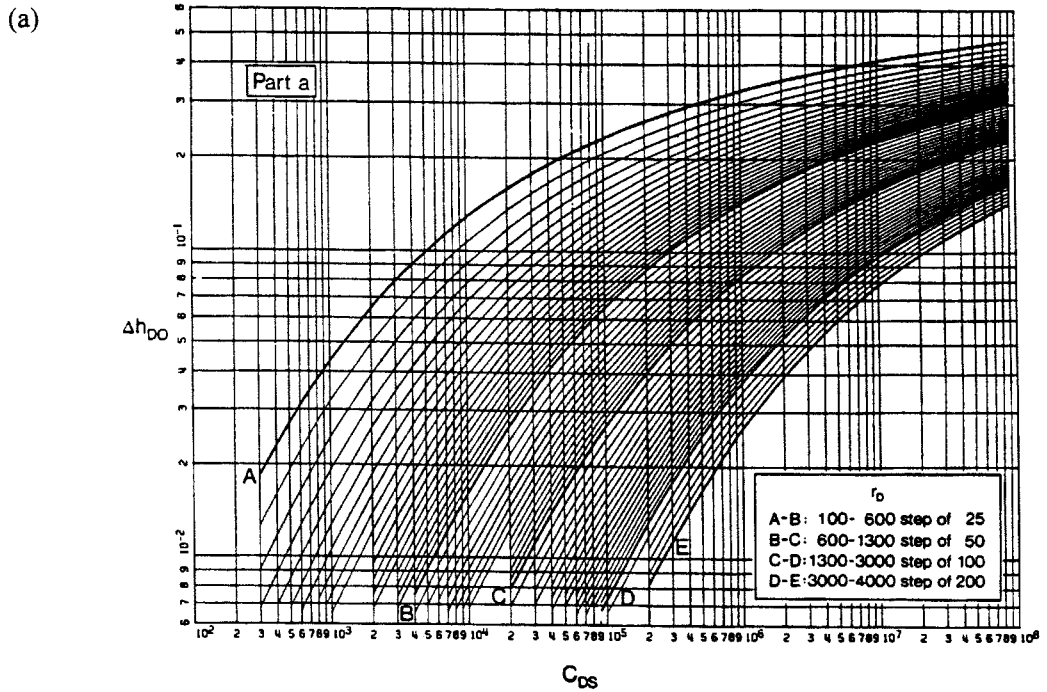
FIGURE 17.14 Type curves used to interpret pulse interference tests by graphical method when there is no observation well storage (a) curves used to calculate S and (b) curves used to calculate T. (From Novakowski, K. S. 1989. Analysis of pulse interference tests. *Water Resour. Res.* 25(11):2377-2387. With permission.)
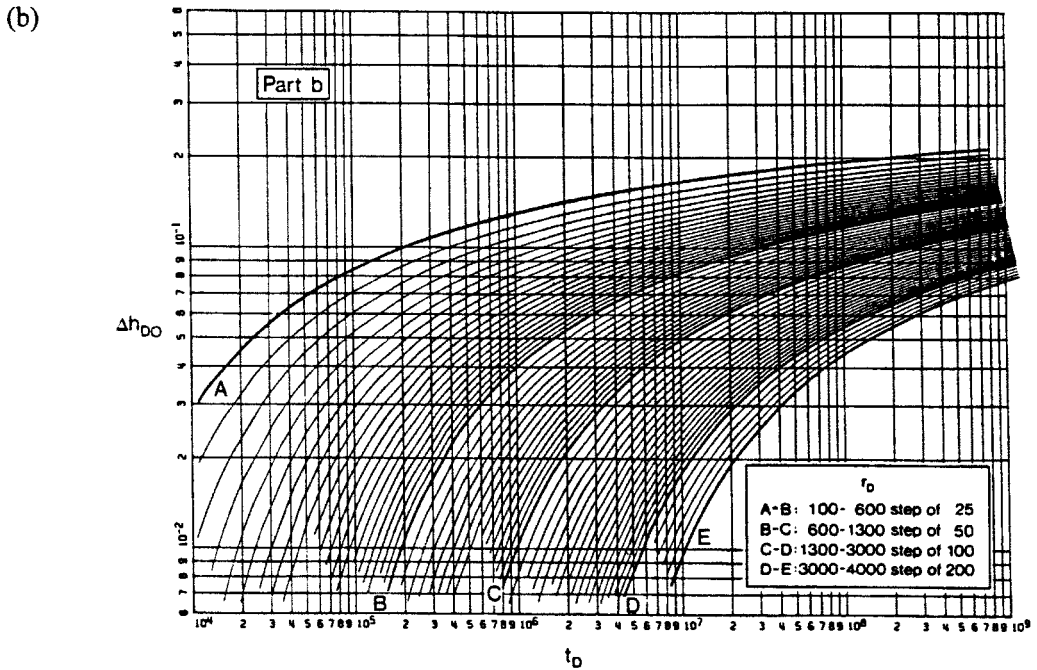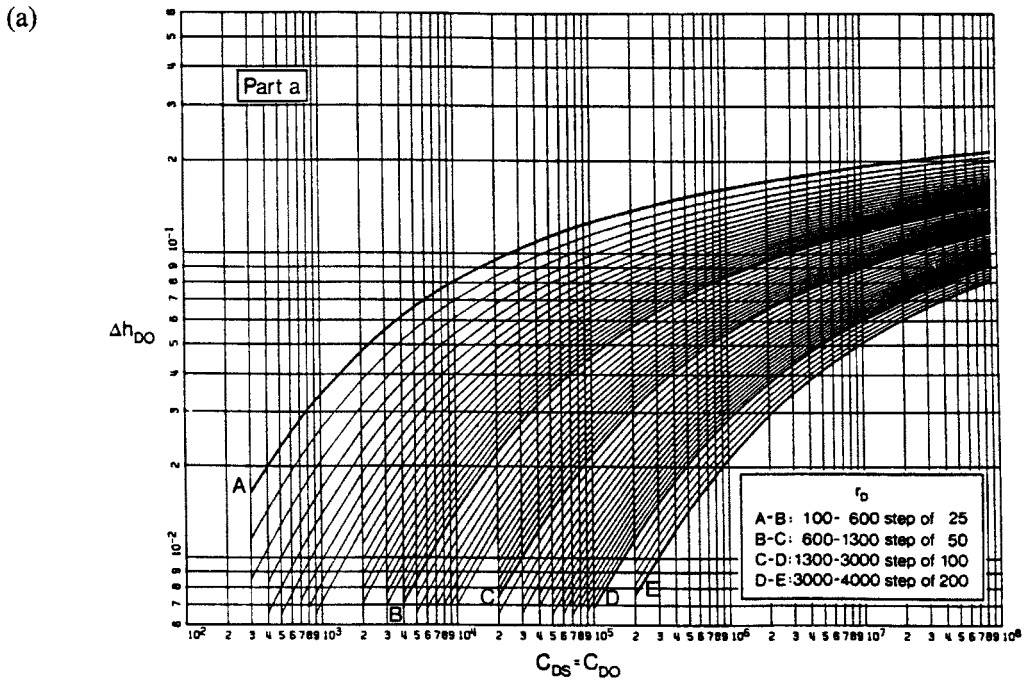
FIGURE 17.15 Type curves used to interpret pulse interference tests by graphical method when observation well storage = source well storage (a) curves used to calculate S and (b) curves used to calculate T. (From Novakowski, K. S. 1989. Analysis of pulse interference tests. *Water Resour. Res.* 25(11):2377-2387. With permission.)
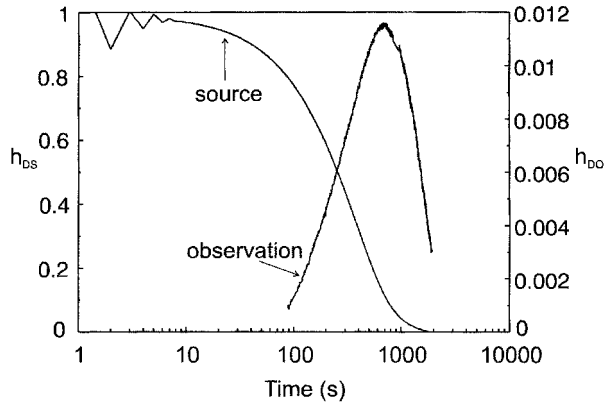
FIGURE 17.16 Example of source well pulse and observation zone response. Test was conducted in a 500-μm fracture where the observation zone was approximately 22 m from the source zone.
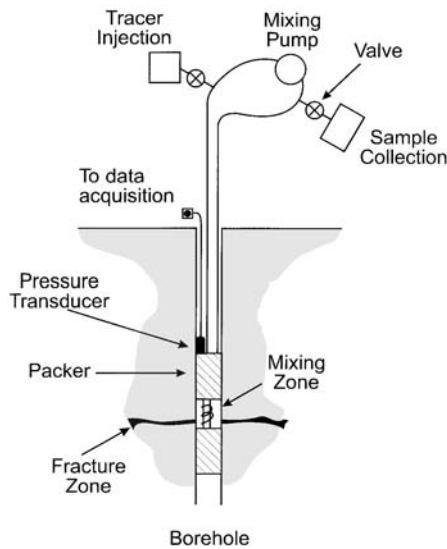
FIGURE 17.17 Schematic diagram of the experimental apparatus used for a point dilution.

In the case of a zone containing a single fracture, $A = 1/2$ the circumference of the borehole times the hydraulic aperture and $\xi = 2$. Where more than one fracture intersects the borehole, the interpretation is more complicated. To accurately estimate velocity, the aperture of each individual fracture should be known, otherwise the result will represent an average.

17.4.8 Tracer Experiments

Tracer experiments conducted in fractured rock provide a means of determining solute transport parameters (i.e., velocity, dispersivity, and matrix porosity) of individual fractures or fracture zones at the field scale. Similar to tests in porous aquifers, tracer experiments conducted in fractures can be carried out using either forced hydraulic gradient or natural gradient flow conditions. Because of the difficulty and cost associated with natural gradient tracer experiments, very few of these have been conducted and no further discussion of this method will be presented here. In the following, the common methodologies used to carry out multiwell tracer experiments under conditions of forced gradient are outlined.
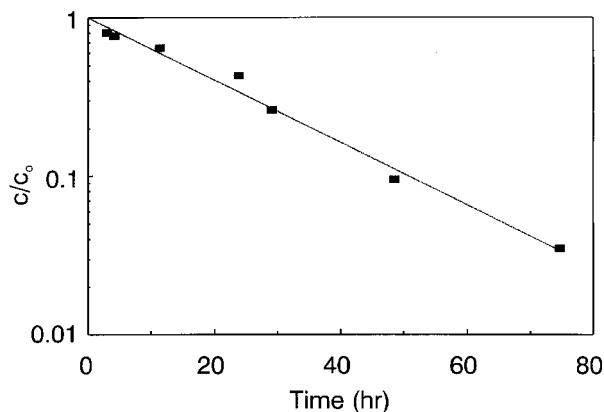
FIGURE 17.18 Example point dilution response curve. In this test a velocity of 4 m/day was measured in a 220-μm fracture.

The three flow field geometries which are commonly used when conducting tracer experiments under forced gradient in fractured rock are: (1) injection-withdrawal, (2) radial convergent, and (3) radial divergent. In the injection-withdrawal format, an artificial steady-state flow field between two wells is established where the injection flow rate is equal to the withdrawal flow rate (Figure 17.19a). Tracer is introduced either as a finite slug or continuously into the injection well, and the concentration is monitored at the withdrawal well. An example of this type of experiment can be found in Novakowski et al. (1985). The water removed from the withdrawal well can be re-circulated back into the system via the injection well, but this will complicate the interpretation of the test. The advantage to this test method is the development of a well controlled flow field where recovery of the tracer should approach 100%. A radial-convergent experiment is conducted by passively injecting tracer into one borehole and withdrawing by pumping at a second borehole (Figure 17.19b). Multiple source boreholes can be used by introducing a different tracer in each. An example of the use of this experimental method is outlined in Shapiro and Nicholas (1989). Radial-divergent experiments are conducted by injection of tracer in a single well and passively monitoring the tracer arrival in one or more observation wells (Figure 17.19c). An example of this experiment is presented in Novakowski and Lapcevic (1994). The choice of methodology will depend on the objectives of the experiment, the number of boreholes available for study, the ease of pumping and sampling, and the type of tracer used.

To date, most forced-gradient experiments have been conducted in discrete fractures or fracture sets in which the geometry of the fracture is well known (e.g., Raven et al., 1988; Shapiro and Nicholas, 1989; Abelin et al., 1991). The scale of these experiments has been limited to interwell travel distances of 50 m or less. In future, tracer experiment methodology should prove very useful in confirming the geometry of fracture intersections, the presence of which may be surmised from the results of interwell hydraulic tests. Interwell distances for tracer experiments will likely increase.

17.4.9 Borehole Instrumentation

Permanent completion of boreholes drilled in bedrock is generally carried out in one of two ways: (1) using technology designed for piezometer construction in unconsolidated porous media or (2) through the use of multilevel borehole casing. Only the latter will be discussed in this chapter; the former is covered in discussions of well completions for unconsolidated porous media. Multilevel completions are primarily designed to: (1) obtain measurements of hydraulic head and (2) obtain representative samples of groundwater. This is accomplished by isolating sections of the borehole using a series of permanent packers joined by casing or riser pipe. Access to the isolated zones is through manometers or access ports. Currently, two commercial systems of multilevel completions are widely used around
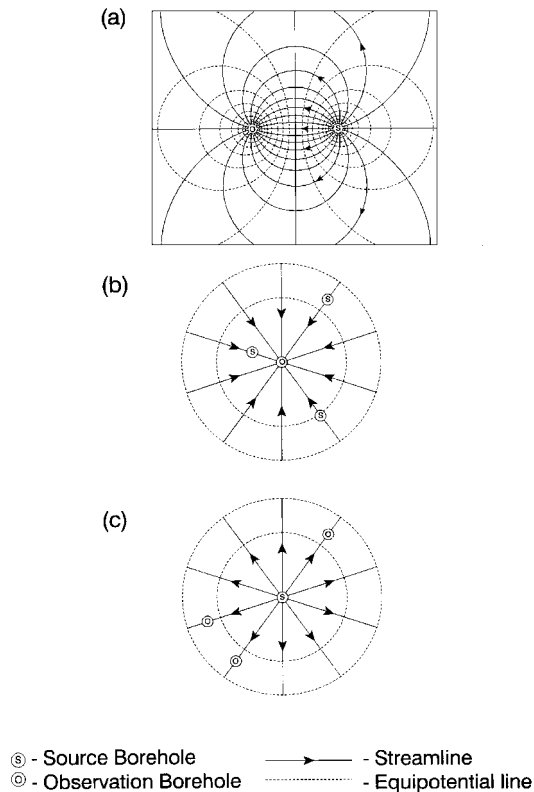
FIGURE 17.19 Flow field geometry commonly used for tracer experiments: (a) injection-withdrawal, (b) radial convergent, and (c) radial divergent.

the world. These are the Solinst® system (Cherry and Johnson, 1982) and the Westbay® system (Black et al., 1987). The advantage to multilevel casing is that numerous distinct zones may be accessed in a single borehole.

The Solinst system (Figure 17.20) is based on packers that are filled with a sealant chemical that is activated by the introduction of water into the casing after the complete casing string has been lowered into the borehole. Sampling ports and manometers of thin nylon or Teflon® tubing provide access to each sealed-off zone to measure water levels and withdraw groundwater samples. Triple-tube sampling pumps can either be installed in each zone or lowered into the manometers. These pumps allow sampling in zones where pumping from the surface is not feasible. Additionally, pressure transducers or transmitters can be included in each zone to electronically measure water pressure. Depending on the objectives of the study and specific site conditions, the modular instrumentation system can be customized to maximize the amount of data obtained. The number of zones sampled or measured is limited by the diameter of the borehole (and thus the diameter of the casing) as each transducer, pump, or sampling port requires access to the surface through tubing or electrical cable.

The Westbay system is also a modular design having water-filled packers connected by casing elements and specially designed pumping and measurement/sampling ports (Figure 17.21). Water pressure is measured and representative groundwater samples are obtained using a submersible probe which is lowered into the casing and connected to an electronic data acquisition device on the surface. The probe has a small arm which is used to locate each measurement port inside the casing. Once positioned on the port, a mechanical foot on one side of the probe is activated and presses against the inner casing wall. This causes an o-ring on the opposite side of the probe to seal around a ball bearing and exposes the pressure transducer in the probe to water pressure outside the casing. Groundwater samples can be
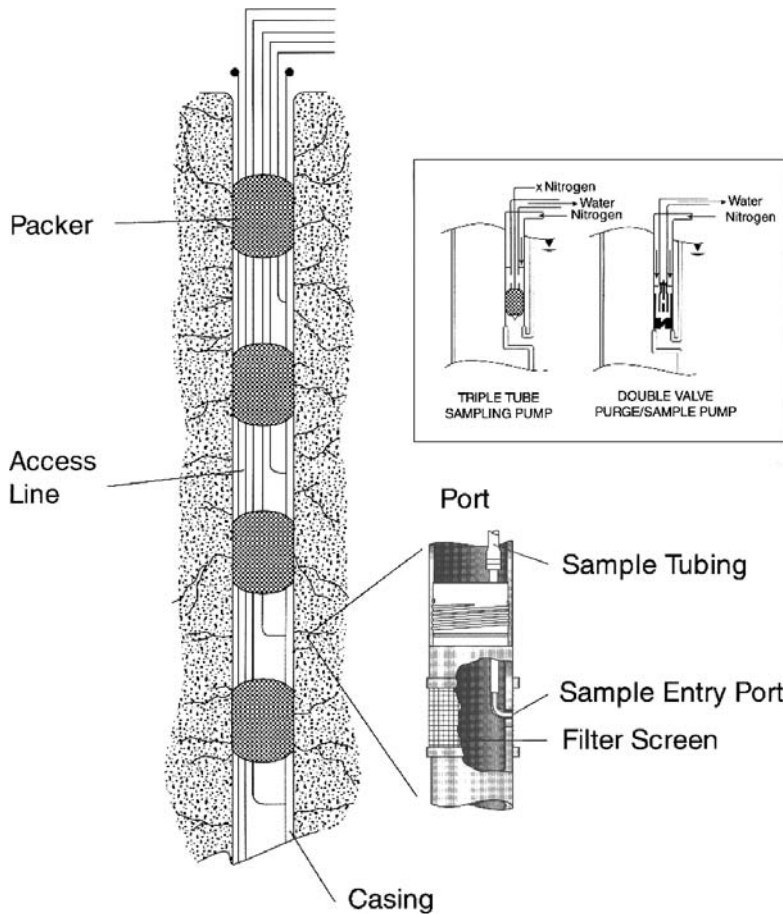
FIGURE 17.20 Components of the Solinst® system of permanent multilevel instrumentation for use in the rock boreholes. Access lines may be electrical cables for pressure transducers installed in the ports or tubing to sample the isolated zones. Pumps (triple tube or double valve types) may be dedicated to specific zones or lowered into the sampling tubing as required. Manual water level measurements are obtained through the tubing.

obtained using the same port by opening a valve from the surface unit and allowing water to fill a sample container attached below the probe (Figure 17.21).

Additionally, in some instances, hydraulic tests can be conducted after permanent multilevel completions have been installed. For example, a borehole instrumented with several zones, used as an observation well during a pumping test can provide information on the vertical connection of fracture zones. To maximize the data obtained from permanently instrumented boreholes, care must be taken in designing appropriate isolated zones. Careful examination of rock core, geophysical surveys, and hydraulic testing results prior to installation of the instrumentation will ensure that the isolated zones appropriately define the three-dimensional nature of the groundwater flow system.

17.5 Conceptual Models

In formulating an understanding of groundwater flow and solute transport for a given site or region at which flow is dominated by the presence of fractures, it is important to develop a reasonable conceptual model as a starting point. In many cases, the conceptual model may incorporate elements of the flow system at a variety of scales. This may include processes at the microscopic scale, such as matrix diffusion,
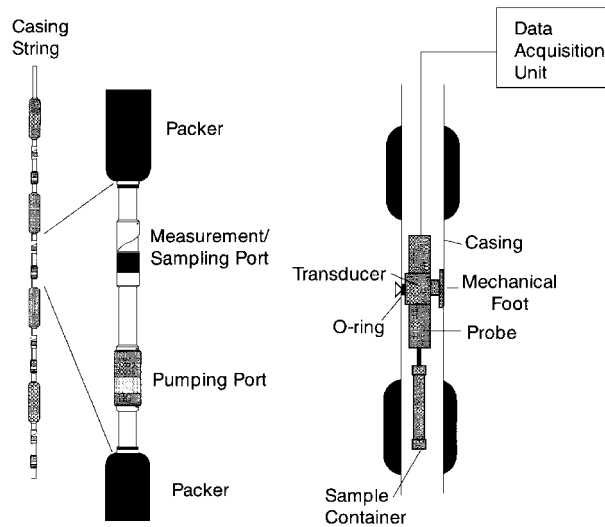
FIGURE 17.21 Components of the Westbay® system of permanent multilevel instrumentation for use in the rock boreholes. Note that no tubing or electrical cable is permanently installed in the casing string since the probe is lowered to each zone to either measure water pressure or obtain a groundwater sample.

and processes at the single-continuum scale, such as recharge or discharge. In the following section, a variety of conceptual models for flow and transport at the scale of a discrete fracture and a fracture network are discussed. In all cases, it is assumed that the cubic law, hydrodynamic dispersion, and matrix diffusion apply at the microscopic scale.

Bear (1993) defined four operational scales in a fractured medium, at which different conceptual approaches might be applied (Figure 17.22). At the very-near field scale, flow and solute transport are dominated by a single fracture. At the near field, flow and transport are dominated by a few well-defined fractures, and interaction between the fractures and the matrix may play a role. At this scale, discrete fracture models may be used in which the major fractures are defined deterministically and the minor fractures are specified stochastically. Depending on the type of fractured rock, it is likely the near-field scale that is of most interest to practicing hydrogeologists. At the far-field scale, multiple continua are defined using at least one continuum for the unfractured matrix, possibly one for the minor fractures and another for the major fractures. At the very-far-field scale, a single continuum can be applied. In particular cases, it may be found that a mixture of conceptual approaches is required. For example, in some sedimentary rock environments, sheeting fractures may predominate in one stratigraphic horizon, while in another the fracturing is so frequent as to warrant the use of a continuum or a multicontinuum approach.

17.5.1 Conceptual Models for a Single Fracture

Numerous studies have been conducted in which the roughness of natural fracture walls and the distribution of fracture aperture have been measured (e.g., Brown et al., 1986; Gentier and Billaux, 1989; Piggott, 1990). These studies have been conducted on fracture samples ranging in scale between 0.1 and 1 m. The results indicate that the surfaces of fracture walls can be rough and undulating with numerous contact points between the walls. In general, it is concluded that (1) fracture aperture may fit into one of several statistical distributions, (2) aperture is correlated spatially, and (3) aperture distributions are scale invariant.

The distribution of apertures at the laboratory scale has been observed to follow either a log-normal distribution (Gentier and Billaux, 1989; Hakami, 1995), a Gaussian distribution (Piggott, 1990; Brown, 1995; Hakami, 1995), or a gamma distribution (Tsang, 1984). Additionally, spatial distributions of
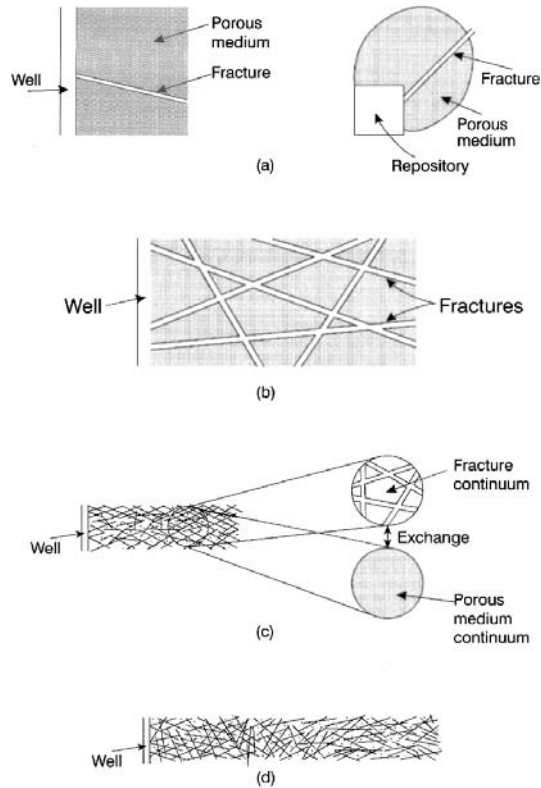
FIGURE 17.22 Operational scales in a fractured medium: (a) very-near field, (b) near field, (c) far field, and (d) very-far field. (From Bear, J. 1993. Modeling flow and contaminant transport in fractured rocks, in *Flow and Contaminant Transport in Fractured Rocks*. Eds. J. Bear, C.-F. Tsang, and G. de Marsily. Academic Press, San Diego, 1-37. With permission.)

aperture are often observed to have a large variance, indicating a high level of heterogeneity in the fracture plane (e.g., Hakami, 1995). The roughness of fracture surfaces has also been determined to display fractal properties implying that variable aperture distributions may be scale invariant (e.g., Brown et al., 1986; Wang et al., 1988).

Fractures with similar statistical distributions of aperture may have very different flow properties due to the spatial relation of the aperture variations. Fracture surfaces have been determined to be correlated spatially with the degree of correlation related to the scale of measurement. Brown et al. (1986) determined correlation lengths from between 0.5 and 5.0 mm for sample lengths and diameters of 25.4 mm. Vickers et al. (1992) studied surface roughness of a single fracture in a block of welded tuff (0.15 × 0.40 m) and found the resulting aperture distributions close to normal except at the tails. The apertures were also found to be correlated at two scales, one on the order of millimeters and the other on the order of tens of centimeters. In addition, Vickers et al. (1992) found that apertures increased consistently along the entire length of the fracture, suggesting that a third spatial correlation occurs at a scale well in excess of the sample dimensions. Hakami (1995) compiled experimental results of aperture measurements at the lab scale and concluded that both the variance and spatial correlation length increase with increasing mean aperture.

It is also well recognized that the regions in which the fracture surfaces are in contact and closed to water flow will strongly influence flow and transport. Experiments conducted by injecting wood's metal into fractures in granitic cores have shown contact areas ranging from 8 to 15% with a normal stress of 3 Mpa (Pyrak-Nolte et al., 1987). At the field scale, hydraulic aperture measurements in two fractures at

depths less than 15 m in sedimentary rock (shale/limestone) in an area roughly 30 m × 30 m, indicate fracture closure between 20 and 30% (Lapcevic et al., 1990). Smooth fractures will display many small areas of contact with complex outlines resulting in a complex distribution of fluid flow among many small channels. In contrast, rough fractures display fewer, larger contact areas with smoother outlines concentrating flow in a few large channels (Odling, 1994). Tsang (1984) observed that increases in contact area results in a reduction in mean aperture, an increase in tortuosity, and a decrease in the connectivity of the fluid flow paths. Tortuosity factors determined through comparisons of measured hydraulic versus physical apertures have been observed to be a good approximation to account for closure (e.g., Piggott and Elsworth, 1993).

Experimental studies of flow and transport in laboratory-scale fractures have been used to investigate the relationship between measured apertures and resultant flow characteristics within the fracture. Results of flow experiments on fractures in laboratory studies with known aperture distribution show that the ratio between mean aperture and hydraulic aperture is 1.1 to 1.7, for 0.1 < 2b < 0.5 mm (Hakami, 1995). The lower value of the hydraulic aperture is to be expected since the variation in aperture of natural fractures forces the flow to be tortuous. It is well recognized that fluid flow in natural fractures will be controlled by the distribution of asperities leading to tortuous paths and areas in which no flow occurs.

There are several conceptual models for transport in a fracture of variable aperture. For example, the "channel model" for flow and transport through a single fracture is based on a series of one-dimensional channels of constant aperture oriented in the direction of flow (Tsang and Witherspoon, 1983). Tsang and Tsang (1987) extended this model to include a limited number of tortuous and intersecting channels each characterized by an aperture density distribution, effective channel width, and correlated in length and aperture (Figure 17.23). A gamma function was used to characterize the aperture distribution and standard geostatistical techniques used to calculate distributions. Johns and Roberts (1991) presented a two-aperture channel model in which lateral transfer of mass from large to small aperture regions of the fracture plane and vertical diffusion into the rock matrix were considered. Moreno et al. (1988) presented a stochastic model in which flow and transport in the entire fracture plane were considered. The aperture of the fracture was lognormally distributed and possessed a spatial correlation length (λ). From this it was concluded that a broad distribution of apertures with spatial correlation length on the same order of magnitude as the scale of measurement is responsible for flow channeling phenomena.
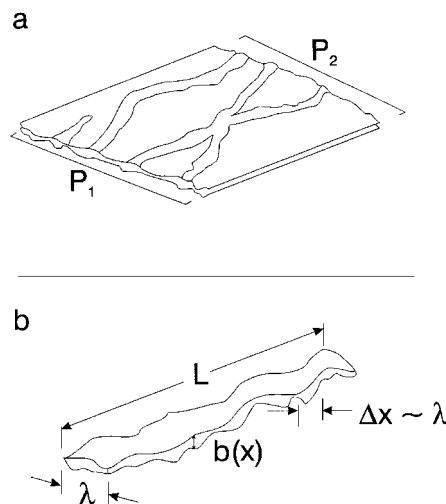
FIGURE 17.23 (a) Channel representation of fluid flow in a single fracture, and (b) schematic sketch of one channel, where b(x) is the aperture distribution, and λ is the spatial correlation length of the distribution. (From Tsang, Y. W. and Tsang, C. F. 1987. Channel model of flow through fractured media. *Water Resour. Res.* 23(3):467-479. With permission.)

Determining the spatial distribution and correlation of fracture aperture at the field scale is difficult because of the extreme cost required to obtain sufficient field data. Consequently, conceptual models of variable aperture based on surface roughness are well verified at scales up to a meter but have yet to be verified with field data at scales in the tens to hundreds of meters. Thus, we are left with using simple correction factors such as macroscopically defined tortuosity to account for this variability.

17.5.2 Conceptual Models for a Fracture Network

In all rock types, fractures of various sizes and lengths combine to form three-dimensional networks of interconnected groundwater pathways. At scales less than that defined for a single continuum, determining an appropriate conceptual model for the three-dimensional arrangement of fractures is essential in understanding and predicting groundwater flow and solute transport in fractured rock. In general, a distinction must be made between conceptual network models for sedimentary rock and those for crystalline rock. In most sedimentary rock environments, the directions of the principal fractures follow bedding planes and other pre-existing planes of weakness related to deposition. These fractures are often connected by near-orthogonal fractures whose genesis is related to paleo- and neotectonic stresses (Holst, 1982; Williams et al., 1985). For flat-lying stratigraphy, this results in a relatively simple fracture framework which can be conceptualized by several discrete horizontal fractures connected by statistically defined sets of subvertical fractures. Figure 17.24 illustrates a typical three-dimensional view of a conceptual fracture framework for a layer-cake stratigraphy. Note that the vertical fractures are constrained by the bedding and have a preferred orientation. There are exceptions to this scenario, however, for sedimentary environments which have undergone considerable deformation. In such cases, the stratigraphy tends to be inclined and folded. This concentrates vertical fractures in the vicinity of fold axes and leads to shearing movement on some bedding planes. Other salient features of the conceptual model for flat-lying stratigraphy are retained, however.
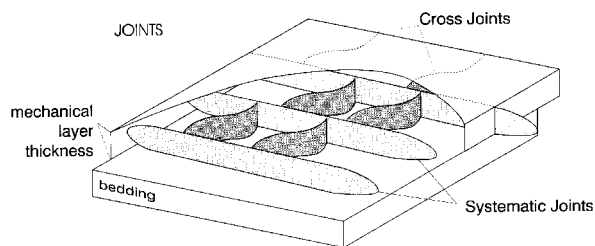
FIGURE 17.24 Typical three-dimensional view of conceptual fracture framework for a layer-cake stratigraphy. (From Engelder, T., Fischer, M. P. and Gross, M. R. 1993. *Geological Aspects of Fracture Mechanics, A Short Course Manual*, Geological Society of America, Boulder, CO. With permission.)

Conceptual models for crystalline rock environments are usually more complex than that for sedimentary rock. In uniform granitic rock, fractures may be oriented in preferred directions according to magma emplacement and cooling (polygonal fracture sets), local deformation, regional deformation, and erosional unloading (sheeting fractures).

One of the earliest conceptual models for flow in crystalline rock was based on the assumption of an equivalent porous medium (Snow, 1969). The concept was developed to relate the results of permeability measurements obtained from boreholes to a hydraulic conductivity ellipsoid which defines the anisotropic permeability of the bulk rock. To conduct the interpretation, it was assumed that flow was predominated by the fractures, all fractures contributed to the flow system, and the orientation of the fractures intersecting the boreholes was known. The hydraulic conductivity ellipsoid was constructed by summing the permeability contributions of each fracture intersecting each test interval.

It has long been recognized that fractures form disc-shaped discontinuities, particularly in granitic environments. Figure 17.25 shows a three-dimensional conceptual model in which disc-shaped fractures

FIGURE 17.25 Three-dimensional conceptual model of disc-shaped fractures which are of random size and oriented orthogonally. (From Long, J. C. S., Gilmour, P., and Witherspoon, P. A. 1985. A model for steady fluid flow in random three-dimensional networks of disc-shaped fractures. *Water Resour. Res.* 21(8):1105-1115. With permission.)

are of random size but oriented orthogonally. This concept can be extended to the general case for more random orientations of discs. For example, Cacas et al. (1990) and Dverstorp et al. (1992) have developed numerical flow and transport models based on a three-dimensional distribution of discs randomly placed in space and of random radius, but given a specific spatial orientation to simulate fracture sets. Ground-water flow and solute transport were then assumed to follow one of three types of linear flow channels which interconnect the individual fractures in the network (Figure 17.26).

Recently, scaling relations have been introduced into these types of conceptual models through the use of fractals, thus providing a method of using field data in the generation of artificial fracture networks. Figure 17.27 illustrates the generation of a three-dimensional network of fractures based on fracture traces measured from outcrop scan lines, air photographs, and geophysical data (Piggott et al., 1997). From the trace data (Figure 17.27a) statistically generated two-dimensional traces are extended to produce the three-dimensional network of fractures (Figure 17.27b). A detailed three-dimensional view of a portion of the fracture network in Figure 17.27b is shown in Figure 17.27c.

17.6 Modeling Flow and Transport in Fractures and Fracture Networks

In some cases, such as where sheeting fractures dominate the groundwater flow system, it may be necessary only to model these discrete features and not the flow system as a whole. Simple analytical models for flow and transport in a single fracture (e.g., Tang et al., 1981) or a set of parallel fractures (Sudicky and Frind, 1982) may suffice. For example, transport calculations such as those shown in Figure 17.7 may provide the information necessary to illustrate the potential importance of matrix diffusion at a given site.

Nonconstant aperture in the fracture plane, as discussed in the previous section, will influence solute transport in single fractures. Direct solution of the flow and transport in a fully defined aperture distribution can be conducted using numerical models (e.g., Lapcevic, 1997). For example, Figure 17.28 compares a tracer plume created assuming a constant aperture to one simulated assuming a spatially variable aperture. Note that the variability in aperture leads to a plume of much greater irregularity in
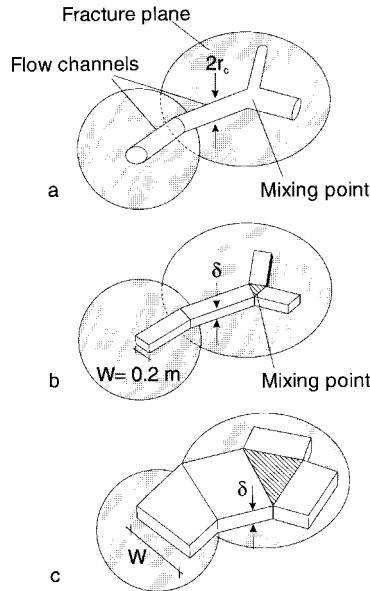
FIGURE 17.26 Three types of interconnecting flow channels in a fracture network: (a) tubular flow channels, (b) parallel plate channels with constant width, and (c) parallel plate channels with width equal to the fracture intersection line. (From Dverstorp, B., Andersson, J., and Nordqvist, W. 1992. Discrete fracture network interpretation of field tracer migration in sparsely fractured rock. *Water Resour. Res.* 28(9):2327-2343. With permission.)
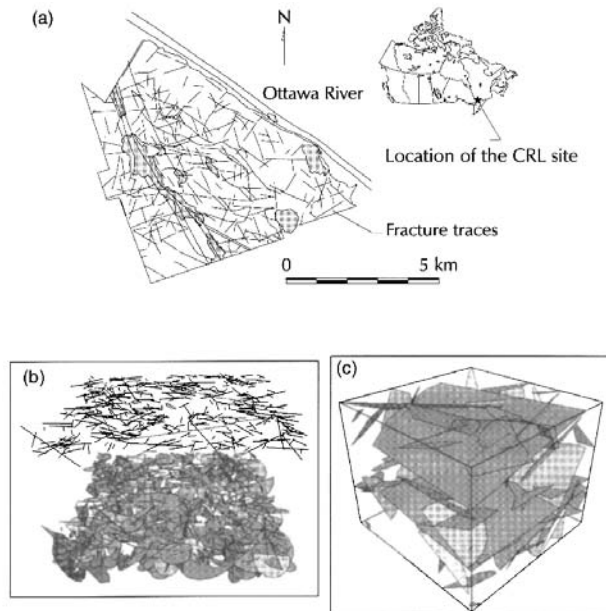
FIGURE 17.27 Generation of a conceptual model of fracture network using field data and fractal relations: (a) fracture traces measured from outcrop scan lines, air photographs, and geophysical surveys, (b) statistically generated two-dimensional traces (top) extended to form three-dimensional network of fractures (bottom), and (c) detailed portion of network shown in bottom part of (b). (From Piggott, A., Moltyaner, G., Yamazaki, L., and Novakowski, K. 1997. Preliminary characterization of fracturing at the site of the Chalk River Laboratories, in *Proceedings of 50th Canadian Geotechnical Conference*, Ottawa, Canada (also NWRI No. 97-126). With permission.)
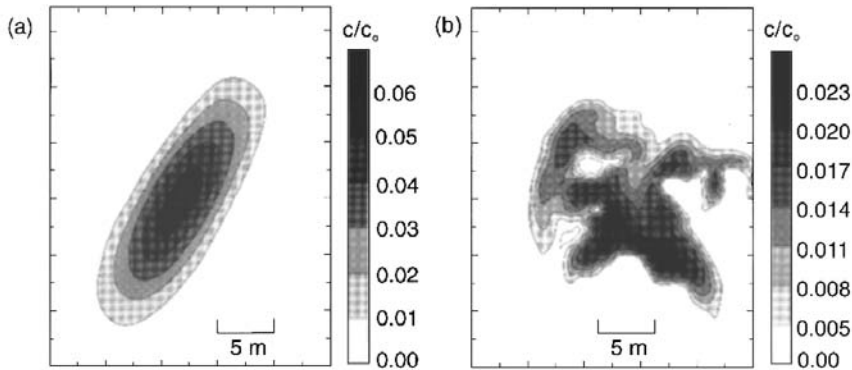
FIGURE 17.28 Two-dimensional concentration plumes in a single fracture illustrating the effect of variable aperture on plume dimensions: (a) tracer plume assuming a constant aperture, and (b) tracer plume assuming aperture of the fracture is variable.

shape and of lower peak concentration. Thus for some distributions of aperture it may be necessary to consider variable aperture in simulation at the discrete fracture scale. Unfortunately, in almost all field studies there is insufficient information to generate a defensible model of variable aperture.

Modeling approaches to simulate flow and transport in fracture networks fall into one of three categories within the range of conceptual models for fractured rock: (1) equivalent porous medium (EPM), (2) dual porosity, (3) discrete fracture representation. In addition, fracture network models may be implemented in either two or three dimensions.

Models based on equivalent porous medium treat the fractured porous rock as equivalent to a non-fractured continuum. Bulk parameters for the permeability of the rock mass are used, and the geometry of individual fractures or the rock matrix is not considered. This is a reasonable approach if fracturing is intense or the study domain is sufficiently large such that individual fractures have no influence on the overall flow system (e.g., some regional systems of the scale of kilometers). Berkowitz et al. (1988) and Schwartz and Smith (1988) discuss the use of continuum models for fractured rock systems.

Raven (1986) applied a continuum model using a detailed set of hydraulic testing results obtained from a small flow system in a monzonitic gneiss. Constant-head injection tests were conducted using contiguous test intervals in 17 boreholes of approximately 50-m depth. Using BHTV, fracture orientation and frequency was determined for each test interval. By assuming each fracture intersecting a given test interval was of equal aperture, the CHIT result was related to an effective permeability tensor for the interval (Snow's method described in the previous section). The tensor was diagonalized to determine the principal directions and principal values of the hydraulic conductivity ellipsoid. These values were then used in a two-dimensional finite-element model based on a single continuum to evaluate the degree to which the conceptual model (i.e., equivalent porous media) could predict the flow of groundwater through the rock mass. Results indicated that the conceptual model and numerical approach could adequately describe flow conditions in the shallow subsurface (<30 m), but could only poorly describe these at greater depths. This was attributed to the inadequacy of the conceptual model in incorporating vertical flow conditions which were more prevalent at depth in this site. It should also be noted that the conceptual model was not tested against the transport properties of the fracture system.

Double porosity models are used to attempt to bridge the gap between the simplifications of EPM models and the details of the discrete fracture models by treating the fracture system and the porous matrix as two separate inter-related continua (Barenblatt et al., 1960). In this approach, equations of flow and transport for each system are linked by a source/sink term that describes the fluid or solute exchange between the two systems each of which may have very different properties relative to the other. Examples of semianalytical and numerical models having this approach can be found in Huyakorn et al. (1983), Rowe and Booker (1990), and Sudicky (1990). Some limitations to the double porosity approach
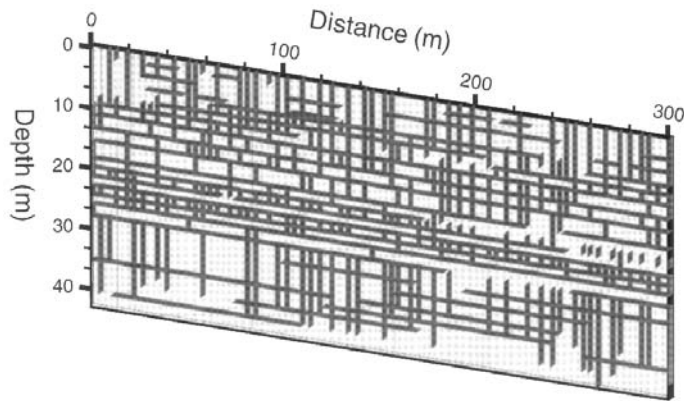
FIGURE 17.29 Network of orthogonal fractures generated based on a conceptual model of layered sedimentary rock where each layer has differing fracture spacing and connectivity parameters. System is divided into three layers: a densely fractured zone (0 to 26 m), a relatively unfractured zone (26 to 28 m), and a more sparsely fractured zone (28 to 45 m).

include the assumption that the matrix blocks are of simple (and possibly unrealistic) geometry and that advection of solutes in each block is ignored.

The discrete fracture approach will result in the most physically representative simulation of flow and transport processes at the subcontinuum scale. However, discrete fracture models require the generation of fracture networks based on a working conceptual model developed using information on both the individual fractures and the geometry of interfracture relationships. A network of fractures is characterized by a distribution of fractures of fixed or variable aperture, finite length, regular or random orientations, and some degree of connectivity with other individual fractures. Thus, the development of the conceptual model strongly influences the outcome of the model simulations. Figure 17.29 illustrates a fracture network generated using a conceptual model of fracturing in a layered sedimentary sequence of dolostone and shale. To generate this network minimum horizontal and vertical fracture spacings of 1.5 m and 3.5 m were set for the entire sequence. Within each geologic unit a range of fracture lengths and fracture density were set. Note the denser fracture zones at the top of the sequence (0 to 26 m), a relatively unfractured zone at 26 m, and a more sparsely fractured zone from 28 m to 45 m.

17.6.1 Flow in Discrete Fracture Networks

The modeling of flow in discrete fracture networks may be conducted to estimate groundwater flux for the purpose of groundwater resource evaluation or to provide an estimate of the distribution of groundwater velocity in individual fractures as input to a solute transport model. In the former case, transient hydraulic head conditions may be of interest, whereas in the latter, only steady conditions are considered. In the following, two approaches to the modeling of flow in discrete features are described. Both are suitable for either transient or steady flow conditions.

Barker (1991) developed a robust semianalytical method which can be used to determine the distribution of groundwater flux and velocity in a two-dimensional network of fractures having random orientations. The fracture network is conceptualized as a linked system of linear elements for both steady and transient flow conditions (e.g., Dverstorp et al., 1992). Flux into each node is calculated by summing the contribution from each fracture using Darcy's law. The cubic law is used to relate the fracture aperture to elemental hydraulic conductivities. The system of Laplace transformed equations is solved by direct or iterative methods depending on the size of the network and the results numerically inverted from the Laplace domain to obtain values in real space.

Modeling groundwater flow in three-dimensional fracture networks is considerably more difficult than for two-dimensional slices. For conceptual models involving sparsely distributed fractures in impermeable

rock, randomly oriented fracture discs and fracture intersections of various orientations must be discretized. However, the mesh generation required for solution with standard finite element or finite difference techniques is prohibitively difficult, even for small solution domains. Alternative solution methods have been attempted including a hybrid analytical–numerical scheme (Long et al., 1985) and a boundary element formulation (Elsworth, 1986), but these were found not to be generally useful for large complex networks.

Therrien and Sudicky (1996) derived a model for a variably saturated fracture-matrix system having three dimensions in which advective fluid exchange is allowed between the fractures and matrix. A modified form of the Richards' equation was used for determining hydraulic head in the matrix, and an extension to the variably saturated flow equations was used for hydraulic head in the fracture. The fractures were idealized as two-dimensional parallel plates, and fluid leakage flux was used to link the equations for the fracture with that for the matrix. The resulting system of governing equations was solved using the control-volume finite-element method in conjunction with Newton–Raphson linearization. Although there is no technical limitation to the method in the simulation of flow in a nonorthogonal fracture system, as mentioned earlier, any nonorthogonal problem simulated using this model would be significantly limited in scale.

17.6.2 Solute Transport in a Discrete Fracture Network

Solute transport in fracture networks can be simulated in two ways: (1) using particle tracking, and (2) direct solution of the governing equations for solute transport. Of the two methods, the former has received more widespread use. In the following, both particle tracking and methods for direct solution will be described, with focus on the former.

In modeling solute transport in fracture networks, the presence of fracture intersections, where solute must be apportioned according to the flux entering and leaving the intersection node, must be incorporated. There are two approaches to this: (1) assume complete mixing at the node, and (2) use streamtube routing. For the former, concentration entering the intersection is assumed to completely mix, resulting in a uniform distribution of concentration leaving the intersection. For the latter, mass is distributed at the intersection according to the distribution of streamtubes, and no mixing is assumed to occur (Figure 17.30). Complete mixing at an intersection having incoming fractures is described by (Küpper et al., 1995a):

$$c = \frac{\sum_{i=1}^{n} c_i Q_i}{\sum_{i=1}^{n} Q_i} \tag{26}$$

where c is the concentration in the intersection, and Q_i and c_i are the volumetric flux and concentrations in the incoming fractures, respectively. The development required for a rigorous implementation of streamtube routing is considerably more complex (e.g., Philip, 1988).

Berkowitz et al. (1994) suggested that diffusional transfer between streamtubes at intersections may smear the distinction between streamtubes, resulting in a distribution of concentration more like that observed with complete mixing. However, Küpper et al. (1995b) showed that, even in the absence of diffusion, for three of four possible flow conditions at intersections the complete mixing assumption and streamtube routing produce identical results. In addition, it was also shown that for some types of flow (i.e., radial flow in a network), the flow condition giving rise to the difference is rare. Thus, in practical terms, the degree of error introduced into a model by assuming complete mixing at every intersection is likely not significant. However, it is important to note that these concepts apply only to mixing at linked one-dimensional elements. In the more realistic case where two-dimensional planes intersect, the
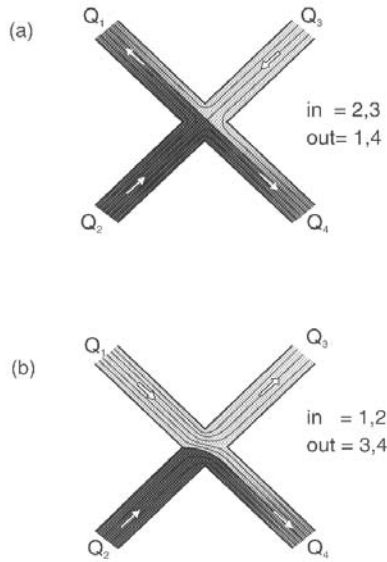
FIGURE 17.30 Schematic diagram showing mass transfer at a fracture intersection according to the distribution of streamtubes. (a) $Q_2 = Q_3$ and transport is split evenly into Q_1 and Q_4 and (b) $Q_1 > Q_2$ and transport is split unevenly between Q_3 and Q_4.

hydraulic head will be nonconstant along the intersection, resulting in much more complex mixing arrangements.

Particle tracking is a means to simulate solute transport by following the residence times of particles released into the flow system at a given boundary or internal location. In the most rudimentary implementation, a given number of particles, usually 20,000 to 30,000, are released into the flow system, routed at fracture intersections according to volumetric flux (i.e., assume perfect mixing at the intersection node) and eventually tracked to an exit boundary (Schwartz et al., 1983). The number of particles are summed at the exit boundary resulting in a breakthrough curve equivalent to a dirac input. Integration of the breakthrough curve with respect to time yields a breakthrough curve equivalent to a step-function input.

The use of particle tracking methods has been expanded in recent years to account for solute retardation (Dverstorp et al., 1992) and variable aperture fracture elements (Nordqvist et al., 1992). Retardation is accounted for by manipulating the travel time in individual fracture elements using the expression:

$$t_R = t_C R_a \tag{27}$$

where t_R is the travel time adjusted for sorption, t_C is determined from the cubic law, and R_a is as defined in Equation (12). Nordqvist et al. (1992) also manipulated residence time in fracture elements to account for the influence of variable aperture fractures. This was conducted by generating a spectrum of residence times using a variety of realizations of a variable aperture fracture having different mean apertures and different properties of the lognormal aperture distribution. To develop the residence time distribution for the system of linked linear elements, the residence time spectrum was randomly sampled for each fracture element (Figure 17.31). Matrix diffusion could also be accounted for in a similar fashion although the implementation would not be as rigorous as that achieved using the methods described in the following.

For two-dimensional networks of linked linear fracture elements, direct solution offers the most robust means of simulating transport in the fractures and interactions with the matrix. Sudicky and McLaren (1992) and Bogan (1996) describe finite-element and semianalytical approaches, respectively.
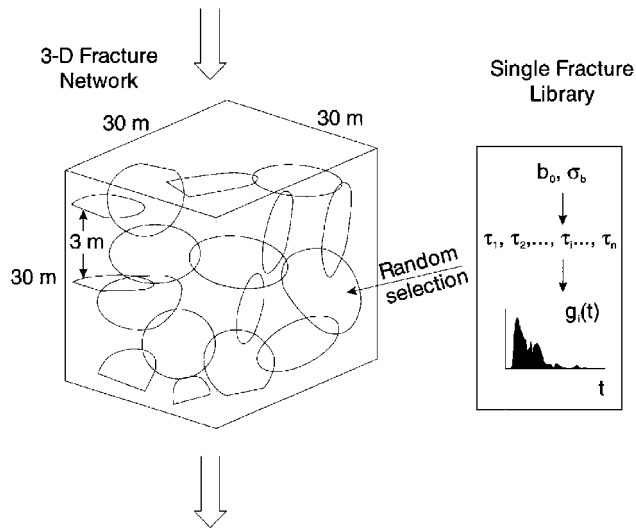
FIGURE 17.31 The incorporation of variable aperture fractures is introduced in a fracture network model. Here, b_o and σ_b are the two parameters describing the lognormal aperture distribution (mean aperture and standard deviation). The τ_i are effective transmissivities of the fractures in the library. The $g_i(t)$ plots are graphical illustrations of the residence time spectra for the fractures in the library. (From Nordqvist, A. W., Tsang, Y., Tsang, C.-F., Dverstorp, B., and Andersson, J. 1992. A variable aperture fracture network model for flow and transport in fractured rocks. *Water Resour. Res.* 28(6):1703-1713. With permission.)

The solution method described by Sudicky and McLaren (1992) accommodates advection-dispersion, retardation, and decay in the fracture and the same in the matrix. A third-type boundary condition is used for the exchange between fracture and matrix. This solution method can therefore simulate fractured domains having a permeable matrix (e.g., fractured clays). The finite element method is used to discretize the equations in the spatial coordinates, and the Laplace transform is used to manage the time derivative. To minimize the difficulties in discretization and to maximize the domain size for tractable problems, fracture geometry has been limited to orthogonal systems.

For two-dimensional networks having more random fracture orientations, the solution method of Bogan (1996) offers a practical alternative. In this case, only diffusive transport is allowed between the fracture and matrix. A semi-infinite domain is also defined for the diffusive transport in the matrix. Clearly, this is an approximation that is viable only for geological material having sparse fractures and low matrix porosity.

In three dimensions, the direct solution method of Therrien and Sudicky (1996) for a variably saturated fractured medium can be employed. The two governing equations for the fracture planes and matrix blocks are linked using continuity between the concentration in the fracture and that in the matrix at the fracture wall. Because the transport equations are linear, solution was based on a standard time-marching Galerkin scheme. Two-dimensional rectangular elements were used for the fractures, and three-dimensional rectangular prisms were used for the matrix.

Presently in the general practice of hydrogeology, the modeling of flow and transport in fractured rock systems is often conducted using EPM models. As we have suggested above, this approach will lead to error in the prediction of both groundwater flux or solute travel time, depending on the size of the domain modeled. At a domain scale of several hundreds of meters, or more, however, prediction of groundwater flux may be conducted with only minor error using EPM models, provided the characteristics of the major fracture zones and unfractured rock are known and appropriately represented. Thus, the hydraulic effects of short or long-term pumping or changes in boundary conditions such as recharge can be reliably explored using this approach.

Unfortunately, the prediction of solute transport using the EPM approach at this scale can be substantially erroneous. Thus, as a general rule of thumb, all predictions of contaminant transport and the design of remedial facilities at the scale of a contaminated site should be conducted using a model which, at minimum, incorporates multicontinuum concepts. In some cases, a defensible modeling strategy may be to simulate transport in the fracture system with simple analytical models which incorporate a parallel system of fractures and matrix diffusion. Although the fracture system may not be well represented by this approach, it is likely that this approach will lead to substantially more accurate predictions of contaminant transport than were employed in an EPM model.

For conditions where more accurate simulations of contaminant transport are required, such as for large-scale groundwater remediation programs or investigations of intrinsic bioremediation, there are now very robust modeling tools available which account for fractures on a discrete basis. Unfortunately, because few of these are widely available in commercial packages, attempts to conduct this type of modeling should be left to specialists.

Acknowledgment

For Further Information

Engelder et al. (1993) presents a thorough overview of the geological aspects of fracture mechanics and includes a comprehensive glossary of terms. de Marsily (1986) is a general text in quantitative hydrogeology which discusses flow and transport in fractures in an easy-to-follow manner. Kulander et al. (1990) discusses in detail core analysis and the identification and classification of breaks in rock core. Doe et al. (1987) discusses various techniques and strategies for hydraulic testing in fractured rock. Earlougher (1977) provides a comprehensive look at standard well-testing methods in the petroleum industry which can be adapted to hydrogeological investigations. Bear et al. (1993) provides a compilation of chapters investigating flow and contaminant transport in fractured rock, and in addition to theoretical discussions has sections on flow and solute transport in fracture networks.

References

Abelin, H., Birgersson, L., Gidlund, J., and Neretnicks, I. 1991. A large-scale flow and tracer experiment in granite 1. Experimental design and flow distribution. *Water Resour. Res.* 27(12):3107-3117.

Barenblatt, G. I., Zheltov, I. P., and Kochina, I. N. 1960. Basic concepts in the theory of seepage if homogeneous liquids in fissured rocks. *J. Appl. Math. Mech. Engl. Transl.* 24(5):852-864.

Barker, J. A. 1991. Reciprocity principle and analytical solution for Darcian flow in a network. *Water Resour. Res.* 27(5):743-746.

Bear, J. 1993. Modeling flow and contaminant transport in fractured rocks, in *Flow and Contaminant Transport in Fractured Rocks*. Eds. J. Bear, C.-F. Tsang, and G. de Marsily. Academic Press, San Diego, 1-37.

Bear, J. 1972. *Dynamics of Fluids in Porous Media*. Elsevier, New York.

Bear, J., Tsang, C.-F., and de Marsily, G. 1993. *Flow and Contaminant Transport in Fractured Rocks*. Academic Press, San Diego.

Berkowitz, B., Naumann, C., and Smith, L. 1994. Mass transfer at fracture intersections: an evaluation of mixing models. *Water Resour. Res.* 30(6):1765-1773.

Berkowitz, B., Bear, J., and Braester, C. 1988. Continuum models for contaminant transport in fractured porous formations. *Water Resour. Res.* 24(8):1225-1236.

Black, W. H., Smith, H. R., and Patton, F. D. 1987. Multiple-level groundwater monitoring with the MP system, in Proceedings NWWA-AGU Conf. Surface and Borehole Geophysical Methods and Groundwater Instrumentation, NWWA, Dublin, Ohio.

Bliss, J. C. and Rushton, K. R. 1984. The reliability of packer tests for estimating the hydraulic conductivity of aquifers. *Q. J. Engrg. Geol.*, 17:88-91.

Bogan, J. D. 1996. *A Semi-Analytical Model for the Simulation of Solute Transport in a Network of Fractures.* M.Sc. Thesis, University of New Brunswick, Fredericton NB.

Bredehoeft, J. D. and Papadopulos, S. S. 1980. A method for determining the hydraulic properties of tight formations. *Water Resour. Res.* 16(1):233-238.

Brown, S. R. 1995. Simple mathematical model of a rough fracture. *J. of Geophys. Res.* 100(B4):5941-5952.

Brown, S. R., Kranz, R. L., and Bonner, B. P. 1986. Correlation between the surfaces of natural rock joints. *Geophys. Res. Letters.* 13(13):1430-1433.

Cacas, M. C., Ledoux, E., de Marsily, G., Tillie, B., Barbareau, A., Calmels, P., Gaillard, B., and Margritta, R. 1990. Modelling fracture flow with a stochastic discrete fracture network: calibration and validation 1. The flow model. *Water Resour. Res.* 26(3):479-489.

Cherry, J. A. and Johnson, P. E. 1982. A multilevel device for hydraulic head monitoring and groundwater sampling in fractured rock. *Ground Water Monitoring Rev.* 2(3):41-44.

Cooper, H. H., Bredehoeft, J. D., and Papadopulos, I. S. 1967. Response to a finite diameter well to an instantaneous charge of water. *Water Resour. Res.* 3(1):263-269.

de Marsily, G. 1986. *Quantitative Hydrogeology.* Academic Press, Orlando, FL.

Doe, T. and Remer, J. 1980. Analysis of constant-head well tests in non-porous fractured rock, in *Third Invitational Well-Testing Symposium — Well Testing in Low Permeability Environments.* Berkeley, CA, 84-89.

Doe, T., Osnes, J., Kenrick, M., Geier, J., and Warner, S. 1987. Design of well testing programs for waste disposal in crystalline rock, in *Proceedings of 6th Congress of the International Society for Rock Mechanics.* Montreal, Canada.

Dougherty, D. E. and Babu, D. K. 1984. Flow to a partially penetrating well in a double porosity reservoir. *Water Resour. Res.* 20(8):1116-1122.

Drost, W., Klotz, D., Koch, A., Moser, H., Neumaier, F., and Werner, R. 1968. Point dilution methods of investigating ground water flow by means of radioisotopes, *Water Resourc. Res.* 4(1):125-146.

Dunnicliff, J. 1988. *Geotechnical Instrumentation for Monitoring Field Performance.* John Wiley & Sons, New York.

Dverstorp, B., Andersson, J., and Nordqvist, W. 1992. Discrete fracture network interpretation of field tracer migration in sparsely fractured rock. *Water Resour. Res.* 28(9):2327-2343.

Earlougher, R. C., Jr. 1977. *Advances in Well Test Analysis, Monogr. 5.* Society of Petroleum Engineers, Dallas, TX.

Elsworth, D. 1986. A model to evaluate the transient hydraulic response of three-dimensional sparsely-fractured rock masses. *Water Resour. Res.* 22(13):1809-1819.

Engelder, T., Fischer, M. P., and Gross, M. R. 1993. *Geological Aspects of Fracture Mechanics, A Short Course Manual.* Geological Society of America. Boulder, CO.

Faust, C. R. and Mercer, J. W. 1984. Evaluation of slug tests in wells containing a finite-thickness skin. *Water Resour. Res.* 20(4):504-506.

Gentier, S. and Billaux, D. 1989. Caracterisation en laboratoire de l'espace fissural d'une fracture, in *Proceedings of the International Symposium on Rock at Great Depth.* 1:425-431, A. A. Balkema, Rotterdam, The Netherlands.

Grisak, G. E., Merritt, W. F., and Williams, D. W. 1977. A fluoride borehole dilution apparatus for groundwater velocity measurements. *Can. Geotech. J.* 14:554-561.

Hakami, E. 1995. *Aperture Distribution of Rock Fractures*, Ph.D. Thesis, Department of Civil and Environmental Engineering, Royal Institute of Technology, Stockholm.

Hess, A. E. and Paillet, F. L. 1990. Applications of the thermal-pulse flowmeter in the hydraulic characterization of fractured rocks, in ASTM STP 1101:99-112.

Holst, T. B. 1982. Regional jointing in the northern Michigan basin. *Geology.* 10:273-277.

Holzhauser, G. R. 1989. Origins of sheet structure, 1. Morphology and boundary conditions. *Engineer. Geol.* 27:225-279.

Huyakorn, P. S., Lester, B. H., and Mercer, J. W. 1983. An efficient finite element technique for modeling transport in fractured porous media, 1. Single species transport. *Water Resour. Res.* 19(3):841-854.

Johns, R. A. and Roberts, P. V. 1991. A solute transport model for channelized flow in a fracture. *Water Resour. Res.* 27(8):1797-1808.

Jouanna, P. 1993. A summary of field test methods in fractured rocks, in *Flow and Contaminant Transport in Fractured Rocks.* Eds. J. Bear, C.-F. Tsang, and G. de Marsily. Academic Press, San Diego, 437-543.

Karasaki, K., Long, J. C. S., and Witherspoon, P. A. 1988. Analytical models of slug tests. *Water Resour. Res.* 24(1):115-126.

Kulander, B. R., Dean, S. L., and Ward, B. J. 1990. *Fractured Core Analysis: Interpretation, Logging and Use of Natural and Induced Fractures in Core: AAPG Methods in Exploration Series.* No. 8. American Association of Petroleum Geologists, Tulsa, OK.

Küpper, J. A., Schwartz, F. W., and Steffler, P. M. 1995a. A comparison of fracture mixing models, 1, A transfer function approach to mass transport modeling. *J. Cont. Hydrol.* 18:1-32.

Küpper, J. A., Schwartz, F. W., and Steffler, P. M. 1995b. A comparison of fracture mixing models, 2, Analysis of simulation trials. *J. Cont. Hydrol.* 18:33-58.

La Pointe, P. R. and Hudson, J. A. 1985. Characterization and interpretation of rock mass joint patterns. *Geological Society of America Special Paper 199.*

Lapcevic, P. A. 1997. *Tracer Experiments Conducted in a Discrete Horizontal Fracture Under Conditions of Forced Hydraulic Gradient and Natural Groundwater Flow,* M.Sc. Thesis, University of Waterloo, Waterloo, Ontario.

Lapcevic, P. A., Novakowski, K. S., and Cherry, J. A. 1990. The characterization of two discrete horizontal fractures in shale, in *Proc. Technology Transfer Conference, Ontario Ministry of Environment.* November, Toronto. Vol II, 486-495.

Lau, J. S. O. 1983. The determination of true orientations of fractures in rock cores. *Canadian Geotech. J.* 20:221-227.

Long, J. C. S., Gilmour, P., and Witherspoon, P. A. 1985. A model for steady fluid flow in random three-dimensional networks of disc-shaped fractures. *Water Resour. Res.* 21(8):1105-1115.

Moench, A. F. and Hsieh, P. A. 1985. Analysis of slug test data in a well with finite thickness skin, in *Memoirs of the 17th International Congress on the Hydrogeology of Rocks of Low Permeability,* 17:17-29, International Association of Hydrologists, Tucson, AZ.

Moreno, L., Tsang, Y. W., Tsang, C. F., Hale, F. V., and Neretnieks, I. 1988. Flow and tracer transport in a single fracture: a stochastic model and its relation to some field observations. *Water Resour. Res.* 24(12):2033-2048.

Neuzil, C. E. 1982. On conducting the modified 'slug' test in tight formations. *Water Resour. Res.* 18(2):439-441.

Nordqvist, A. W., Tsang, Y., Tsang, C.-F., Dverstorp, B., and Andersson, J. 1992. A variable aperture fracture network model for flow and transport in fractured rocks. *Water Resour. Res.* 28(6):1703-1713.

Novakowski, K. S. 1989. Analysis of pulse interference tests. *Water Resour. Res.* 25(11):2377-2387.

Novakowski, K. S. and Lapcevic P. A. 1994. Field measurement of radial solute transport in fractured rock. *Water Resour. Res.* 30(1):37-44.

Novakowski, K. S., Evans, G. V., Lever, D. A., and Raven, K. G. 1985. A field example of measuring hydrodynamic dispersion in a single fracture. *Water Resour. Res.* 21(8):1165-1174.

Odling, N. E. 1994. Natural rock profiles, fractal dimensions and joint roughness coefficients. *Rock Mech. Rock Engng.* 27(3):135-153.

Paillet, F. L. and Kapucu, K. 1989. Characterization of fracture permeability and fracture flow modeling at Mirror Lake, New Hampshire. *U.S. Geological Survey Water Resources Investigations Report* 89-4058. U.S. Government Printing Office. Denver, CO.

Philip, J. R. 1988. The fluid mechanics of fractures and other junctions. *Water Resour. Res.* 24(2):239-246.

Piggott, A. R. 1990. *Analytical and Experimental Studies of Rock Fracture Hydraulics.* Ph.D. thesis, Pennsylvania State University, University Park, PA.

Piggott, A., Moltyaner, G., Yamazaki, L., and Novakowski, K. 1997. Preliminary characterization of fracturing at the site of the Chalk River Laboratories, in *Proceedings of 50th Canadian Geotechnical Conference*, Ottawa, Canada (also NWRI No. 97-126).

Piggott, A. R., Huynh, T. N. T., Lapcevic, P. A., and Novakowski, K. S. 1996. Automated analysis of hydraulic and tracer tests conducted in fractured rock. *Hydrogeology J.* 4(3):84-93.

Piggott, A. R. and Elsworth, D. 1993. Laboratory assessment of the equivalent apertures of a rock fracture. *Geophys. Res. Letters.* 20(13):1387-1390.

Pyrak-Nolte, L. J., Myer, L. M., Cook, N. G. W., and Witherspoon, P. A. 1987. Hydraulic and mechanical properties of natural fractures in low permeability rock, in *Proceedings of 6th Congress of the International Society for Rock Mechanics*, Montreal, Canada, 225-231.

Ragan, D. M. 1973. *Structural Geology: An Introduction to Geometrical Techniques.* John Wiley & Sons, New York.

Raven, K. G. 1986. Hydraulic characterization of a small ground-water flow system in fractured monzononitic gneiss. *Nat. Hyd. Res. Inst. Scientific Series No. 149*, No. 30. Government of Canada, Ottawa.

Raven, K. G., Novakowski, K. S., and Lapcevic, P. A. 1988. Interpretation of field tracer tests of a single fracture using a transient solute storage model. *Water Resour. Res.* 24(12):2019-2032.

Rowe, R. K. and Booker, J. R. 1990. Contaminant migration through fractured till into an underlying aquifer. *Can. Geotech. J.* 27(4):484-495.

Sageev, A. 1986. Slug test analysis. *Water Resour. Res.* 22(8):1323-1333.

Sanford, B. V., Thompson, F. J., and McFall, G. H. 1985. Phanerozoic and recent tectonics of the Canadian Craton, in *The Geoscience Program — Proceedings of the Seventeenth Information Meeting of the Nuclear Fuel Waste Management Program*, Atomic Energy of Canada Limited, TR-299:334-352. Winnipeg.

Schwartz, F. W. and Smith, L. 1988. A continuum approach for modeling mass transport in fractured media. *Water Resour. Res.* 24(8):1360-1372.

Schwartz, F. W., Smith, L., and Crowe, A. 1983. A stochastic analysis of macroscopic dispersion in fractured media. *Water Resour. Res.* 19(5):1253-1265.

Shapiro, A. M. and Nicholas, J. R. 1989. Assessing the validity of the channel model of fracture aperture under field conditions. *Water Resour. Res.* 25(5):817-828.

Smith, L. and Schwartz, F. W. 1984. An analysis of the influence of fracture geometry on mass transport in fractured media. *Water Resour. Res.* 20(9):1241-1252.

Snow, D. T. 1969. Anisotropic permeability of fractured media. *Water Resour. Res.* 5(6):1273-1289.

Sudicky, E. A. 1990. The Laplace transform galerkin technique for efficient time-continuous solution of solute transport in double-porosity media. *Geoderma.* 46:209-232.

Sudicky, E. A. and McLaren, R. G. 1992. The Laplace transform Galerkin technique for large-scale simulation of mass transport in discretely fractured porous formations. *Water Resour. Res.* 28(2):499-514.

Sudicky, E. A. and Frind, E. O. 1982. Contaminant transport in fractured porous media: analytical solution for a system of parallel fractures. *Water Resour. Res.* 18(6):1634-1642.

Tang, D. H., Frind, E. O., and Sudicky, E. A. 1981. Contaminant transport in fractured porous media: analytical solution for a single fracture. *Water Resour. Res.* 17(3):555-564.

Therrien, R. and Sudicky, E. A. 1996. Three-dimensional analysis of variably-saturated flow and solute transport in discretely-fractured porous media. *J. of Cont. Hydrol.* 23(1-2):1-44.

Trainer, F. W. 1988. Hydrogeology of the plutonic and metamorphic rocks, in *Hydrogeology.* Eds. Back, W., Rosenshein, J. S., and Seaber, P. R. Geological Society of America, The Geology of North America, Boulder, CO, v. O-2:367-380.

Tsang, Y. W. 1984. The effect of tortuosity on fluid flow through a single fracture. *Water Resour. Res.* 20(9):1209-1215.

Tsang, Y. W. and Tsang, C. F. 1987. Channel model of flow through fractured media. *Water Resour. Res.* 23(3):467-479.

Tsang, Y. W. and Witherspoon, P. A. 1983. The dependence of fracture mechanical and fluid flow properties on fracture roughness and sample size. *J. of Geophys. Res.* 88(B3):2359-2366.

Vickers, B. C., Neuman, S. P., Sully, M. J., and Evans, D. D. 1992. Reconstruction and geostatistical analysis of multiscale fracture apertures in a large block of welded tuff. *Geophys. Res. Letters.* 19(10):1029-1032.

Wang, J. S. Y., Narasimhan, T. N., and Scholz, C. H. 1988. Aperture correlation of a fractal fracture. *J. Geophys. Res.* 93(B3):2216-2224.

Williams, H. R., Corkery, D., and Lorek, E. G. 1985. A study of joints and stress-relief buckles in Paleozoic rocks of the Niagara Peninsula, southern Ontario. *Can. Geotech. J.* 22:296-300.

Witherspoon, P. A., Wang, J. S. Y., Iwai, K., and Gale, J. E. 1980. Validity of cubic law for fluid flow in a deformable rock fracture. *Water Resour. Res.* 16(6):1016-1024.

Ziegler, T. W. 1976. *Determination of Rock Mass Permeability.* Waterways Experiment Station, Technical Report S-76-2, Vicksburg, Mississippi.

Zemanek, J., Caldwell, R. L., Glenn, E. E., Holcomb, S. V., Norton, L. J., and Strauss, A. J. D. 1969. The borehole televiewer — a new logging concept for fracture location and casing inspection. *J. Pet. Technol.* 21(12):762-774.

Glossary

Aperture   Separation distance between two fracture surfaces; used as measure of fracture width.

Borehole   Drilled open hole in rock. Boreholes may be of any orientation and plunge from vertical to horizontal in relation to the earth's surface.

Channel   Preferred pathway in a single fracture plane or fracture network.

Conceptual Model   Geological and/or hydrogeological image representing natural system.

Fracture   Any planar or curviplanar discontinuity or break in a rock mass that has formed as a result of a brittle deformation process. Joints, shear fractures, faults, microcracks, etc., are all examples of fractures.

Fracture Network   Two- or three-dimensional arrangement of intersecting fractures. Fractures in a network may be of identical aperture, length, or orientation or may be represent a distribution of fractures of varying parameters.

Hydraulic Test   Test involving the injection or withdrawal of water into a given test zone to determine the permeability of the zone. Either the transient or steady-state response can be used.

Interference Test   Hydraulic test conducted between two or more hydraulically connected zones. A vertical interference test can be conducted between two isolated zones in a single borehole.

Joint   A mesoscopic fracture in rock exhibiting purely opening mode displacement, with no appreciable shear offset.

Matrix Diffusion   Process whereby solute from water in fracture is transferred to water in the pore spaces of rock through chemical diffusion.

Rock Matrix   Structure of unfractured bulk rock mass composed of mineral grains, cement, pore water, and pore space.

Sheeting Structure   Mesoscopic fractures formed due to the vertical expansion of the rock resulting from erosional unloading. Sheeting structures are often found parallel to the earth's surface in all types of rock and are common conduits for groundwater flow at the regional scale.

Skin   Zone around borehole which has different properties than the bulk rock mass. A positive skin suggests the permeability around the borehole is less than the bulk rock, whereas a negative skin suggests greater permeability. Skin zones may be caused by fracture infilling due to rock flour, drill bit damage, or invasion of drilling muds.

Surface Roughness   Texture of fracture surface due to asperities in the plane. Roughness can range from smooth planes to highly irregular surfaces.

Wellbore Storage   Volume of water in the borehole which, when pumping is initiated, is first depleted prior to water stored in fracture or rock matrix.

18
Groundwater Flow in Karstic Aquifers

William B. White
The Pennsylvania State University

18.1 Introduction

Karstic aquifers are those that contain solution-generated cavities that permit rapid transport of groundwater, often in turbulent flow, and often carrying a sediment load. Such aquifers are found in gypsum, limestone, and dolomite rocks. In most aquifers there is a mismatch of many orders of magnitude between groundwater flow rates within the aquifer and stream flow rates on the land surface above. In karstic aquifers, there is a continuum of flow rates such that the groundwater system in karstic aquifers takes on the characteristics of both surface water and groundwater. There is also a rapid interchange of surface water and groundwater.

Karst regions display a characteristic suite of landforms consisting primarily of deranged surface drainage with sinking streams connected by underground flow routes to large springs, closed depressions on many size scales, sculptured bedrock surfaces, and caves. Karst regions are often classified by their dominant landforms such as fluviokarst for mixed carbonate–clastic rock regions, cone-and-tower karst, pavement karst, doline karst, and others. They are also sometimes classified by climate as in alpine karst, tropical karst, arid karst, and others. The emphasis of this chapter is on fluviokarst. The study of karst has traditionally been the province of physical geographers and geomorphologists, and most books approach the subject from this point of view. Books that emphasize the hydrologic and hydrogeologic aspects of karst are Milanovic, White, Dreybrodt, and Ford and Williams.[1-4]

18.2 Karst Aquifers: The Conceptual Framework

18.2.1 Groundwater Basins

The physical framework for discussion of groundwater is the aquifer. Usually, an aquifer is a specific rock formation with a well-defined thickness but an indefinite lateral extent. One can speak of the "area" of an aquifer only in a rough sense. The physical framework for discussion of surface water is the drainage basin. Drainage basins have an arbitrary downstream limit determined by the choice of gauge point but once the gauge point is fixed, a drainage divide can be drawn and the basin has a well-defined area. Karst aquifers share the properties of both groundwater and surface water which leads to the concept of a groundwater basin with well-defined boundaries which do not necessarily coincide with the surface basin boundaries.

The essential components are sketched in Figure 18.1. The sketch is appropriate to the mixed carbonate/clastic rock terrain that makes up the fluviokarst of the eastern U.S. There are three main components to the catchment system: allogenic recharge, internal runoff, and diffuse infiltration. The groundwater basin arises when the aquifer acquires a well-developed conduit system. The conduit system provides a flow path of low hydraulic resistance. It also acts as a drain for matrix flow within the fracture and pore system so that there is the subsurface equivalent of a drainage divide. However, unlike surface divides which are fixed by topography, groundwater basin divides may shift depending on groundwater levels. High-level spillovers between adjacent basins are common.

18.2.2 Allogenic Recharge

Streams rising on adjacent or overlying clastic rocks often sink underground near the clastic rock/carbonate rock contact when they flow onto the karst. These sink points are called swallow holes or swallets. Each sinking stream forms its own small surface catchment with a well-defined area. The sink point acts as the gauge point. The sinking streams inject large quantities of water into the carbonate aquifer at single points. This is the component of the groundwater referred to as allogenic recharge. Allogenic recharge often has a different chemistry, is exceptionally vulnerable to pollution, and exhibits the rapid response to storm flow characteristic of surface runoff. The total area providing allogenic recharge is simply the sum of the areas of the individual sinking stream basins:

$$A_a = \sum_i A_i \tag{1}$$

Many relationships between surface water and karst groundwater are possible. In some cases the allogenic streams have downcut upstream from the swallet forming blind valleys so that all allogenic runoff feeds into the groundwater regardless of the magnitude of flow. In other cases, a dry surface channel continues downstream from the swallet so that during flood flows, only a portion of the flow enters the aquifer and the remainder remains as surface water.
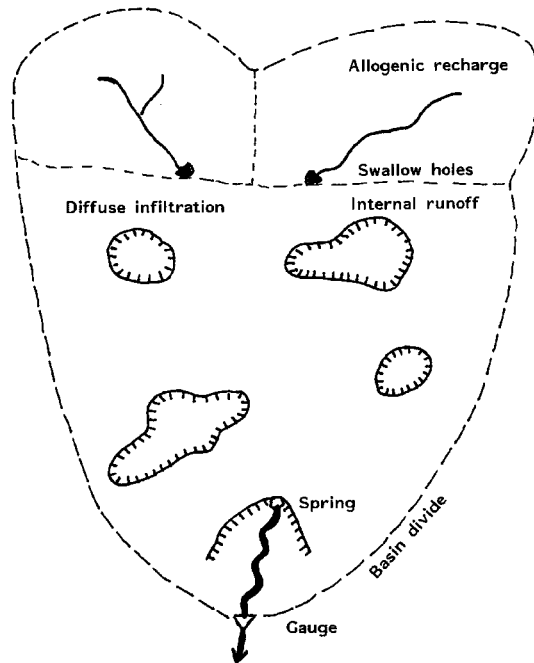
FIGURE 18.1 Sketch map showing groundwater basin with areas of allogenic recharge, internal runoff, and diffuse infiltration.

In regions such as the Appalachian Plateaus, the Illinois Basin, and the Ozark Plateau, karstic carbonate aquifers underlie extensive plateaus capped with clastic rocks. Surface water catchments on the plateaus provide allogenic runoff that enters the carbonate aquifer from the top. These waters tend to be chemically aggressive and enlarge the fractures into open shafts, which in turn permit very rapid movement of water through the unsaturated zone.

18.2.3 Internal Runoff

Karst terrains are characterized by closed depressions often so closely packed that overland storm flow is entirely captured by the depressions and does not reach surface stream channels. These depressions, known as sinkholes or dolines, contain a drain at the bottom which often efficiently carries water into the underlying aquifer (Figure 18.2). The drains serve to localize the recharge into small tributary channels of the underground drainage system. This component of the recharge, called internal runoff, is chemically distinct from diffuse infiltration and mainly contributes storm flow to the karst aquifer.

The drains can become plugged with soil so that storm runoff is temporarily ponded. The plugs can be sufficiently impervious that the closed depressions contain sinkhole ponds which remain filled even during drought periods.

Calculation of the contribution of internal runoff to the overall water balance is made difficult by the soil that mantles many closed depressions. Extended periods of gentle rainfall saturate the soils and introduce a diffuse infiltration through the walls of the depression. Intense storms may be required to produce enough overland flow to reach the bottom of the depression.

18.2.4 Diffuse Infiltration and the Epikarst

Storm water infiltrates slowly downward as the soil becomes saturated. In karst terrain, the soils are strongly leached, with few carbonate rock fragments remaining in the soil. The C-horizon is usually missing so that there is a sharp contact between the B-horizon and unweathered limestone or dolomite
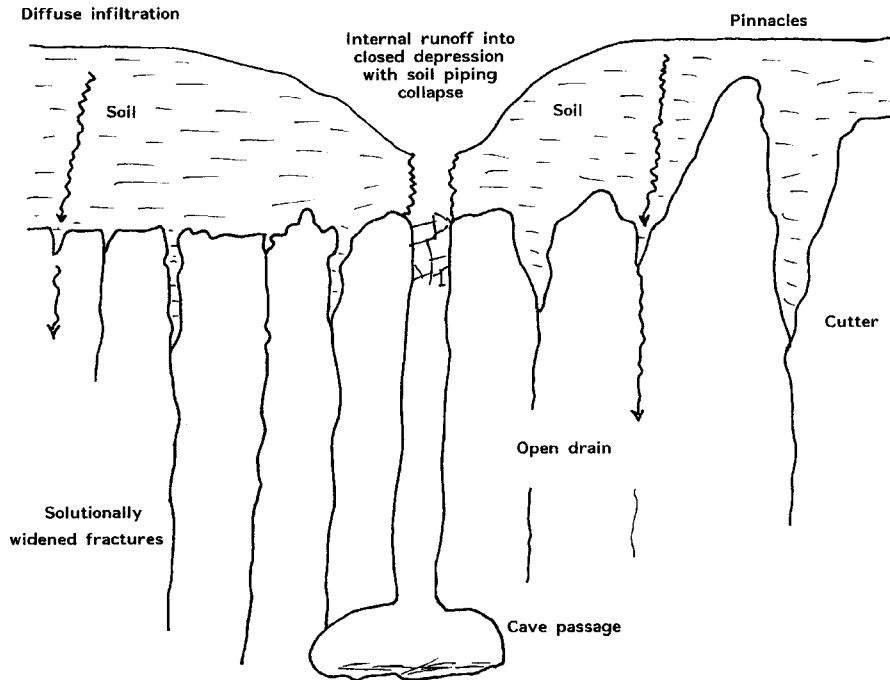
FIGURE 18.2 Cross-section of karst surface showing closed depression with internal runoff and the epikarst collecting diffuse infiltration.

bedrock. However, the bedrock surface is often highly irregular with deep channels and trenches dissolved along fractures. Between the fractures are irregular pinnacles of residual bedrock. Cavities may occur within the bedrock mass. The zone of soil plus irregular bedrock channels and pinnacles has a thickness on the order of meters to tens of meters and is known as the epikarst[5,6] (Figure 18.2).

The bedrock usually has a low permeability so that the infiltrating water can reach the subsurface only by moving laterally along the crevices and trenches until it reaches open fractures or fracture intersections. Because the irregular arrangement of channels and trenches are soil filled, flow rates are slow and storm water may be stored for up to several weeks. The epikarst is the reservoir that supplies cave dripwater during drought periods.

The catchment area for diffuse infiltration is simply the remaining area of exposed carbonate rock within the defined groundwater basin after the internal runoff catchments have been deducted. The proportion of the overall catchment assigned to internal runoff and to diffuse infiltration varies depending on storm intensity.

18.2.5 Infiltration Pathways in the Vadose Zone

Movement of water through the unsaturated zone is highly anisotropic and often very rapid. Pathways through the vadose zone include:

1. Slow percolation through the matrix porosity of the bedrock. This component is important in young limestones, as in Florida and the Bahamas for example, but is usually negligible in the dense Paleozoic limestones of the eastern U.S.
2. Vertical movement through fractures at the base of the epikarst
3. Solutionally widened fractures and open chimneys that act as sinkhole drains

4. Larger shafts that carry the runoff from caprock upland surfaces and the discharge from overlying perched aquifer systems
5. Open channels that carry the water from sinking streams

Allogenic streams often sink into open cave systems at elevations well above regional base levels. The underground streams descend rapidly, often in an open channel with a stair-step profile. In high-relief terrains the profile may consist of a sequence of steep channels with interspersed waterfalls. These input channels are often large enough for human exploration and many have been mapped. The explorable allogenic stream inputs and some large closed depression drains have been classified as vadose caves. The internal relief varies from a few meters in low-relief terrains to more than 1000 meters in high-relief terrains.

Water draining from upland areas underlain by clastic rock is injected into vertical pathways when it reaches the top of the carbonate aquifer. Because allogenic recharge tends to be highly aggressive, dissolution by the water moving along vertical pathways enlarges the initial fractures into shafts with diameters of meters to tens of meters and with depths of tens to hundreds of meters. Some shafts have irregular shapes dictated by the shape of the fracture system that provided the initial pathway. With time, the shafts take the form of near perfect cylinders with vertical walls that slice across beds of varying lithology.[7]

Drains from closed depressions have dimensions from centimeters to meters and are sometimes large enough to permit human exploration. Flow velocities are high and sufficient to transport sediment. Soil piping processes in closed depressions and above open fractures are an important source of sediment injected into the conduit system. The walls of closed depression drains are often jagged and corroded showing the highly aggressive chemistry of the water collected in the closed depression and injected into the drain.

The Paleozoic carbonate aquifers of the eastern U.S. are highly fractured. The fractures provide effective pathways for infiltration water into the subsurface. If the chemical reaction between the aggressive water of the epikarst with the carbonate bedrock goes to equilibrium, there may be relatively little dissolutional enlargement of these fractures. If the fractures are closely spaced, they provide a very important pathway for diffuse infiltration. Because of the long residence time at the base of the epikarst, the fracture water is near saturation with calcium carbonate.

18.2.6 Subsurface Drainage Patterns

The mixed groundwater/surface water character of karst aquifers is best revealed in the flow pattern in the subsurface. Instead of a uniform flow field mapped by isopotential contours, groundwater in karstic aquifers moves through an interconnected system of fractures, small solution channels, and large conduits. Travel times in the conduit system are often very short, hours to days per kilometer. The water may flow as open channel underground streams or as pipe flow with varying degrees of pressure head. Gradients within the matrix and fracture flow portions of the aquifer are oriented toward the conduits.

Sinking surface streams for the most part inject water directly into the conduit system. Drains from closed depressions also feed into the conduit system. Diffuse infiltration recharges the matrix and fracture system which in turn discharges into the conduit system. Each of these subsystems has different storage capacities and travel times.

18.2.7 Groundwater Discharge: Karst Springs

Groundwater discharge from karst aquifers is through high discharge springs. Sometimes a groundwater basin discharges through a single spring, and sometimes the water is distributed over multiple springs including wet-weather spillover springs. As more tracer data accumulates, groundwater distributed to multiple springs is being shown to be common. Diffuse discharge of groundwater in valley bottoms certainly occurs, but in well-developed karst aquifers, it constitutes only a small portion of the total discharge.

As an example, measurements show that 95% of the discharge from the groundwater basins in the south central Kentucky karst aquifer is through a few large springs.[8] The springs along the Green River in south central Kentucky can be divided into a small number of large regional springs draining groundwater basins of more than 100 km², a somewhat larger number of valley drain springs with catchments of a few tens of km², and a large number of very small local springs fed by small catchment areas near the river. Water balance shows, however, that the large springs account for most of the discharge.

18.2.8 Geologic Settings

The hydrologic components of the karst groundwater system must be superimposed on the geologic framework of the drainage basin because much of the detail in the pattern and development of the internal drainage system is guided and controlled by the local geology. The details of the geologic framework are highly site specific, but the main factors can be listed.

1. Physiographic factors
   (a) Placement of the karst aquifer within the regional drainage basin
   (b) Relief between recharge area and springs
2. Structural factors
   (a) Folded vs. nonfolded rocks
   (b) Dip with respect to groundwater flow directions
   (c) Faults
   (d) Extent of joint and fracture development
3. Stratigraphic and lithologic factors
   (a) Thickness of carbonate rocks
   (b) Location of carbonate sequences in the regional drainage basin
   (c) Lithologic characteristics of carbonate rock (pure limestones, shaley limestones, dolomitic limestones, dolomites)
   (d) Presence, location, and thickness of shale or sandstone beds within the carbonate sequence

Conduit development seeks out optimum flow paths through the available fractures and bedding plane partings. Conduits are deflected by relatively minor impermeable layers and by unfavorable lithologies. Fine-grained (micritic) pure limestones are favorable; shaley limestones, dolomitic limestones, and dolomites are unfavorable. Folded and faulted rocks guide possible flow paths by controlling the placement of favorable rock lithologies. The final arrangement of internal flow paths represents a balance between hydraulic gradients and the geologic constraints.

18.2.9 The Conceptual Model

The hydrogeologic concepts can be reduced to a conceptual model based on overall water balance and internal flow paths. Using the groundwater basin concept, the spring is an effective gauge point from which to determine the overall discharge from the aquifer.

The first part of the conceptual model is the water budget (Figure 18.3) which can be written

$$Q_B = Q_A + Q_R + Q_I + Q_D - \Delta Q_S \tag{2}$$

where Q_B is the overall basin discharge as gauged at the spring or at the base level surface stream carrying all runoff from combined surface and groundwater basin.

Q_A is allogenic recharge. The combined runoff from surface catchments drained by sinking streams.

Q_R is surface runoff. The discharge provided by surface streams tributary to the base level stream which do not sink into the karst.
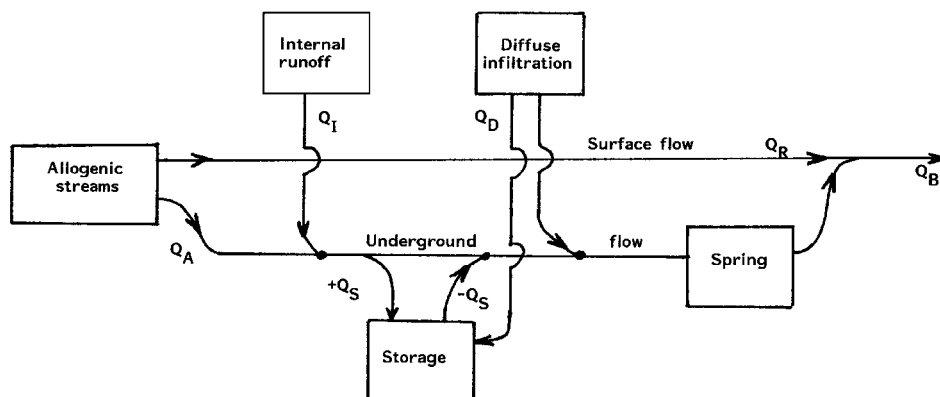
FIGURE 18.3 Flow chart showing groundwater budget in karst aquifers.

Q_I is internal runoff. Overland flow captured by closed depressions and injected into the aquifer through a drain at the bottom.

Q_D is diffuse recharge. Infiltration through soil and bedrock.

ΔQ_S is change in storage within the karst aquifer.

All terms in the water budget equation are complicated functions of time and discharge. Internal runoff only contributes during storms. Except for sinkhole ponds, closed depressions drain rapidly when rainfall ceases. The contribution from surface flow is oversimplified by the terms Q_A and Q_R. One must distinguish surface tributaries in blind valleys that inject all of their water into the aquifer regardless of discharge, losing streams that remain on the surface but lose part of their flow to the subsurface, and sinking streams that overflow into surface routes. Equation (2) provides some guidance for writing a water budget, but the details will be highly site specific.

The second part of the conceptual model is the arrangement of internal flow paths (Figure 18.4). Again, these flow path components are schematic and generalized. Allogenic recharge and internal runoff feed mainly into the conduit system. Throughput times are short. A well-developed conduit system exhibits a fast response to storm recharge. Rising and falling stage levels in the conduit system force the exchange of water with the matrix and fracture system. Diffuse infiltration recharges mainly the matrix and fracture system. There is a substantial time lag due to water stored in the epikarst. Overall throughput times are long, generally longer than the mean spacing between storms.

Not all of these components are present in all karst aquifers. There are carbonate aquifers to which there is no allogenic recharge and in which the conduit system is absent. These are the diffuse flow aquifers. In some aquifers, notably the aquifers in the young limestones of Florida and the Caribbean, matrix flow is a dominant component of the groundwater system. In the dense Paleozoic limestones of the eastern U.S., matrix flow rarely makes much of a contribution.

18.3 Permeability and Flow Dynamics in Karst Aquifers

A central requirement for any theoretical description of groundwater flow is that the permeability remains a fixed quantity characteristic of particular rock units. Discontinuities in permeability across geologic boundaries can be accommodated. Likewise, expressing permeability in tensor form accounts for the anisotropy of bedded rock units. Within these constraints, permeability remains scale-independent and time-independent. Aquifer behavior can be described theoretically or modeled by computer programs with reasonable accuracy.
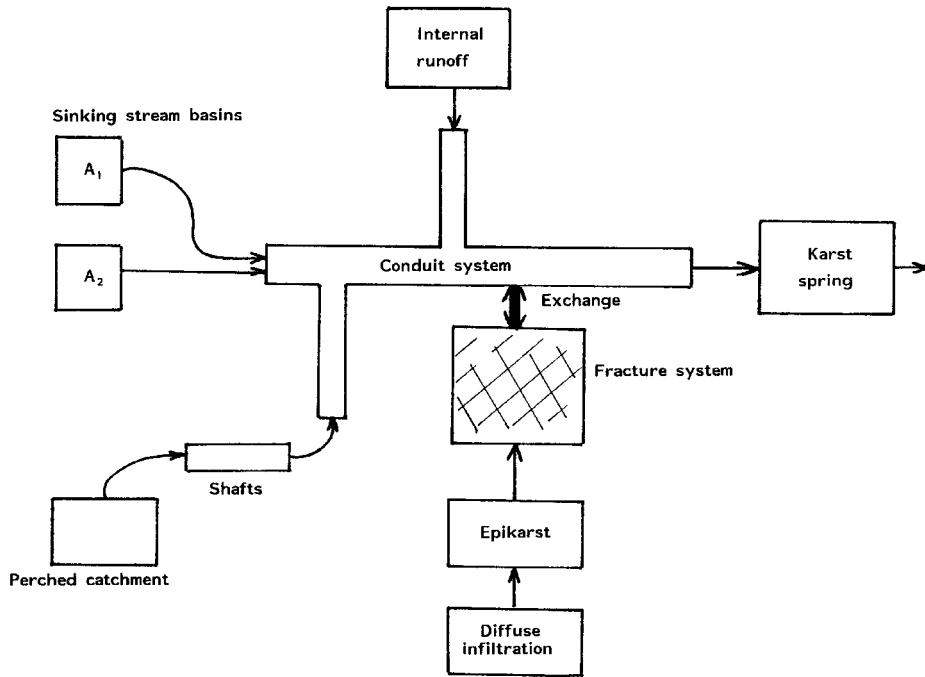
FIGURE 18.4 Internal flow paths within karst aquifers.

The permeability of karst aquifers consists of three overlapping categories: matrix permeability, fracture permeability, and conduit permeability. It is the conduit permeability that sets karst aquifers apart from other water-bearing formations (Table 18.1).

As a consequence of the permeability distribution, the hydraulic conductivity varies by many orders of magnitude on different size scales.[9] Karst aquifers are not only anisotropic with respect to orientation

TABLE 18.1.A Classification of Permeability

	Porous Media Aquifers	Fracture Aquifers	Karstic Aquifers
Matrix Permeability	Pores and vugs	Pores and vugs	Pores and vugs
Fracture Permeability		Mechanical joints, fractures, and bedding plane partings	Solutionally enlarged joints, fractures, and bedding plane partings Aperture < 1 cm
Conduit Permeability			Integrated conduits of various sizes from 1 cm to 10's of meters

TABLE 18.1.B Flow Characteristics of Permeability Types

Characteristic	Matrix Permeability	Fracture Permeability	Conduit Permeability
Regime	Laminar, Darcy flow	Laminar, may be deviations from linear behavior	Turbulent, nonlinear
Response to storms	Slow	Moderate	Fast

but they are also anisotropic with respect to measurement scale and with respect to the specific point in the aquifer where a measurement is to be made.

1. Laboratory scale, 0.1 m. Permeability tests on core samples. Most likely to measure permeability of rock matrix.
2. Test block scale, 1 to 3 m. Matrix permeability plus some fracture permeability.
3. Borehole scale, 3 to 100 m. Pumping, injection, and packer tests. Hugely variable results depending on location of borehole. Result could include pure matrix flow, fracture flow, or possibly conduit flow.
4. Groundwater basin scale, 1 to 10 km. Methods have not been developed for total basin modeling although analysis of spring discharge has promise. Only at the basin scale are all three contributions to the permeability taken into account.

18.3.1 Matrix Flow and Fracture Flow

Groundwater flow through the bulk rock mass and through fractures that have not been significantly modified by dissolution is not different in principle from matrix flow or fracture flow in other kinds of aquifers. What is different is that hydraulic gradients within the matrix and fracture system are often directed toward the conduit system rather than to the surface stream where the groundwater is discharged.

18.3.2 Conduit Permeability: Flow Dynamics

For reasons to be discussed later, the aperture size marking the transition from fracture permeability to conduit permeability is taken to be one centimeter. Within a mature karst aquifer is a network of solutionally widened fractures and bedding plane partings, small solution channels with dimensions of centimeters to tens of centimeters, and conduits with dimensions of meters to tens of meters. All of these are interconnected in complex ways, sometimes in an ordered branchwork pattern as an underground analogy of a surface stream, but more often with multiple interconnections of closed loops. The conduits have very low hydraulic resistance and can carry large volumes of water at high velocities. Knowledge of the conduit system is mainly obtained by tracer tests and by direct observation and survey. Given 0.5 meter as a practical lower limit for human exploration, there is a very large portion of the conduit permeability in the scale range from 0.01 to 0.5 meter which cannot be reliably probed by any presently available technique.

Because hydraulic conductivity has units of velocity, there is a temptation to equate travel time to an effective "hydraulic conductivity." This is erroneous. Use of hydraulic conductivity implies a linear relationship between hydraulic head and discharge. Most flow in conduits is turbulent. Further, the travel time represents an average. The actual flow path may contain reaches of high-velocity flow interspersed with reaches where the water is ponded.

Conduit geometry is highly variable with cross-sections ranging from uniform elliptical shapes, to high and narrow fissures, to irregular shapes partly filled with rock fragments. Most conduits contain a bed of clastic sediments on their floors and some of these are armored with cobble and boulder-size material in the manner of surface streams. Water may flow in conduits in an open-channel regime subject only to gravitational gradients, or as pipe flow under substantial pressure heads. Both regimes frequently occur in different reaches of the same conduit. Rising and falling water levels in response to storm recharge can shift the regime from open channel flow to pipe flow and back again. Passages that flood to the ceiling during storms are well known to cave explorers.

The flow regime in closed pipes is described by a Reynolds number defined as

$$N_R = \frac{\rho v R}{\eta} \qquad (3)$$

where ρ = density of fluid, η = viscosity of fluid, v = velocity, and R is the hydraulic radius. Although the hydraulic radius for circular pipes is often defined as the pipe diameter, it is more useful to define the hydraulic radius of irregular karst conduits as the cross-sectional area divided by the perimeter. With this definition R = D/4 (D = pipe diameter) with a corresponding scaling of the numerical value of the Reynolds number for circular conduits. If Equation (3) is written in consistent units, the Reynolds number is dimensionless. At low Reynolds numbers, the flow regime will be laminar; at high Reynolds numbers it will be turbulent. The transition from laminar to turbulent occurs over a range of Reynolds numbers. The onset depends on the wall roughness of the conduit and will be in the range of N_R = 50 to 500.

In the laminar regime, the fundamental energy balance equations can be solved exactly leading to a description of head loss along a reach of conduit as

$$h_f = \frac{12\eta vL}{\rho g B^2} \qquad \text{Parallel walled fracture} \tag{4}$$

$$h_f = \frac{8\eta vL}{\rho g r^2} \qquad \text{Circular conduit} \tag{5}$$

$$h_f = \frac{\eta vL}{\rho g\left(Nd^2\right)} \qquad \text{Porous medium} \tag{6}$$

where h_f = head loss, η = viscosity of fluid, L = length of conduit, ρ = density of fluid, and g = gravitational acceleration. B = aperture of parallel walled fracture, r = radius of circular conduit, and d = grain size in porous medium with N as a geometrical shape factor. The head loss equation for porous media is simply a restatement of Darcy's law. The challenge in modeling the matrix and fracture permeability portion of a karst aquifer is to find an adequate description of fracture spacings and the distribution of fracture apertures. Equation (4) provides an accurate description of the flow dynamics.

There is no corresponding equation for turbulent flow derived from first principles. Instead, the usual starting point is an empirical relation, the Darcy–Weisbach equation

$$h_f = \frac{fLv^2}{4gr} \tag{7}$$

where f is a dimensionless friction factor. The Darcy–Weisbach equation can also be applied to laminar flow for which the friction factor is a function of the Reynolds number

$$f = \frac{64}{N_R} \tag{8}$$

If the wall of the conduit is smooth, the friction factor is given by the Prandl–Von Karman equation

$$\frac{1}{\sqrt{f}} = 2\log\left(N_R\sqrt{f}\right) - 0.8 \tag{9}$$

Turbulent flow over a smooth wall produces a laminar layer separating the mass of the fluid from the rock wall. The wall roughness of rock conduits depends on the irregularities on the conduit wall and on the lithologic characteristics of the bedrock. The thickness of the laminar layer, δ, is determined by the dimensionless equation

$$\frac{\delta}{2r} = \frac{32.8}{N_R\sqrt{f}} \tag{10}$$

The relief of surface irregularities, e, is given in the same units as the conduit dimensions. A rough surface is one where the surface roughness, e, is greater than the thickness of the laminar sublayer. For rough surfaces the friction factor is given by

$$\frac{1}{\sqrt{f}} = 2\log\frac{2r}{e} + 1.14 \tag{11}$$

Depending on hydraulic radius, hydraulic gradient, and wall roughness, three flow regimes may occur in karst conduit systems:

1. Laminar flow
2. Turbulent flow with laminar sublayer
3. Totally rough flow at high Reynolds number

Open-channel flow is described by the Manning equation

$$v = \frac{M}{n}R^{2/3}S^{1/2} \tag{12}$$

where v = mean velocity, R = hydraulic radius, and S = channel slope. The constant, n (known as "Manning's n"), is a roughness factor. It has been the custom to let n have the same numerical values regardless of units. The parameter, M, adjusts the units. $M = 1$ for velocity in m/sec and hydraulic radius in m. $M = 1.486$ for velocity in ft/sec and hydraulic radius in ft. The hydraulic radius of an irregular channel is defined as the cross-sectional area divided by the wetted perimeter. Values for Manning's n have been determined for a range of surface stream bed materials[10,11] which are the best available approximation for underground channels (Table 18.2). No direct experimental determinations of roughness characteristics of karst conduits seem to have been made.

TABLE 18.2 Manning's n for Channels

Bed Material	n[10,11]
Smooth concrete	0.010
Good ashlar masonry	0.013
Rubble masonry	0.017
Firm gravel	0.020
Clay	0.026
Clay and silt	0.046
Sand and clay	0.030
Sand and gravel	0.042
Gravel over rock	0.027
Coarse sand with some gravel	0.049
Gravel, rock, and boulders	0.055
Large angular boulders	0.079

Water may flow in open channels in either laminar or turbulent regimes, although most underground free surface streams have dimensions and velocities that place them in the turbulent regime. Open-channel flows are also subject to a balance between inertial forces and gravity forces described by a dimensionless Froude number

$$N_F = \frac{v}{\sqrt{gD}} \tag{13}$$

where D is the hydraulic depth (equal to the water depth in a rectangular channel). Flows with $N_F < 1$ are termed subcritical; flows with $N_F > 1$ are termed supercritical. The transition from supercritical to subcritical flow is abrupt, with considerable release of energy producing a hydraulic jump.

Most underground streams flow in a subcritical regime, but supercritical flows occur in steep gradient vadose streams carrying water from allogenic catchments down to base level. Supercritical flows also occur in the shafts and drains that carry internal runoff through the unsaturated zone down to the water table.

18.3.3 Caves and Conduits: Hydrologic Interpretation of Cave Patterns

It is important to distinguish between caves and conduits. Conduits are solutionally formed pathways on many size scales that serve to transmit water from recharge points to discharge points. Some portions of the conduit system are below base level and are continuously water-filled. Some portions lie between the elevations of base flows and flood flows and are therefore used intermittently depending on ground-water stage. Still other portions are abandoned as base levels are lowered and groundwater levels fall below the elevation of the conduit.

Caves are those fragments of the conduit system that are accessible to human exploration. Cave explorers traditionally include all accessible passages as part of the same cave. Thus a cave map may show fragments of several conduits interconnected by collapse or secondary solution. The cave map may also include conduit segments of greatly different ages ranging from abandoned high-level passages that may date from the early Pleistocene to vertical shafts that are part of the contemporary groundwater system. Cave maps, in brief, are composites of all parts of conduit systems that are interconnected and which can be accessed by cave exploration. Although they provide valuable data on the conduit system, they must be interpreted with some care.

Although cave maps reveal a great diversity of patterns, some common features can be recognized. The important variables are the local geologic structure which determines many of the details of the cave pattern, the overall hydraulic gradient of the drainage basin, and the relative contributions of allogenic, internal runoff, and diffuse infiltration recharge.[12] Some of the possibilities are sketched in Figure 18.5.

The conduit systems within aquifers dominated by recharge through internal runoff often exhibit branchwork patterns as small tributaries draining individual closed depressions merge to form larger drainage channels which ultimately lead to a spring. The smallest tributaries of these systems are often too small for human exploration. Branchwork patterns require sufficient gradient to allow a free-flowing drainage to develop. There may appear closed loops and spillover routes used during periods of high discharge. In detail, the branching tributaries are guided by joints or bedding plane partings.

If the aquifer is dominated by allogenic recharge, the underground drainage may be through a single large conduit with few tributaries. In the limit, a large surface catchment could have a short segment of underground route that would be almost completely decoupled from the groundwater system. These systems are most like underground rivers and least like aquifers. Large conduits can be nearly linear, can be angulate following structural features in the bedrock, or can be sinuous.

Caves frequently occur with a maze pattern consisting of many closed loops along the same flow direction. Maze caves may take the form of anastomotic mazes or network mazes depending on whether the fracture permeability is dominated by bedding plane partings or by joints. Both require that the dissolutional process proceed along all possible pathways without the concentration of flow along a single pathway that leads to branchwork and single conduit caves. Anastomotic mazes generally have their greatest cross section developed along bedding planes. Freedom for water to select optimum pathways along the bedding plane parting leads to the gently curving loops on many size scales. Network mazes generally have their greatest cross section developed along joints so that the pattern is constrained by the
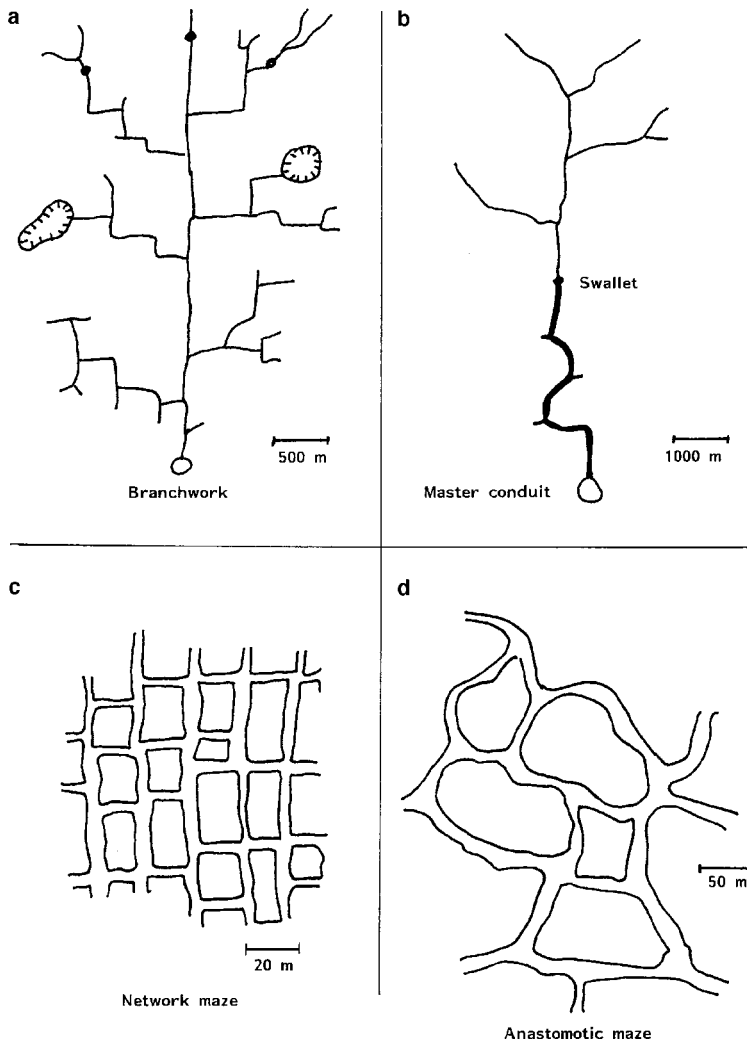
FIGURE 18.5 Patterns of conduit drainage system: (a) Branchwork pattern. (b) Master conduit dominated by allogenic recharge. (c) Network maze. (d) Anastomotic maze.

joint pattern to produce the characteristic "city block" pattern. Anastomotic mazes generally occur in low gradient systems in nearly flat-bedded rock where many percolation paths would be possible. Network mazes are sometimes related to diffuse infiltration through an overlaying permeable caprock into a jointed limestone mass below the water table.[13] Other network mazes are the result of backflooding from surface streams, a form of bank storage. In a different category are floodwater mazes that occur in very high gradient conduit systems with large contributions of allogenic recharge where high hydraulic heads from flood flows can force water along all possible joints and fractures.

To a certain extent, branchwork patterns are the subsurface analogs of surface tributary stream patterns. Single conduit patterns are the subsurface analogs of high-order surface streams. Anastomotic maze patterns may be taken as the subsurface equivalent of braided rivers, but most generally, maze patterns are the subsurface equivalent of swamps.

18.3.4 Paleohydrology

Unlike deepening surface valleys and their drainage systems which usually destroy their past as the valleys downcut, caves provide a useful record of the history of the conduit drainage system. Vertical sequences of abandoned conduit fragments (tiered caves) may reveal changing drainage patterns and changing discharges. To be useful, the paleohydrologic record must be tied to the time sequence of the evolving surface drainage basin. U/Th dating of calcite speleothems, useful in the age range back to 0.5 Myr, can place some limits on the ages of various passage levels.[4] Magnetic reversals preserved in cave sediments, particularly the reversal at 0.72 Ma, provide absolute time markers when they can be found and provide some insight into base level lowering rates.[14] Geophysical, geochemical, and geomorphological evidence suggests that the karstic drainage basins of the eastern U.S. with 100 to 200 meters of internal relief have been evolving since the late Pliocene with some evidence for systems extending back into the Miocene.

Cave walls often exhibit a pattern of scallops (Figure 18.6) which record the paleoflow velocity. The scallop's length is inversely related to velocity through a Reynolds number.[15]

$$\frac{\rho v L_{32}}{\eta} = N_R = 22,500 \tag{14}$$

where L_{32} is the Sauter mean of a population of measured scallop lengths

$$L_{32} = \sum_i \frac{\ell_i^3}{\ell_i^2} \tag{15}$$

A detailed analysis of the transport processes shows a dependence on conduit size (Figure 18.7). Measurement of the Sauter mean on a population of scallops gives a paleoflow velocity, which, combined with conduit cross-sectional area, gives a paleodischarge. With some assumptions the paleodischarge allows an estimate of paleocatchment areas. Calculations of hydraulic parameters from scallops in both paleoconduits and active conduits have produced reasonable agreement with the parameters determined by other measurements.[16-18]
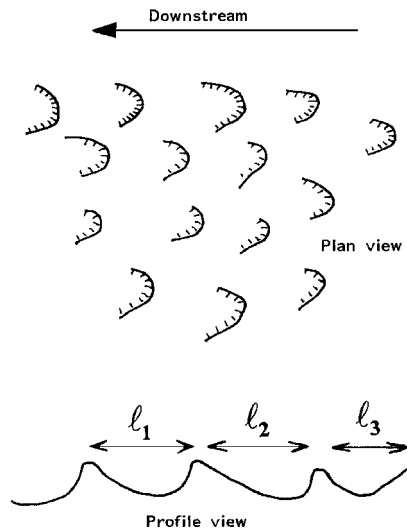
FIGURE 18.6 Plan and profile of a scalloped cave wall. Characteristic lengths of scallops are measured along the flow direction.
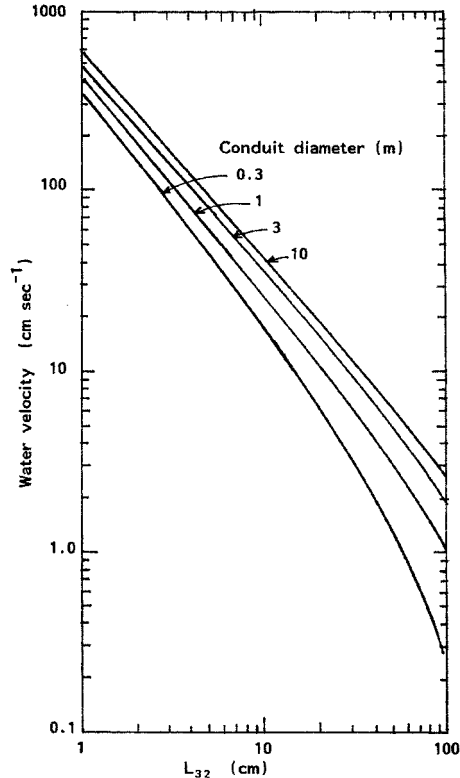
FIGURE 18.7 Calculated water velocity as a function of mean scallop length. L_{32} = Sauter mean.

18.4 Chemistry of Carbonate Rock Dissolution

18.4.1 Chemical Equilibrium

The primary process in karstic weathering and in the development of fracture and conduit permeability in karst aquifers is the dissolution of carbonate rocks in weak solutions of carbonic acid. Carbonate chemistry has been extensively investigated, and more detail may be found in Chapter 4 of Drever[19] and Chapter 6 of Langmuir.[20]

Carbon dioxide occurs in the atmosphere at a volume fraction of 3.47×10^{-4}. This reservoir provides a constant atmospheric background that CO_2 partial pressures in karstic waters rarely fall below. Much greater concentrations of CO_2 are generated in the soil due to respiration from roots, from decay of organic matter, and from the action of microorganisms. CO_2 generation in the soil is a sensitive function of temperature and, in temperate and subpolar climates, exhibits a seasonal periodicity. The volume fraction of CO_2 in the soil varies from 0.01 to 0.1 with excursions to both greater and lesser values. Some of the soil CO_2 diffuses upward to be lost to the atmosphere, while some dissolves in infiltrating groundwater.

Carbonic acid is formed from gaseous CO_2 by a two-step process

$$CO_2\left(gas\right) \leftrightarrow CO_2\left(aqueous\right)$$

$$CO_2\left(aqueous\right) + H_2O \leftrightarrow H_2CO_3$$
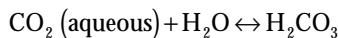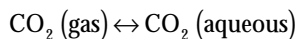
Molecular carbonic acid is a minor species and comprises only 0.17% of the dissolved CO_2. Nonetheless, it is convenient to consider both reactions together and write the equilibrium constant as

$$\frac{a_{H_2CO_3}}{P_{CO_2}} = K_{CO_2} \tag{16}$$

H_2CO_3 ionizes

$$H_2CO_3 \leftrightarrow H^+ + HCO_3^-$$

$$HCO_3^- \leftrightarrow H^+ + CO_3^{2-}$$
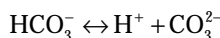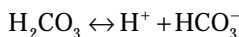
with the equilibrium expressions

$$\frac{a_{H^+} a_{HCO_3^-}}{a_{H_2CO_3}} = K_1 \tag{17}$$

$$\frac{a_{H^+} a_{CO_3^{2-}}}{a_{HCO_3^-}} = K_2 \tag{18}$$

Calcite (limestone) is only slightly soluble in pure water.

$$CaCO_3 \leftrightarrow Ca^{2+} + CO_3^{2-}$$
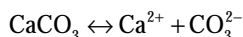
$$a_{Ca^{2+}} a_{CO_3^{2-}} = K_c \tag{19}$$

Solving Equation (19) gives the concentration of dissolved calcite as 6 to 7 mg/L, less than the solubility of quartz. Likewise, the dissolution of dolomite is described by the reaction

$$CaMg(CO_3)_2 \leftrightarrow Ca^{2+} + Mg^{2+} + 2CO_3^{2-}$$
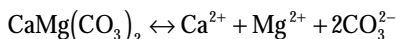
$$a_{Ca^{2+}} a_{Mg^{2+}} a_{CO_3^{2-}}^2 = K_d \tag{20}$$

The equilibrium constants for Equations (16) through (20) have been refined and evaluated.[21] Accepted values are given in Table 18.3.

In overall terms, the dissolution of limestone and dolomite are described by the reactions

$$CaCO_3 + CO_2 + H_2O \leftrightarrow Ca^{2+} + 2HCO_3^-$$

$$CaMg(CO_3)_2 + 2CO_2 + 2H_2O \leftrightarrow Ca^{2+} + Mg^{2+} + 4HCO_3^-$$
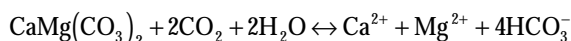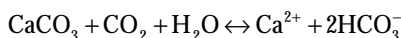
To actually calculate the concentration of dissolved carbonate, one must combine the various reactions. Three cases are of interest:

1. Available CO_2 is a fixed quantity (closed system).

TABLE 18.3 Selected Equilibrium Constants for Carbonate Reactions

T (°C)	log K_{CO_2}	log K_1	log K_2	log K_c	log K_d
0.0	–1.11	–6.58	–10.63	–8.38	–16.56
5.0	–1.19	–6.52	–10.55	–8.39	–16.63
10.0	–1.27	–6.46	–10.49	–8.41	–16.71
15.0	–1.34	–6.42	–10.43	–8.43	–16.79
20.0	–1.41	–6.38	–10.38	–8.45	–16.89
25.0	–1.47	–6.35	–10.33	–8.48	–17.00
30.0	–1.52	–6.33	–10.29	–8.51	–17.12
35.0	–1.58	–6.31	–10.25	–8.54	–17.25
40.0	–1.68	–6.30	–10.22	–8.58	–17.39

Data extracted from a compilation by White, W. B., *Geomorphology and Hydrology of Karst Terrains*, Oxford University Press, New York, 1988.

2. Carbon dioxide partial pressure is fixed by the environment (open system).
3. pH is fixed by the environment.

Case (2) is of most interest in karst hydrology because the CO_2 partial pressure is usually fixed by the rate of CO_2 generation in the soil. The calculations (see Drever,[19] Chap. 4 for detail) give the concentration of dissolved Ca^{2+} ions (as molal concentration)

$$m_{Ca^{2+}}^3 = P_{CO_2} \frac{K_1 K_c K_{CO_2}}{4 K_2 \gamma_{Ca^{2+}} \gamma_{HCO_3^-}^2} \tag{21}$$

If the groundwater is saturated with calcite at a given CO_2 partial pressure, the pH is also fixed

$$a_{H^+}^3 = P_{CO_2}^2 \frac{K_1^2 K_2 K_{CO_2}^2 \gamma_{Ca^{2+}}}{2 K_c \gamma_{HCO_3^-}} \tag{22}$$

Allowing for the temperature dependence of the equilibrium constants, Equation (21) describes the solubilty of calcite as a function of the carbon dioxide partial pressure (Figure 18.8). The solubility, expressed as concentration of $CaCO_3$ increases from 6 mg/L in the absence of CO_2 to 50 mg/L in equilibrium with atmospheric CO_2, to several hundred mg/L in equilibrium with the CO_2 pressures found in many soils.

18.4.2 Chemical Parameters for Karst Aquifers

As a practical tool in karst hydrology, neither the underlying physical chemistry as described above nor raw chemical analyses of groundwater is directly useful. Parameters that have been helpful for aquifer interpretation are the hardness, the Ca/Mg ratio, the saturation indices for calcite and dolomite, and the calculated CO_2 partial pressure.

Hardness is a measure of the total dissolved carbonate, usually expressed as mg/L $CaCO_3$. Hardness is defined as

$$Hd = \left(\left[Ca^{2+}\right] + \left[Mg^{2+}\right]\right) MW_{CaCO_3} \tag{23}$$

The quantities in brackets are the analyzed concentrations of Ca^{2+} and Mg^{2+} expressed as molar concentration. MW_{CaCO_3} is the molecular weight of $CaCO_3$. Although of no theoretical significance, hardness is a useful parameter for describing total concentration of dissolved carbonates.
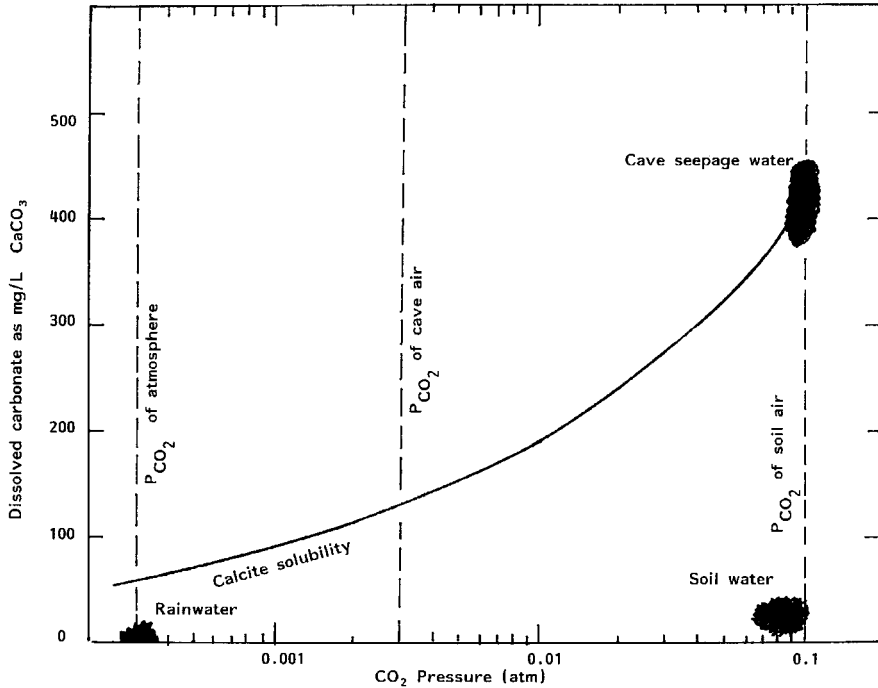
FIGURE 18.8 Solubility of calcite as a function of CO_2 partial pressure. The temperature has been set to 10°C. Chemistry of diffuse infiltration water is shown at three points along its flow path.

The Ca/Mg ratio is simply the relative concentrations of Ca^{2+} and Mg^{2+} on an atomic scale.

$$\frac{Ca}{Mg} = \frac{\left[Ca^{2+}\right]}{\left[Mg^{2+}\right]} \tag{24}$$

The Ca/Mg ratio provides information on the rock units through which the groundwater has passed. The Ca/Mg ratio = 1 to 1.5 for dolomite aquifers and 6 to 8 for limestone aquifers. Intermediate values imply a dolomitic limestone or a mixed limestone/dolomite sequence.

The degree of disequilibrium between groundwater and carbonate wall rock is given by the saturation index defined as

$$SI_c = \log \frac{a_{Ca^{2+}} a_{CO_3^{2-}}}{K_c} = \log \frac{K_{iap}}{K_c} \tag{25}$$

where K_{iap} is the activity product determined from actual analyses of the water and K_c is the calcite solubility product constant. If the water is exactly at equilibrium the experimental ion activity product will be exactly equal to the solubility product constant and $SI_c = 0$. If the water is undersaturated with respect to calcite, K_{iap} will be less than K_c and the saturation index will be negative. For a supersaturated water, SI_c will be positive.

Calcium ion activity can be determined directly from the analyzed calcium concentration. The carbonate ion, however, is a minority species in most carbonate groundwaters. Its activity must be calculated from the measured bicarbonate ion concentration and the pH. A final expression for saturation index as a function of experimentally determined quantities is

$$SI_c = \log \frac{\gamma_{Ca^{2+}}\left[Ca^{2+}\right]\gamma_{HCO_3^-}\left[HCO_3^-\right]K_2}{10^{-pH}K_c} \tag{26}$$

See Drever,[19] White,[2] or Langmuir[20] for a full discussion of the thermodynamics.

The effective concentration of carbon dioxide in groundwater is related to the CO_2 pressure in the source area through Equations (16) and (17). These reactions can be combined to allow the calculation of carbon dioxide pressure from measured quantities.

$$P_{CO_2} = \frac{\gamma_{HCO_3^-}\left[HCO_3^-\right]10^{-pH}}{K_1 K_{CO_2}} \tag{27}$$

Because the observed CO_2 pressures are rarely lower than the atmospheric background, it is often helpful to normalize the calculated CO_2 pressure to the atmospheric CO_2 pressure. This gives an enhancement factor that can be used to interpret source areas for the groundwater.

18.4.3 Dissolution Kinetics for Limestone and Dolomite

The central role of dissolution kinetics is a unique characteristic of karst aquifers. In porous media aquifers and in most fracture aquifers, travel times and surface/volume ratios are both large. The groundwater is in intimate contact with the mineral grains of the aquifer rock and remains in contact for long periods of time. As a result, there is sufficient time for chemical reactions to proceed to completion, and it is generally safe to assume that chemical reactions are at equilibrium. Geochemical modeling of aquifer chemistry can make use of speciation and reaction path programs such as WATEQ4F, PHREEQE, and MINTEQAZ.[22] In karst aquifers, travel times are measured in hours or days which is on the same order of magnitude as the reaction time of the water with the rock. As a result, karst groundwater can travel the entire width of the aquifer without coming into chemical equilibrium.

The characteristic time scales for the various carbonate reactions are:

1. Dissociation of carbonic acid into bicarbonate ions and bicarbonate ions into carbonate ions: milliseconds.
2. Hydration of aqueous CO_2 to form carbonic acid: 30 sec.
3. Dissolution of calcite: 2 to 3 days.
4. Dissolution of dolomite, 1st stage: 3 to 5 days.
5. Dissolution of dolomite, 2nd stage: months to years.

It can be assumed from (1) that the concentrations of H_2CO_3, HCO_3^- and CO_3^{2-} in solution are always in equilibrium. The 30-second time scale for producing carbonic acid from aqueous CO_2 is important for sheet flow of thin films of rapidly moving water in shafts and chimneys in the vadose zone. The dissolution kinetics of calcite and dolomite are central to the interpretation of the evolution of karst aquifers.

There have been a great many investigations of the dissolution kinetics of calcite and limestone in both carbonic and in mineral acids. A useful rate equation was developed by Plummer, Wigley, and Parkhurst.[23,24]

$$\text{Rate} = k_1 a_{H^+} + k_2 a_{H_2CO_3} + k_3 a_{H_2O} - k_4 a_{Ca^{2+}} a_{HCO_3^-} \tag{28}$$

where k_1, k_2, and k_3 are forward rate constants and k_4 is the backward rate constant. The rate is given in units of millimoles cm^{-2} sec^{-1}.

The first term dominates in strongly acid solutions and describes, for example, the interaction of acid mine water with limestone. This component of the rate is mass-transport controlled so that rate depends on rate of flow of water across the rock surface. The second term describes the dissolution process driven by carbonic acid. The third term describes a constant background rate of limestone dissolving in water. All three of the forward rate constants are known functions of temperature

$$\log k_1 = 0.198 - \frac{444}{T} \tag{29}$$

$$\log k_2 = 2.84 - \frac{2177}{T} \tag{30}$$

$$\log k_3 = -5.86 - \frac{317}{T} \quad \left(T < 298 \; K \right) \tag{31}$$

$$\log k_3 = -1.10 - \frac{1737}{T} \quad \left(T > 298 \; K \right) \tag{32}$$

The backward reaction constant depends on the forward rate constants

$$k_4 = \frac{K_2}{K_c} \left\{ k_1' + \frac{1}{a_{H^+}} \left[k_2 a_{H_2CO_3} + k_3 a_{H_2O} \right] \right\} \tag{33}$$

where K_2 and K_c are equilibrium constants (Table 18.3) and k_1' represents the hydrogen ion concentration at the dissolving interface. Experimenal comparisons have been made by setting $k_1' = k_1$.

Important to the interpretation of karst aquifer evolution is a change in rate mechanism that takes place when the saturation index exceeds $SI_c = -0.3$. Dissolution in more undersaturated waters is described by Equation (28). As the saturation index becomes greater than $SI_c = -0.3$ and the water appoaches saturation, the rate of dissolution drops precipitously. The final approach to equilibrium requires many days.

The dissolution of dolomite is not completely understood. Dolomite dissolves in highly unsaturated water following a rate equation of the form[25,26]

$$\text{Rate} = k_1 a_{H^+}^n + k_2 a_{H_2CO_3}^n + k_3 a_{H_2O}^n - k_4 a_{HCO_3^-} \tag{34}$$

where the exponent, $n = 0.5$ for temperatures below 45°C. However, unlike the calcite rate curve which approaches the equilibrium value on times on the order of a few days, the dolomite rate curve flattens out at a saturation index of $SI_d = -2$. A different and slower mechanism takes effect. The final approach to equilibrium requires times on the order of months to years. Groundwater in dolomite aquifers does eventually come into equilibrium with the wall rock because well waters from dolomite are found to be near saturation.

As a result of the slower dissolution kinetics, conduit systems are generally less well developed in dolomite aquifers. Dissolution tends to follow pathways through limestones if these are available. If there are no alternative pathways available, large conduits do develop in dolomite but appear to do so at a much slower rate.

18.5 Evolution of Karst Aquifers

18.5.1 Life History of a Conduit

A particular pathway through the mass of soluble rock will have a "life history" as sketched in Figure 18.9. It is assumed that the initial state is a mechanical opening in the bedrock: a sequence of joints, fractures, or bedding plane partings, more rarely a fault or a sequence of primary vuggy porosity. The *initiation* phase is that time during which the primary mechanical opening is enlarged by dissolution from its initial aperture of ten to hundreds of micrometers to a width or the order of 5 to 10 millimeters. When the aperture reaches 5 to 10 millimeters, the flow regime crosses a set of critical thresholds in the dissolution and transport processes so that a 10-mm aperture is the practical boundary between openings that would be considered part of the fracture permeability and openings that would be part of the conduit permeability.

Once the critical thresholds are passed, the conduit enters an *enlargement* phase and grows to dimensions typical of cave passages by the continued solutional retreat of its walls. As the enlargement phase comes to a close, several alternative scenarios are possible. If the conduit is so situated that it remains water-filled, solution will continue until the ceiling span becomes mechanically unstable and the passage will collapse. If the recharge source for the conduit persists as base levels are lowered, the passage will eventually grow to a size where the available recharge will not fill it and a free air surface will develop. If the recharge source is not diverted, there will be a transition from a tubular conduit to a canyon morphology and the canyon will continue deepening as base level continues to lower. If the recharge source is diverted to some lower route with lowering base levels, the conduit will be drained but will persist as an abandoned dry tube. Those tubes and canyons, which by chance circumstances develop entrances, are the population of explorable caves.

After the conduit has been drained, it enters a *stagnation* phase that may persist for a very long period of time. Eventually, however, erosional lowering will sweep the land surface past the cave passage. The conduit will be fragmented and ultimately destroyed. The later part of this sequence has been sketched separately as a final *decay* phase when cave passages are truncated by collapse and broken into short passage fragments, as many caves are observed to be, before they are destroyed completely.
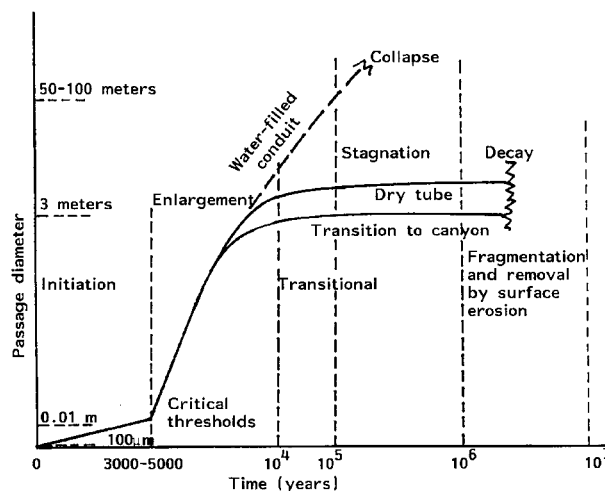
FIGURE 18.9 Schematic illustration of the "life history" of a conduit. The time scale is derived from geochemical modeling in the initiation and enlargement phases and from geologic and age-dating evidence in the stagnation and decay phases.

One of the objectives of cave development theories is to assign a time scale to the various phases in the life history of the conduit, particularly the early ones where solution of the bedrock is the dominant process. The time scale of Figure 18.9 is based on calculations for the early stages and on geomorphic and age-dating evidence for the later stages. The stagnation and decay phases are determined by events elsewhere in the drainage basin or by regional tectonics. It is not possible to calculate what happens from a general model although events in specific drainage basins can sometimes be interpreted.

A question not yet addressed by any model is whether, given the calculations and analysis for single passages, the results could be put together to obtain the equivalent "life history" for an entire conduit system and its associated groundwater basin. If one could, it would allow us to say something about the historical development of the drainage basin, and still further extrapolated, to say something about events in the regional climatic and tectonic setting in which the drainage basin is embedded. Overall, a formidable task, but an important one.

18.5.2 Initiation and Critical Thresholds

In compact and impermeable limestones, the initial hydraulic pathway from input to output is along a sequence of joints, fractures, and bedding plane partings with mechanical apertures in the range of 10 to 500 micrometers. Flow along the hydraulic pathway is laminar. The friction head is described by Equation (4) from which the hydraulic conductivity becomes

$$K = \frac{\rho g B^2}{12\eta} \tag{35}$$

where B is the full width (aperture) of the fracture. From Equation (35) it is possible to describe the discharge through a single fracture as the "cube law" of fracture flow.[27]

$$Q = \frac{C}{f} B^3 \, \Delta h \tag{36}$$

The volume of water flowing through the fracture and its velocity both increase rapidly as the fracture enlarges. When unsaturated water from the recharge source reaches the carbonate bedrock, much of the dissolutional capacity is consumed by reaction at the bedrock interface because of the rapid kinetics at high undersaturation. Water that enters the fracture system is close to saturation but remains slightly undersaturated because of the very slow rate of reaction in water close to equilibrium. In spite of the small fracture aperture and long travel times, slightly undersaturated water penetrates deep into the aquifer where the fractures are slowly enlarged. As the fractures enlarge, the volume of flow increases with the cube of the aperture, and the saturation index becomes more negative. The initiation phase ends when water at the critical saturation of $SI_c = -0.3$ penetrates the entire length of the hydraulic pathway.

Calculations based on a range of reasonable hydraulic gradients and carbon dioxide partial pressures gave a time scale for the initiation phase in the range of 5000 years,[28] a number that has been supported by more elaborate computer models.[29]

When the enlarging fracture reaches an aperture of 1 cm, three thresholds are crossed. These thresholds that mark the transition from fracture permeability to conduit permeability are:

1. The onset of turbulence
2. The penetration of the aquifer by undersaturated water
3. The transport of clastic sediments

For a fixed input and a fixed output defining the hydraulic gradient across an aquifer, the flow velocity within a growing fracture can be calculated. From the velocity and fracture aperture, the Reynolds number

can be calculated and tracked as both velocity and aperture increase as dissolution proceeds. For a range of hydraulic gradients from 10^{-4} to 10^{-2}, the onset of turbulence occurs at an aperture of 10 mm to 5 mm, respectively. The onset of turbulence marks the beginning of a regime where a linear relation between head and flow velocity can no longer be used to construct models of groundwater flow.

When one of the possible hydraulic pathways through the aquifer achieves an aperture that permits highly undersaturated water to penetrate the aquifer, the dissolution kinetics shift from the slow rate regime to the fast rate regime. Dissolution rates then increase rapidly so that the hydraulic pathway that first achieves breakthrough will undergo a runaway process and grow at the expense of alternate hydraulic pathways. It is this kinetic trigger that is responsible for the aquifer containing a small number of large conduits rather than an arrangement of small openings along all possible pathways.[3,29,30,31] For the kinetic trigger and associated runaway process to function, the aquifer must have sufficient hydraulic gradient to permit velocity to increase in response to increasing aperture. A failure of the kinetic trigger to function leads to the development of maze caves.

Carbonate rocks contain from a little to as much as 50% insoluble residue in the form of clay minerals, quartz grains, silicified fossil fragments, and chert. This material must be transported out of the aquifer by mechanical action of the flowing water. Soils and other weathered material is flushed down shafts and sinkhole drains as are the weathering products of the overlying caprock. Further adding to the underground sediment load are sinking streams which carry clastic material generated by weathering in the sinking stream basins. Mass balance arguments[32] show that the karst system must be capable of carrying the total clastic load if cave systems are not to choke up.

Sediment can be transported as bedload and as suspended load. Suspended load moves in karst aquifers mainly during flood flows. Bedload transport requires a finite boundary shear between the bedded sediment and the flowing water. The velocity to initiate sediment movement is on the order of 0.1 m/sec; this will occur, depending on hydraulic gradient, when the aperture is in the range of 5 to 10 mm.

18.5.3 Enlargement, Maturity, and Decay

Once the kinetic threshold has been passed, the conduit enlarges rapidly by fast kinetics. Calculations[30] indicate that the enlargement is uniform along the entire conduit. The retreat of the walls was calculated to be in the range of 400 mm/1000 years, a rate that would produce a conduit 4 m in radius in 10,000 years. These results represent an upper limit. Other factors such as walls coated with the insoluble residue from the limestone or with mud from the surface would slow down the dissolution process. Generally, the rapid enlargement converts the initial fracture geometry into a more nearly tubular cross section. The rate of enlargement depends on the path length and the hydraulic gradient. Under the high gradients and short path lengths that would be found in such locations as dam abutments, significant dissolutional enlargement could take place within the design lifetime of the structure.[33]

Development of conduit permeability in karst aquifers takes place on very short time scales. Continuous conduit development at successively lower elevations can easily keep pace with downcutting of surface river valleys. Geological evidence that the theoretical predictions are reasonable is provided by caves that are known to have developed in the 20,000 years since the retreat of the last Pleistocene ice sheet and also by caves in the Bahamas that have developed in limestones that themselves are only 85,000 years old.[34]

In the eastern U.S. many conduits are graded to local base levels. The conduits that can be readily explored are in a mature stage, often occupied by free-surface streams and subject to flooding. Many conduits have alternating reaches of air-filled passage (at least during base flow) and permanently water-filled passage. The streams in these conduits and the springs where the water returns to the surface are usually undersaturated with respect to calcite, with saturation indices typically in the range of –0.5 to –0.3, exactly what is expected for continuous enlargement by dissolution in the fast kinetics mode.

Mature conduits lying close to regional base levels act as low gradient master drains for the matrix and fracture portion of the aquifer. Conduits may be blocked by rockfall which ponds the water, which in turn acts as a sediment trap. Bypass routes and flood spillover routes evolve. Segments of the conduit

that lie below regional base level remain water-filled and continue to enlarge. This process continues until the conduit becomes mechanically unstable and collapses.

Continually lowering base levels shift hydraulic gradients. The new gradients initiate new conduit development at lower elevations, and eventually the recharge sources are diverted into the new conduits. The original conduit may survive as a dry cave passage for time periods of several million years based on the geomorphological interpretation of eastern U.S. landscapes. Shafts and chimneys carrying vadose water from overlying perched aquifers or from closed depressions sometimes intersect the older conduit where the water may be seen as waterfalls or segments of very high gradient streams as the water, moving nearly vertically in the vadose zone, passes the abandoned conduit.

18.6 Quantitative Hydrology of Karst Aquifers

18.6.1 Water Budgets

The concept of the groundwater basin allows the water budget to be used as a powerful tool in the evaluation of karst aquifers. The overall balance equation (Equation [1]) balances inputs to the aquifers against discharge. Spring discharge and allogenic recharge can be gauged. Diffuse infiltration and internal runoff must be estimated from catchment areas corrected for evapotranspiration losses. High internal runoff tends to lower evapotranspiration losses because of the rapid transport of water underground.

The divides for groundwater basins must be drawn from a consideration of incomplete evidence:

1. Results of tracer tests linking sinking streams, cave streams, and other accessible inputs to the discharging spring.
2. Water table slopes determined from well inventories.
3. Geological constraints imposed by folds and faults, regional dips, and fracture patterns.

The groundwater basin as finally drawn must be consistent with the water budget. A seriously unbalanced budget is an indication that the boundaries of the groundwater basin have not been drawn correctly.

Water budgets should be constructed under high and low flow regimes. Spillover routes that transport water across basin boundaries are common in the small groundwater basins.

18.6.2 Base Flow/Runoff Relations

Base flow in surface streams is maintained by groundwater discharge. If the region has a distinct dry season, the mean base flow for the surface streams gives a measure of regional base flow. This is normalized to basin area to yield a unit base flow:

$$Q_N = \overline{Q_B}/A \qquad (37)$$

Springs fed by conduits behave much like surface streams and exhibit a low storage and therefore a low base flow. For those aquifers for which the spring discharge reaches base flow conditions, Equation (37) can be used to estimate the area of the groundwater basin.[35] Tests of the relationship by Quinlan and Ray[36] suggest that the method is reasonably accurate provided that it is calibrated for the climate and geologic setting of the aquifer of interest.

18.6.3 Tracer Studies

Much of the advance in understanding karst aquifers in the past several decades has come from tracer studies.[37-39] An identifiable substance is placed in a sinking stream, cave stream, well, or flushed into a closed depression. The substance is later detected at an outlet point, usually a spring but sometimes an underground stream or another well. Tracers that have been used include:

- Fluorescent dyes
- Spores
- Rare elements (e.g., bromine)
- Noble gases
- Other substances

Although tracing using colored lycopodium spores received a great deal of attention in Europe, it has not proved useful in the conduit aquifers of the eastern U.S. where springs are frequently muddy and the plankton nets used to capture the spores easily become clogged. Traces using NaBr combined with neutron activation analysis to detect low levels of Br⁻ have been used but are expensive because of the analytical methods. Traces with other substances such as a massive injection of salt would now be unacceptable from environmental considerations.

Injections of noble gases are useful for tracing fracture flows. A noble gas, typically helium or neon, is bled continuously into the groundwater through a well drilled along the fracture. Water collected from another well or spring is then analyzed for the noble gas by mass spectrometry. Unlike other tracers which use a single spike of the tracer material, noble gas traces use a continuous injection over a long period of time. It is thus suitable for traces through tight fractures where long travel times are expected.

By far the most important tracers for characterizing karst aquifers are fluorescent dyes. The technique is the same as with other tracers tests. Dye is injected into the aquifer at points of interest and detected at springs or other discharge points. With increasing levels of sophistication these tests are:

1. Visual observation of dye
2. Visual test with dye receptor
3. Fluorescence detection with dye receptor
4. Quantitative test with automatic sampler
5. Quantitative test with on-site spectrofluorophotometer

Visual tests require the injection of sufficient dye to visibly color the water at the discharge point. Other than the inconvenience of positioning observers and the possibility of dye coming through at night, visual traces require such large quantities of dye that the test would be a source of groundwater contamination. Direct visual observation has been used historically but is now rarely used.

A major advance made in the late 1950s[40] was the introduction of packets of activated charcoal as a dye receptor. A few grams of high-grade coconut charcoal contained in mesh bags are placed in all suspected resurgent points. Dye is strongly adsorbed onto the charcoal and is not desorbed by continuing flow of water. The packets are collected at convenient intervals and replaced by fresh packets. The dye is elutriated with an alcoholic solution of strong base. The detection of dye by visual observation of the elutriate can be improved by observing the fluorescence of the solution under ultraviolet light. Use of the charcoal receptors permits many possible discharge points to be tested at the same time.

Optical brighteners can also be used for qualitative traces. These are agents widely used in laundry detergents so background should be checked. The receptors are wads of unsized cotton onto which the brightening agent strongly adsorbs. The presence of the brightener is determined by observing its bright blue luminescence under long wave ultraviolet radiation.

A more quantitative trace can be made by installing an automatic water sampler at the suspected discharge point. These can be programmed to collect water samples at prescribed intervals. The water samples are analyzed by spectrofluorimetry. Since the analysis is of the water itself rather than an elutriate, the analysis can be calibrated to determine fraction of dye recovered. A plot of dye concentration as a function of time gives a breakthrough curve (Figure 18.10) that shows the exact travel time. The shape of the dye pulse often gives other clues to the nature of the flow path.

To carry out tests on more than one injection point requires the use of multiple dyes. The fluorescence bands are broad and overlap so that the identification of mixed dyes is difficult by visual observation alone. However, fluorescence spectra can be deconvoluted with suitable software, allowing quantitative determination of the different dye combinations.
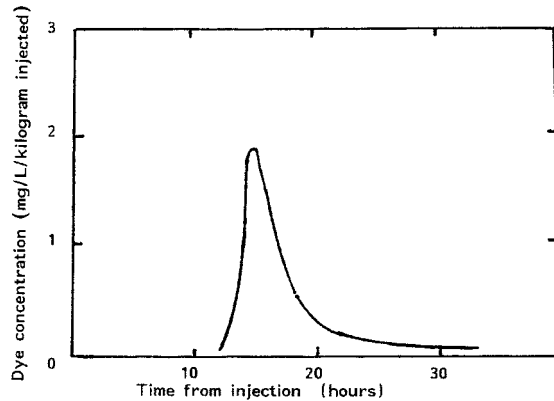
FIGURE 18.10 A typical dye breakthrough curve. (Adapted from Mull, D. S., Liebermann, T. D., Smoot, J. L., and Woosley, L. H., Jr., Application of dye-tracing techniques for determining solute-transport characteristics of ground-water in karst terranes, *U.S. Environmental Protection Agency Report,* EPA 904/6-88-001, 1988.)

A recent development is the use of a fiber optic probe and portable spectrofluorophotometer directly at the detection site. The fluorescence signal is recorded on a data logger and transferred to a computer. This technique permits a much greater density of measurement points and eliminates the need for collecting large numbers of samples.

The requirements for tracers is that they be nontoxic, stable in the groundwater environment, non-adsorbent on clays, silts, and the carbonate wall rock, and detectable at very low concentrations. Tracers that are most widely used are listed in Table 18.4. Details on the chemistry, properties, and toxicity of these and other tracer dyes may be found in Smart and Laidlaw,[41] and Smart.[42] Use of fluorescein as a water tracer has a 100-year history and remains the most popular tracer. It decomposes in sunlight and so does not persist after emerging from the groundwater system. The main drawback of fluorescein is that it is used in many common products such as antifreeze and toilet bowl cleaners, which often produce an unacceptable background.

TABLE 18.4 Commonly Used Tracer Dyes for Karst Hydrology

Common Name	Color Index	Fluorescence Wavelength	
		Elutriate	Water
Sodium fluorescein	Acid yellow 73	515.5	508
Eosine	Acid red 87	542	535
Rhodamine WT	Acid red 388	568.5	576
Sulphorhodamine B	Acid red 52	576.5	585
Optical brightener	Tinopal CBS-X	398.0	397

Data courtesy of Crawford & Associates, Bowling Green, Kentucky.

If there are many potential discharge points to be tested, a qualitative investigation with dye receptors can be used to delineate the overall drainage pattern. These results can be followed up with quantitative studies to obtain travel times and flow characterization. When legal and regulatory issues are involved, the quantitative trace provides an objective record to be entered into evidence.

Tracer testing is by far the most powerful technique available for working out the conduit system of a carbonate aquifer. Combined with direct exploration and survey of such parts of the active conduit as may be available, it should be possible to delineate the groundwater basin, determine whether it is conservative or nonconservative, specify allogenic recharge and base level discharge points, and set down at least a rough plan of the conduit system.

18.6.4 Well Tests and Water Table Mapping

The literature is ambiguous concerning the properties of the karst water table and even its very existence. Many examples are known where air-filled cave passages lie below flowing streams. Certain alpine caves ingest snow melt and runoff as turbulent frothing streams and crashing waterfalls. These have been cited as evidence for the absence of a water table in karst aquifers. Much of the confusion appears to arise from a failure to appreciate the dynamic response of the conduit system to storm runoff and failure to appreciate the mismatch in hydraulic conductivity between the conduit system and the associated fracture and pore system.

Because of their high flow capacity, conduits cannot support much hydraulic head and thus create a trough in the groundwater system during periods between storms. Because of their rapid response to recharge, conduits quickly fill during storm flow and the groundwater trough may become a groundwater mound. Loosing surface streams perched over conduits that are air-filled at base flow may become gaining streams when the conduit is flooded. If the hydraulic conductivity of the fracture and pore system is low, the conduit system, particularly one with a large component of allogenic recharge, may act as a decoupled set of pipes with little influence on the water table.

The fracture and matrix permeability portion of the aquifer supports a well-defined water table which can be mapped if a sufficient number of wells is available. Water tables that have been mapped in karst areas appear reasonable and the location of conduits is often revealed by the groundwater troughs.

Drilling test wells is an appropriate way to probe the fracture and matrix component of the aquifer. However, account must be taken of the heterogeneous distribution of fracture permeability on the borehole scale. Wells drilled on fracture intersections may produce several orders of magnitude more water than wells drilled in the same formation in the rock between the fractures. When the bedrock is not masked by thick soils or overlying caprock, and where the permeability is dominated by fractures rather than by bedding plane partings, air photographs can be used to map fracture traces at the land surface.[43] These are then used to select drilling sites.

Standard pump tests can be used to evaluate aquifer yield providing that the pumped well has not tapped a conduit (which would give too much yield with too little drawdown) and providing that it can be shown that the well has tapped the fracture system (drilling into unfractured bedrock would give yields that are too small).

Downhole video cameras are a useful tool for locating fracture zones and directly examining solution features intersected by wells. If the well can be pumped dry, the camera can be used to observe which of the solution openings is actually discharging water into the well. Packers can then be used to isolate productive zones and these can be pump tested or injection tested independently of the remainder of the aquifer.

18.6.5 Hydrographs and Chemographs of Karst Springs

While the exploration and tracer studies reveal something of the conduit system and wells and pump tests characterize the fracture and matrix systems, the only hydraulic response that depends on both systems is that of the spring that drains the aquifer. Spring hydrographs and fluctuations in spring water chemistry provide very useful guides to aquifer hydraulics.

Spring hydrographs can be grouped into three types (Figure 18.11) depending on the response time of the aquifer compared with the mean spacing between storms. In fast response aquifers (Type I), a storm pulse can pass completely through the aquifer and conditions return to base flow before the next storm pulse arrives. Each storm produces its own hydrograph at the spring which is not different in overall characteristics from hydrographs of surface streams. In slow response aquifers (Type III) the throughput time is sufficiently long to completely flatten the individual hydrographs and all that remains is a broad rise and fall relating to wet and dry seasons. In the limit, the slow response aquifer may produce only a horizontal straight line, a spring of constant flow. When the response time is comparable to the mean spacing between storms (Type II), some structure appears in the hydrograph but the individual storm pulses are not resolved.
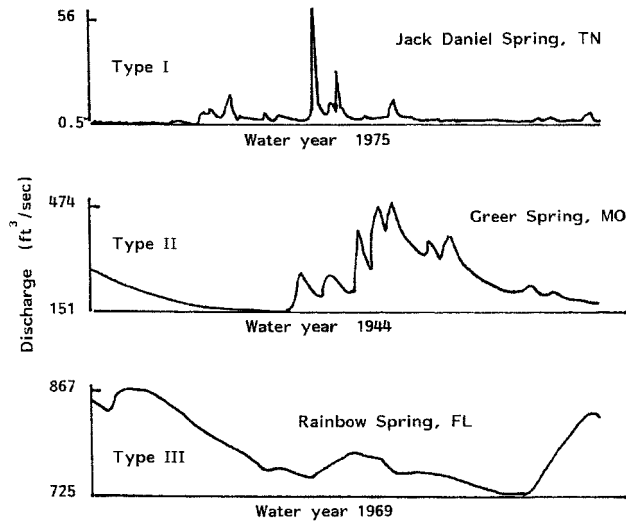
FIGURE 18.11 Types of spring hydrographs. Maximum and minimum discharges are given numerically in cubic feet/second. (Data from U.S. Geological Survey surface water records.)

Factors that determine the type of spring hydrograph include degree of conduit development, area of groundwater basin, proportion of allogenic recharge, and hydraulic gradient from input to output. Small basins with well-developed conduit systems are more flashy than larger basins with the same degree of conduit development. Longer travel times in large basins tend to smooth out the hydrograph even if there are open conduits. Very low gradients reduce travel times and produce the same effect.

A useful parameter for describing the flashiness of an aquifer is the ratio of the maximum annual discharge, the quantity known in surface water hydrology as the annual flood, to the base flow, averaged over the number of years of record.

$$Q_f = \frac{1}{n} \sum_{i=1}^{n} \frac{\left(Q_{\max}\right)_i}{\left(Q_{\text{base}}\right)_i} \tag{38}$$

Type I responses have Q_f in the range of 70 to 100, Type II in the range of 5 to 10, and Type III in the range of 1 to 2, based on relatively sparse data.

Individual storm hydrographs can be further analyzed (Figure 18.12). There is a lag between the arrival of a storm pulse at the swallet or the onset of rainfall and the rising limb of the hydrograph at the spring. This lag, however, does not measure the time required for storm water to traverse the aquifer. The injection of water into the upstream reaches of the aquifer raises hydraulic gradients, floods conduits, and forces out by piston flow water that has been stored deeper in the aquifer. The recession limb of the hydrograph is rarely the expected simple exponential function

$$Q = Q_0 e^{-\varepsilon t} \tag{39}$$

$$\varepsilon = \frac{1}{t_R}$$

Very characteristically, the recession limb will consist of a fast response segment and a slow response segment (Figure 18.13). Sometimes there are more than two distinct segments. The fast response segment can be assigned to the conduit portion of the aquifer which drains rapidly when the storm input ceases.

The figure shows a storm hydrograph with Discharge Q on the vertical axis and Time on the horizontal axis. Labels include Rainfall, Q_{max}, Rising limb, Lag time, Recession limb, Base flow, Q_B, and $t = 0$. Equations shown: $\dfrac{d^2 Q}{d t^2} = 0$ and $Q = Q_0 e^{-t/t_R}$.
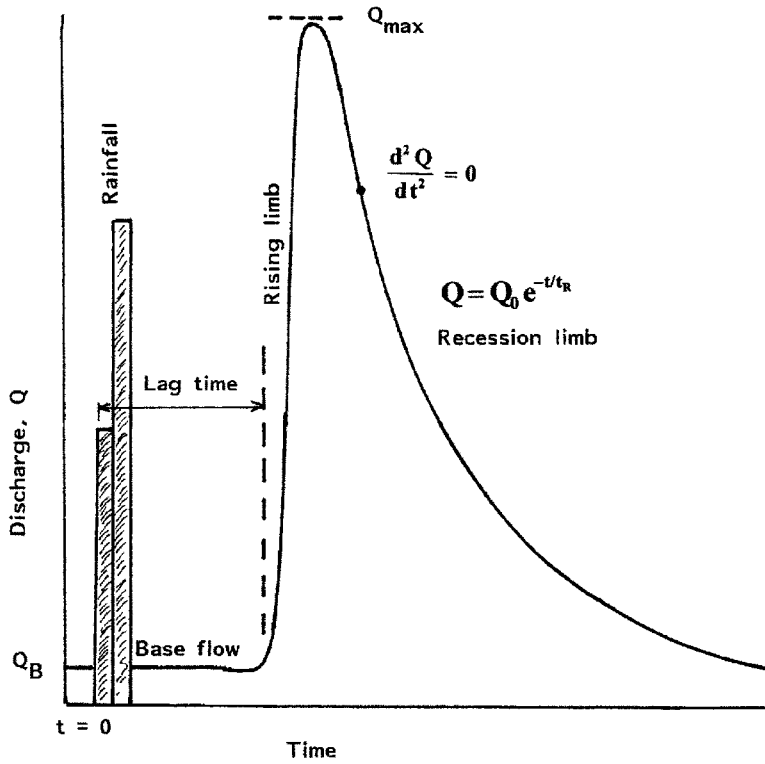
FIGURE 18.12 Sketch showing components of a storm hydrograph at a karst spring.

The semi-log plot is labeled Big Spring, MO, with $\ell n\, Q$ (ft^3/sec) on the vertical axis and Time (days) on the horizontal axis. Two line segments are labeled $t_R = 9$ days and $t_R = 25$ days.
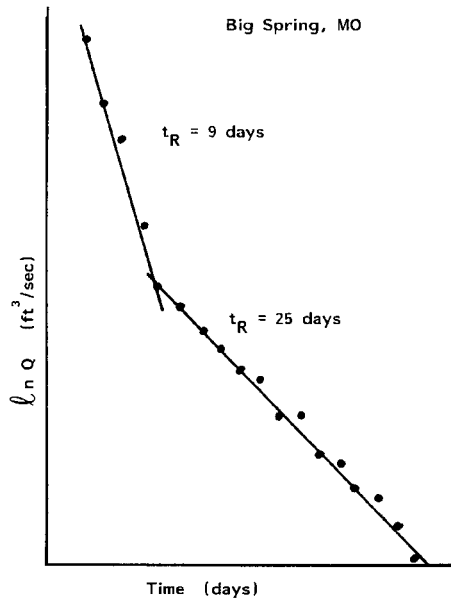
FIGURE 18.13 Semi-log plot of hydrograph recession curve showing fast response and slow response components. (Data from U.S. Geological Survey surface water records.)

The slow response segment can be assigned to the fracture system draining into the conduit after the conduit itself has emptied.

Analysis of the recession limb of the spring hydrograph gives some information on the relative importance of the conduit and fracture systems and the coupling between them. The quantity, ε, in Equation (38) is called the exhaustion coefficient.[44] The volume of water in dynamic storage above regional base level is given as

$$V = \frac{86,400Q_{max}}{\varepsilon} \tag{40}$$

where V is volume in m^3, 86,400 is the number of seconds in a day, Q_{max} is peak discharge in $m^3 sec^{-1}$, and ε is given in units of days. The exhaustion coefficient is also related to the transmissivity and storativity through

$$\varepsilon = \frac{2T}{SL^2} \tag{41}$$

where L is the horizontal distance from recharge point to discharge point.[1] The storage derived from the recession curve is what may be termed "dynamic storage." It is the water stored above the local base level defined by the elevation of the spring. The deep storage below local base level is not probed by these measurements and requires wells of sufficient depth to penetrate the bottom of the aquifer.

It has long been known that the chemical composition, especially the hardness (Equation [23]) of some springs is nearly constant regardless of wet and dry seasons and episodes of storm activity. Other spring waters have a highly variable chemistry. The original interpretation[45] used the coefficient of variation (standard deviation/mean) of hardness data as a quantitative expression of the variation. A coefficient of variation (CV) < 3% indicated a spring draining a fracture aquifer (termed a "diffuse flow" aquifer) and a CV > 20% indicated a spring draining an aquifer containing open conduits (termed a "conduit flow" aquifer). It was soon realized that the variability in spring water chemistry represented a continuum rather than a binary distribution in two categories. The coefficient of variation could be correlated with throughput time[46] and with fraction of groundwater supplied by allogenic recharge.[47] Both throughput time and allogenic recharge are related to the development of conduit permeability.

More information may be obtained from a continuous record of chemical composition. The result is called a chemograph (Figure 18.14) in analogy with a hydrograph. Chemographs may be obtained by obtaining a very large number of chemical analyses collected on a sampling interval short with respect to events on the chemograph. More practical is the continuous monitoring of specific conductance which is proportional to hardness for most groundwater basins. Chemographs record events on many time scales. The chemograph of CO_2 pressure from a large basin reveals the seasonal oscillation due to the coming and going of the growing season.[48] The concentration of Ca^{2+} ions measured on the Maramec Spring, Missouri,[49] exhibits dips corresponding to individual storm events. Like the recession limb of the hydrograph, the chemograph has an exponential recovery tail. However, the recovery time for the chemistry is substantially longer than the hydrograph recession. There are also details within individual storm chemographs with a time resolution as short as 15 minutes.[50] This fine structure has been interpreted as the arrival of storm water from different tributaries of the conduit system.

18.6.6 Modeling of Karst Aquifers

Thus far it is probably fair to say that a completely adequate model for karst aquifers has not been developed. Aquifers with permeability consisting mainly of solutionally widened fractures can be treated with continuum models such as the popular MODFLOW program. The necessary assumption is that the fractures are sufficiently interconnected and closely spaced to justify being treated as a continuum with an average hydraulic conductivity on a regional scale.
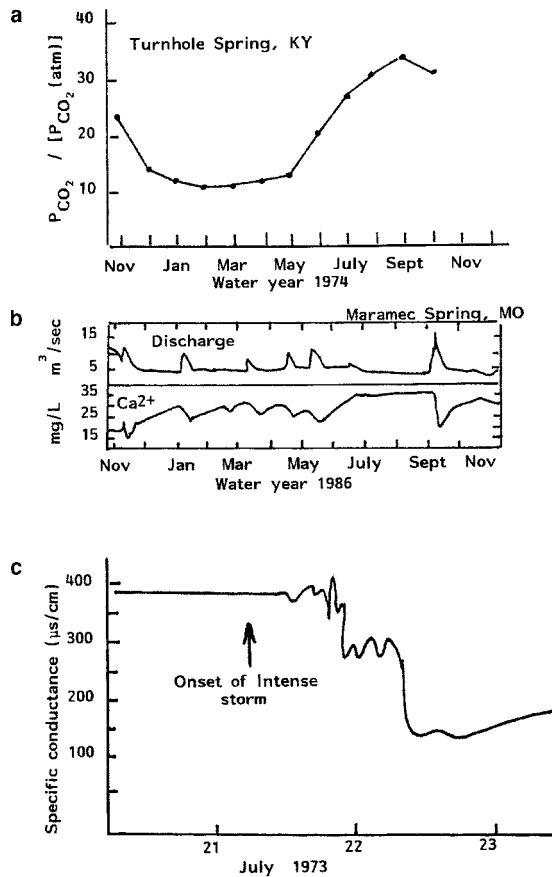
FIGURE 18.14 Chemographs for selected karst springs. (a) CO_2 partial pressure normalized to atmospheric CO_2, Turnhole Spring, KY. (Data from Hess, J. W. and White, W. B., Groundwater geochemistry of the carbonate karst aquifer, Southcentral Kentucky, U.S.A., *Applied Geochemistry*, 8, 189, 1993.) (b) Relation of Ca^{2+} ion concentration to discharge, Maramec Spring, MO. (Adapted from Dreiss, S. J., Regional scale transport in a karst aquifer 1. Component separation of spring flow hydrographs, *Water Resour. Res.*, 25, 117, 1989.) (c) Response of specific conductance to a single intense storm, Turnhole Spring, KY. (Adapted from Hess, J. W. and White, W. B., Storm response of the karstic carbonate aquifer of Southcentral Kentucky, *J. Hydrol.*, 99, 235, 1986.)

Two approaches have been used for aquifers with a well-developed conduit system.[51,52] Both assume a distinct fracture permeability system and a conduit system as suggested by storm recession hydrographs (Figure 18.13). In both models the fracture system is described by an average hydraulic conductivity assumed to hold on a regional scale. In the discrete model the conduit is taken as pipe flow described by the Darcy–Weisbach equation (Equation [7]). In the double continuum model, flow in the conduit system is also treated as Darcian but with a higher value of hydraulic conductivity than the fracture system. In both models it is necessary to define an exchange term that describes the movement of water between the conduit and the fracture system.

Both models reproduce the flashy Type I hydrograph of the sort illustrated in Figure 18.11. In both models the specific arrangement of conduits must be put in "by hand" after field investigation of the specific groundwater basin of interest. An assessment of the status of groundwater modeling of karst aquifers is that great progress has been made but much more progress is necessary before a versatile and generally accepted model is achieved.

18.7 Water Resource Issues in Karst Hydrology

18.7.1 Water Supply in Karst

Water supplies in karst are obtained from springs and wells. Karst springs draining fracture systems, as indicated by uniform discharge, lack of turbidity, and small chemical variation, provide reliable water supplies generally of good quality except for high hardness. Springs draining conduit systems have a highly variable discharge, sometimes become turbid to muddy after storms, and are at risk from pollutants introduced through sinking streams and sinkholes. The conduit systems offer very little filtration. Injected pollutants can be quickly transmitted to the spring with little purification except for dilution.

Location of production wells in karst is challenging. Fractured limestone and dolomite aquifers provide the most reliable supplies of good quality water if the wells are drilled on fracture intersections. In favorable localities these can be determined by mapping lineaments on air photographs. Location of wells in aquifers dominated by conduit permeability is difficult because conduits cannot generally be located from the surface. If the conduit is an accessible cave, wells can be located by use of a magnetic induction device from within the cave. Wells drilled into conduits face the same uncertainty of discharge, turbidity, and pollution risk as conduit springs. Farm wells are sometimes placed in conduits; municipal water supply wells are usually not placed in conduits.

18.7.2 Flood Flows in Karst: Sinkhole Flooding

Karst aquifers serve to modify the flood flows in the associated surface water basins. Flood pulses of intermediate magnitude in tributary streams are injected into the conduit system at swallow holes. The conduit system holds the water in temporary storage and bleeds it out more slowly through the spring. The response time of even the most flashy conduit system is usually slower than the response time of flood pulses in surface channels. This has the effect of clipping the top of the hydrograph and spreading it out, thus reducing the flood crest downstream.[53]

Extreme floods may override swallow holes and send flood water down dry valleys where there is no stream channel. In developed areas, buildings constructed in dry swales and valleys may be damaged by the unexpected invasion of flood waters.

Sinkhole terrain is prone to flooding. One mechanism is that storm water ponds in closed depressions if the drain is plugged with clay. As a second and more important mechanism, the groundwater trough created by the conduit system is subject to rapid filling with the addition of an excess pressure head. As a result, water can be forced up the sinkhole drains. The resulting flooded sinkhole is filled by water injected from below. Sinkhole terrain is not usually thought of as flood prone and thus may be subject to development. Homes and businesses suffer damage when the sinkholes fill following exceptional storms.[54]

18.7.3 Pollution and Pollution Transport in Karst Aquifers

Karst aquifers are highly vulnerable to groundwater contamination because of the negligible filtering of the internal runoff and allogenic recharge components. Pollutants can travel long distances through the conduit system. In broad categories, contaminants include:

1. Barnyard waste and sewage
2. Agricultural chemicals
3. Solid waste from sinkhole dumps and improperly constructed landfills
4. Liquid waste from contaminated sinking streams, injection wells, sinkhole disposals, outfalls, and leaky storage tanks

Fecal material from septic tanks, leaky sewer lines, barnyards, and manure spread on fields migrates downward through thin soils and through the fracture system below. Cave streams and conduit springs

frequently contain colliform bacteria and high nitrate levels. Agricultural chemicals, insecticides, and herbicides have been reported in karst groundwater in many farming areas of the U.S.

Sinkholes have long been used for solid waste disposal in rural areas. Household waste, junk, paint and chemical containers, and dead, often diseased, farm animals end up in sinkholes. The waste materials are leached by the internal runoff draining into the sinkholes and the leachate is carried into the underlying aquifer. Solid waste can be flushed down open drains and dispersed along the conduit system where it is impossible to clean up. Contaminants can be incorporated into the clastic sediments and flushed by flood flows deep into the aquifer where they bleed out leachate over long periods of time.

Liquid wastes from improper disposal of solvents, from industrial outfalls directed into fractures and sinkholes, and from leaking pipelines and storage tanks can make their way into karst aquifers through any of the recharge routes. Soluble wastes dissolve in the karst groundwater and are carried along with it. Insoluble liquids with a density less than water (LNAPLs) float on underground streams or on the karst water table. The most common LNAPLs are gasoline and fuel oil from leaking underground storage tanks.[55] Toxic substances — the BTEX components benzene, toluene, ethyl benzene, and xylene — are found in LNAPL hydrocarbons. Fumes from these and other contaminants can be forced upward by rising water levels and intrude into basements of homes, schools, and office buildings. Flammable liquids ponded on cave streams pose an explosion hazard, especially when the conduits are located under urban areas.[56]

Dense nonaqueous phase liquids (DNAPLs) consist mainly of chlorinated hydrocarbons used as solvents and degreasers. Such materials would sink to the bottom of the cave stream and water-filled conduits. Their movement through the karst aquifer is poorly known. To the extent that they lie in pools in low places in the conduits, they act as a continuous source of contamination to the water in the conduit. During flood flow, however, the clastic sediments and any incorporated DNAPLs are swept down the conduit.

18.7.4 Water Quality Monitoring in Karst

U.S. Environmental Protection Agency and state regulations require legally established landfills and other disposal sites to install a network of monitoring wells. Samples are drawn from these wells at prescribed intervals and analyzed for contaminants. A typical requirement is three down-gradient wells and one up-gradient well to be installed on the perimeter of the waste disposal site. The strategy is to detect any contamination before it moves offsite and to take early remedial action.

The concept does not work in karst. Because of the heterogeneous distribution of conduit permeability and the general impossibility of locating conduits in the subsurface, there is a very real possibility of contaminants moving offsite through the karst conduit while leaving no trace in the monitoring wells.

To monitor waste sites in karst requires a different strategy. Trenches should be excavated to bedrock and tracers injected into the bedrock fractures with a water flush. The spring or springs to which the aquifer drains must be identified, and the spring must be used as a monitoring site regardless of the difficulties of offsite monitoring. Likewise, more frequent monitoring, especially monitoring following storm events is necessary because of the rapid transport of contaminants through the conduit system.[57]

18.7.5 Sediment Transport and Land Instability

During periods of high flow, velocities in conduit systems reach the thresholds necessary to transport clastic sediment. This transport allows piping processes in vertical fractures, shafts, and closed depression drains. Continuous piping of soil into the subsurface creates a void space that migrates upward toward the surface. The final step in this process is the abrupt collapse of the soil void with the creation of a soil piping sinkhole. These can range in size from a few meters to tens and even hundreds of meters. Soil piping sinkholes are one of the most important land use hazards in karst. For a comprehensive discussion of the engineering implications of these indirect products of karst hydrology see Sowers' review.[58]

References

1. Milanovic, P. T., *Karst Hydrogeology*, Water Resources Publications, Littleton, CO, 1981.
2. White, W. B., *Geomorphology and Hydrology of Karst Terrains*, Oxford University Press, New York, 1988.
3. Dreybrodt, W., *Processes in Karst Systems*, Springer Verlag, Berlin, 1988.
4. Ford, D. C. and Williams, P. W., *Karst Geomorphology and Hydrology*, Unwin Hyman, London, 1989.
5. Williams, P. W., The role of the subcutaneous zone in karst hydrology, *J. Hydrol.*, 61, 45, 1983.
6. Williams, P. W., Subcutaneous hydrology and the development of doline and cockpit karst, *Zeitschrift für Geomorphologie*, 29, 463, 1985.
7. Brucker, R. W., Hess, J. W., and White, W. B., Role of vertical shafts in movement of groundwater in carbonate aquifers, *Ground Water,* 10 [6], 5, 1972.
8. Hess, J. W. and White, W. B., Water budget and physical hydrology, in *Karst Hydrology: Concepts from the Mammoth Cave Area*, White, W. B. and White, E. L., Eds., Van Nostrand Reinhold, New York, 1989, Chap. 4.
9. Sauter, M., Assessment of hydraulic conductivity in a karst aquifer at local and regional scale, *Proceedings of the Third Conference on Hydrogeology, Ecology, Monitoring, and Management of Ground Water in Karst Terranes*, Nashville, TN, 39, 1991.
10. Vennard, J. K., *Elementary Fluid Mechanics*, Wiley, New York, 1962.
11. Barnes, H. H., Jr., *Roughness Characteristics of Natural Channels*, U.S. Geological Survey Water Supply Paper 1849, Washington, D.C., 1967.
12. Palmer, A. N., Origin and morphology of limestone caves, *Geological Society of America Bulletin*, 103, 1, 1991.
13. Palmer, A. N., Origin of maze caves, *The NSS Bulletin*, 37, 57, 1975.
14. Schmidt, V. A., Magnetostratigraphy of sediments in Mammoth Cave, Kentucky, *Science*, 217, 827, 1982.
15. Curl, R. L., Deducing flow velocity in cave conduits from scallops, *National Speleological Society Bulletin*, 36, 1, 1974.
16. Gale, S. J., The hydraulics of conduit flow in carbonate aquifers, *J. Hydrol.*, 70, 309, 1984.
17. Lauritzen, S.-E., Abbotto, J., Arnesen, R., Crossley, G., Grepperud, D., Ive, A., and Johnson, S., Morphology and hydraulics of an active phreatic conduit, *Cave Science*, 12, 139, 1985.
18. White, W. B. and Deike, G. H., Hydraulic geometry of cave passages, in *Karst Hydrology: Concepts from the Mammoth Cave Area*, White, W. B. and White, E. L., Eds., Van Nostrand Reinhold, New York, 1989, Chap. 9.
19. Drever, J. I., *The Geochemistry of Natural Waters*, Prentice Hall, Englewood Cliffs, NJ, 3rd Edition, 1997.
20. Langmuir, D., *Aqueous Environmental Geochemistry*, Prentice Hall, Upper Saddle River, NJ, 1997.
21. Plummer, L. N. and Busenberg, E., The solubilities of calcite, aragonite, and vaterite in CO_2–H_2O solutions between 0 and 90°C and an evaluation of the aqueous model for the system $CaCO_3$–CO_2–H_2O, *Geochimica et Cosmochimica Acta*, 46, 1011, 1982.
22. Melchior, D. C. and Bassett, R. L., *Chemical Modeling of Aqueous Systems II*, American Chemical Society Symposium Series 416, 1990.
23. Plummer, L. N., Wigley, T. M. L., and Parkhurst, D. L., The kinetics of calcite dissolution in CO_2-water systems at 5° to 60°C and 0.0 to 1.0 atm CO_2, *Am. J. Sci.*, 278, 179, 1978.
24. Plummer, L. N., Parkhurst, D. L., and Wigley, T. M. L., Critical review of the kinetics of calcite dissolution and precipitation, in *Chemical Modeling in Aqueous Systems*, Jenne, E. A., Ed., American Chemical Society Symposium Series 93, Washington, D.C., 1979, Chap. 25.
25. Busenberg, E. and Plummer, L. N., The kinetics of dissolution of dolomite in CO_2–H_2O systems at 1.5 to 65°C and 0 to 1 atm PCO_2, *Am. J. Sci.*, 282, 45, 1982.
26. Herman, J. S. and White, W. B., Dissolution kinetics of dolomite: Effects of lithology and fluid flow velocity, *Geochimica et Cosmochimica Acta*, 49, 2017, 1985.

27. Witherspoon, P. A., Wang, J. S. Y., Iwai, K., and Gale, J. E., Validity of cubic law for fluid flow in a deformable rock fracture, *Water Resour. Res.*, 16, 1016, 1980.

28. White, W. B., Role of solution kinetics in the development of karst aquifers, in *Karst Hydrogeology*, Tolson, J. S. and Doyle, F. L., Eds., International Association of Hydrogeologists Memoir 12, 503, 1977.

29. Groves, C. G. and Howard, A. D., Early development of karst systems. I. Preferential flow path enlargement under laminar flow, *Water Resour. Res.*, 30, 2837, 1994.

30. Groves, C. G. and Howard, A. D., Minimum hydrochemical conditions allowing limestone cave development, *Water Resour. Res.*, 30, 607, 1994.

31. Dreybrodt, W., Principles of early development of karst conduits under natural and man-made conditions revealed by mathematical analysis of numerical models, *Water Resour. Res.*, 32, 2923, 1996.

32. White, E. L. and White, W. B., Dynamics of sediment transport in limestone caves, *National Speleological Society Bulletin*, 30, 115, 1968.

33. Dreybrodt, W., Dynamics of karstification: A model applied to hydraulic structures in karst terranes, *Applied Hydrogeology*, 3, 20, 1992.

34. Mylroie, J. E. and Carew, J. L., Field evidence of the minimum time for speleogenesis, *The NSS Bulletin*, 49, 67, 1987.

35. White, E. L., Sustained flow in small Appalachian watersheds underlain by carbonate rocks, *J. Hydrol.*, 32, 71, 1977.

36. Quinlan, J. F. and Ray, J. A., Normalized base-flow discharge of groundwater basins: A useful parameter for estimating recharge area of springs and for recognizing drainage anomalies in karst terranes, in *Karst Geohazards*, Beck, B. F., Ed., A. A. Balkema, Rotterdam, pp. 149, 1995.

37. Gospodaric, R. and Habic, P., *Underground Water Tracing*, Institute of Karst Research, Ljubljana, Slovenia, 1976.

38. Jones, W. K., Dye tracer tests in karst areas, *National Speleological Society Bulletin*, 46, 3, 1984.

39. Mull, D. S., Liebermann, T. D., Smoot, J. L., and Woosley, L. H., Jr., Application of dye-tracing techniques for determining solute-transport characteristics of groundwater in karst terranes, *U.S. Environmental Protection Agency Report,* EPA 904/6-88-001, 1988.

40. Haas, J. L., Evaluation of groundwater tracing methods used in speleology, *National Speleological Society Bulletin*, 21, 67, 1959.

41. Smart, P. L. and Laidlaw, I. M. S., An evaluation of some fluorescent dyes for water tracing, *Water Resour. Res.*, 13, 15, 1977.

42. Smart, P. L., A review of the toxicity of twelve fluorescent dyes used for water tracing, *National Speleological Society Bulletin*, 46, 21, 1984.

43. Lattman, L. H. and Parizek, R. R., Relationship between fracture traces and the occurrence of groundwater in carbonate rocks, *J. Hydrol.*, 2, 73, 1964.

44. Burdon, D. J. and Papakis, N., *Handbook of Karst Hydrogeology*, United Nations Special Fund, Institute for Geology and Subsurface Research, Athens, Greece, 1963.

45. Shuster, E. T. and White, W. B., Seasonal fluctuations in the chemistry of limestone springs: A possible means for characterizing carbonate aquifers, *J. Hydrol.*, 14, 93, 1971.

46. Ternan, J. L., Comments on the use of a calcium hardness variability index in the study of carbonate aquifers: With reference to the Central Pennines, England, *J. Hydrol.*, 16, 317, 1972.

47. Worthington, S. R. H., Davies, G. J., and Quinlan, J. F., Geochemistry of springs in temperate carbonate aquifers: Recharge type explains most of the variation, *Colloque d'Hydrologie en Pays Calcaire et en Milieu Fissure*, Nauchatel, Switzerland, 341, 1992.

48. Hess, J. W. and White, W. B., Groundwater geochemistry of the carbonate karst aquifer, Southcentral Kentucky, U.S.A., *Applied Geochemistry*, 8, 189, 1993.

49. Dreiss, S. J., Regional scale transport in a karst aquifer 1. Component separation of spring flow hydrographs, *Water Resour. Res.*, 25, 117, 1989.

50. Hess, J. W. and White, W. B., Storm response of the karstic carbonate aquifer of Southcentral Kentucky, *J. Hydrol.*, 99, 235, 1986.
51. Teutsch, G. and Sauter, M., Groundwater modeling in karst terranes: Scale effects, data acquisition, and field validation, *Proceedings of Third Conference on Hydrogeology, Ecology, Monitoring, and Management of Ground Water in Karst Terranes*, Nashville, TN, 17, 1991.
52. Mohrlok, U. and Sauter, M., Modelling groundwater flow in a karst terrane using discrete and double-continuum approaches — Importance of spatial and temporal distribution of recharge, *Proceedings of the 12th International Congress of Speleology*, La Chaux-de-Fonds, Switzerland, 2, 167, 1997.
53. White, E. L. and Reich, B. M., Behavior of annual floods in limestone basins in Pennsylvania, *J. Hydrol.*, 10, 193, 1970.
54. Quinlan, J. F., Legal aspects of sinkhole development and flooding in karst terranes: 1. Review and synthesis, *Environmental Geology and Water Science*, 8, 41, 1986.
55. Ewers, R. O., Estes, E. K., Idstein, P. J., and Johnson, K. M., The transmission of light hydrocarbon contaminants in limestone (karst) aquifers, *Proceedings of the Third Conference on Hydrogeology, Ecology, Monitoring, and Management of Ground Water in Karst Terranes*, Nashville, TN, 287, 1991.
56. Stroud, F. B., Gilbert, J., Powell, G. W., Crawford, N. C., Rigatti, M. J., and Johnson, P. C., U. S. Environmental Protection Agency emergency response to toxic fumes and contaminated groundwater in karst typography, Bowling Green, Kentucky, *Proceedings of the Environmental Problems in Karst Terranes and Their Solutions Conference*, Bowling Green, KY, 197, 1986.
57. Quinlan, J. F., Special problems of ground-water monitoring in karst terranes, *Ground Water and Vadose Zone Monitoring*, ASTM STP 1053, Nielsen, D. M. and Johnson, A. I, Eds., American Society for Testing and Materials, Philadelphia, 175, 1990.
58. Sowers, G. F., *Building on Sinkholes*, American Society of Civil Engineers, New York, 1996.

19
Holistic Macroengineering Approach for Environmental Information Systems (EIS)

Basile Dendrou and
Stergios Dendrou
ZEi Engineering Inc.

19.1 Introduction

The first 18 chapters of this handbook provide detailed technical information on many facets of groundwater engineering. But information alone cannot be used productively: it takes knowledge and technology to render a new product useful, and therefore marketable. In this chapter, we apply principles of the information revolution to the problem of geoenvironmental engineering. In the early days of mainframe computers, the tool of choice was the numerical simulation model (many of the early codes are still in

use, e.g., MOC, Konikow et al., 1978). As is often the case, advances in computer technology find *ad hoc* implementation in existing technology; e.g., the pre-post-processor cohort of software in support of existing simulators. The quantum leap to the full advances and promise of the information highway require addressing the two pillars of this technology, *integration* and *automation*. Manifestations of this technology are the *object-oriented* software architecture and the freedom from the bondage of the file. In conventional engineering software the user must keep track of a plethora of files: input files, pre-processing files, output files, graphics files, document files. With the new technology, the user operates "under one roof" and can communicate with and access other software while dealing only with physical objects, wells, rivers, strata, log-points. Sophisticated simulation models are fully integrated and can be exercised from within the platform without recourse to input files. Likewise, all simulation results are automatically accessible in tabular and graphical form on screen, in editable documents, and on the web. This chapter gives the theoretical foundation and a blueprint for implementing this new technology. It also gives examples from the prototype software EIS (ZEi, 1994-1997).

19.1.1 Context of the Contamination Problem in Geomedia

Chemical contamination is a well-known environmental problem, whether caused by landfills and other waste sites, agricultural pesticides and herbicides, or by statistically inevitable accidents. Any successful remediation scheme must be based on a thorough understanding of the contaminant's fate and migration. The migration of contaminants is affected by many factors such as the movement of the water mass, the interaction of the bulk mass of the chemical with the water mass, and many chemical and biological processes bringing water, chemicals, and the soil matrix into interaction. The *Geomedia* represent a large system, as shown in Figure 19.1, which includes many components of different characteristics, such as the natural soil and rock formations, surface and subsurface waters, the contaminants in their multiphase form, and man-made structures (for containment or reinforcement).

Due to these inherent complexities of the geomedia and the corresponding natural processes, the existing simplified empirical modeling and simulation tools are no longer sufficient to design, regulate, and manage the contamination problem effectively. Furthermore, sustainable development requires a *holistic macroengineering* approach where interactions of different natural processes are an integral part of the theoretical model, which must be used to simulate actual contamination episodes, identifying
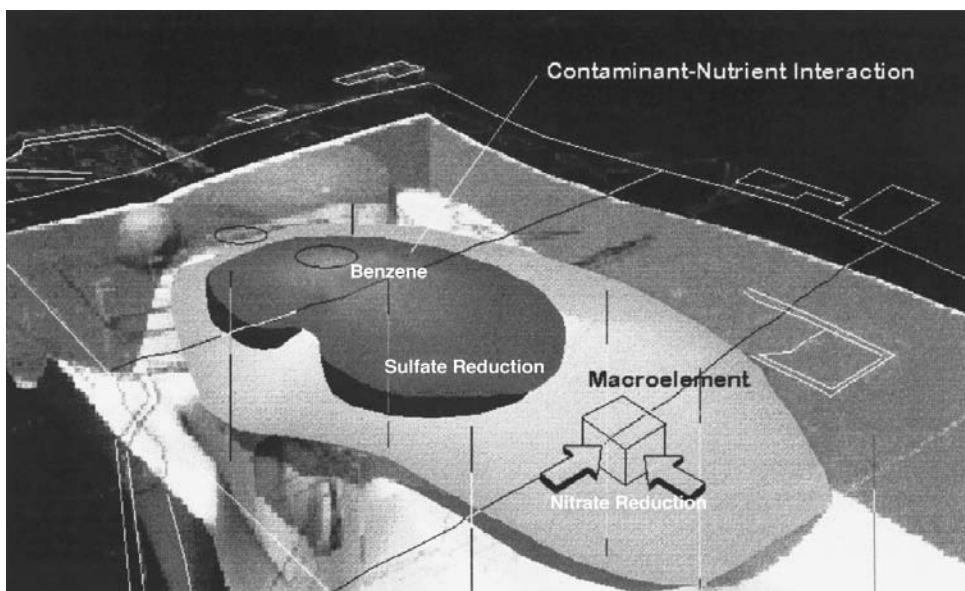
FIGURE 19.1 Geomedia characterization and scaling configurations.

optimum innovative and effective mitigative measures. Holistic means that one must look at the entire picture, not merely at components. And macroengineering (science of large systems) means to study the entire system through its component parts and their interactions at the appropriate geometric and time scales.

For example, it is no longer sufficient in a bioremediation scheme to simulate the contaminant migration using a simple exponential decay law without representing the chemical and biological interactions between the contaminant and the electron acceptors, the soil matrix and the available nutrients. The same can be said for other *in situ* remediation technologies such as bioventing (interaction of thermal and chemical processes), pump-and-treat (interaction of mechanical and chemical processes), vitrification (interaction of thermal, chemical, and mechanical processes), electrokinetics (interaction of electrical, chemical, and mechanical processes), chemical barriers, and *in situ* containment technologies (interaction of chemical and mechanical processes). Three-dimensional simulation models are needed capable of handling these processes, and of quantifying their impact on the geomedia environment. Simplified models on isolated natural processes help sort data for diagnostic purposes but lack the ability to make an accurate prognosis of the impact of the selected remedial alternative on the environment. In a sense, the absence of a reliable simulation model to quantify the interaction of all these different natural processes initiated during the implementation of a particular remedial technology is tantamount to carrying the operations blindfolded. This can result in unacceptable environmental risks *and* excessive clean-up costs.

These considerations provide the motivation to approach the contamination problem from a new perspective, where all natural processes involved in the migration and fate of chemicals in geomedia are handled in a unified computational framework. The prototype Environmental Information System for Contaminant Migration Risk Assessment (EIS/GWM) computer Platform was developed to support this novel approach. The objectives were twofold:

- Develop a comprehensive methodology, based on a holistic macroengineering (science of large systems) framework to simulate the interaction of the competing natural processes affecting contaminant migration and fate.
- Demonstrate the feasibility of embedding this holistic approach in an integrated/ automated computer platform based on the interaction of natural objects in three-dimensional space.

In this chapter we provide a detailed description of the new holistic approach, along with the theoretical background that is needed to mathematically describe the interaction and evolution of these different natural processes. An implementation of this methodology is provided in the EIS integrated modeling platform, written for MS Windows 95, NT. It offers a systematic, hierarchical framework for the assessment of the relative impact of the various processes by supporting and facilitating the analysis of these complex interactions.

19.2 Holistic Framework for the Contaminant Migration Problem

To solve the Contaminant Migration Problem in general one needs to know the concentration C of a given contaminant everywhere in space and time. What makes this task difficult is the large number of features: (1) geologic features (aquifers, faults, aquitards, rock, and soil formations), (2) interacting contaminants (chemical species), and (3) natural processes (controlling the interaction of chemicals with each other and the environment). The problem is essentially one of managing the fate and migration of a contamination plume with all possible interactions of the hazardous chemicals with the existing soil environment.

Each of the above features comes with its own space and time scale of influence, and this raises the level of complexity to a higher degree. In engineering terms, the *control volume* identifies the representative scale at which the balance of energy and mass is satisfied. Different scales are necessary to accurately describe the evolution of different natural processes. For example, while the groundwater flow problem

can be adequately represented at the macroscale, chemical processes necessitate a microscale representation. The automated simulation framework envisioned here must be able to automatically select the appropriate scale, based on the information produced by the diagnostic phase. This phase concerns the collection of information about constitutive parameters (properties) of the geomedia and the contaminants, and the influence of the natural processes on the potential contaminant migration. Clearly, the quality of this information has an impact on the prognosis phase, where predictions are made on the fate and migration of different contaminants. This phase is critical to the success of the overall effort and usually requires selection of a 3-D grid to characterize the topology of the macro-element discretization of the natural domain. To integrate this effort, there is need for an inference engine, which is the building block that distributes (maps) the properties of the media throughout the grid. Finally, an appropriate integrator is needed at the core of the system, operating at the level of each macro-element and taking into account the effects of all natural processes and their interactions described at the macroscale and microscale levels. The most challenging aspect of this approach is keeping track of the level and degree of *interactions* between the geomedia and the contaminant. A holistic approach to the simulation of the contaminant migration problem is the only way to keep an accurate account of these different levels of interaction.

The Modeling Dilemma

Often the definition of a simulation "model" encompasses a cluster of "small" objects built in another, often larger object. In the case of an environmental model, the "often larger object" refers to the enormously complicated interactions of natural systems. For example, a small object can very well be a hydrologic, ecological, land surface, or groundwater model component. The problem then is how to simplify the larger object in such a way as to make the small object manageable, yet useful. It is a problem compounded by the uncertainty of the interactions involved.

The present state-of-the-art in environmental software is to link a Geographic Information System (GIS) with a spatial-temporal data base and create a loose coupling with existing simulation models through software interfaces. This approach is unsatisfactory because it requires a high level of technical expertise beyond that of the average user and also requires much time and effort. A typical assessment, for example, of the natural biodegradation of an oil fuel spill requires the use of highly specialized simulation packages requiring a sustained effort of several months, leading very often to unacceptable remedial actions.

19.2.1 Limitations of Presently Available Environmental Systems

Geographic Information Systems (GIS) technology has tremendous potential for the analysis and modeling of spatial data. However, modeling within a GIS environment remains presently at an unsophisticated level. Goodchild (1991) described GIS as a "…database containing a discrete representation of geographical reality in the form of static two-dimensional geometric objects and associated attributes, with a functionality largely limited to primitive geometric operations to create new objects or to compute relationships between objects and to simple query and summary descriptions." Furthermore, most current GIS implementations are based on conventional databases that are inadequate for environmental applications. The most significant limitation is the lack of expressive modeling power resulting from the simplicity of the relational data model and its inability to handle the complexity of natural data. Also, Egenhofer and Frank (1992) contend that the relational data model does not match the natural concept humans have about spatial data. They claim users are required to artificially transform their mental models into a restrictive set of nonspatial concepts.

19.2.2 The New EIS Paradigm

The basic idea of the EIS platform is that the environment is perceived as consisting of objects that interact in specific ways. It focuses on modeling objects, as humans perceive them in reality. The new concept of object-orientation places the emphasis on modeling the data rather than on the actions
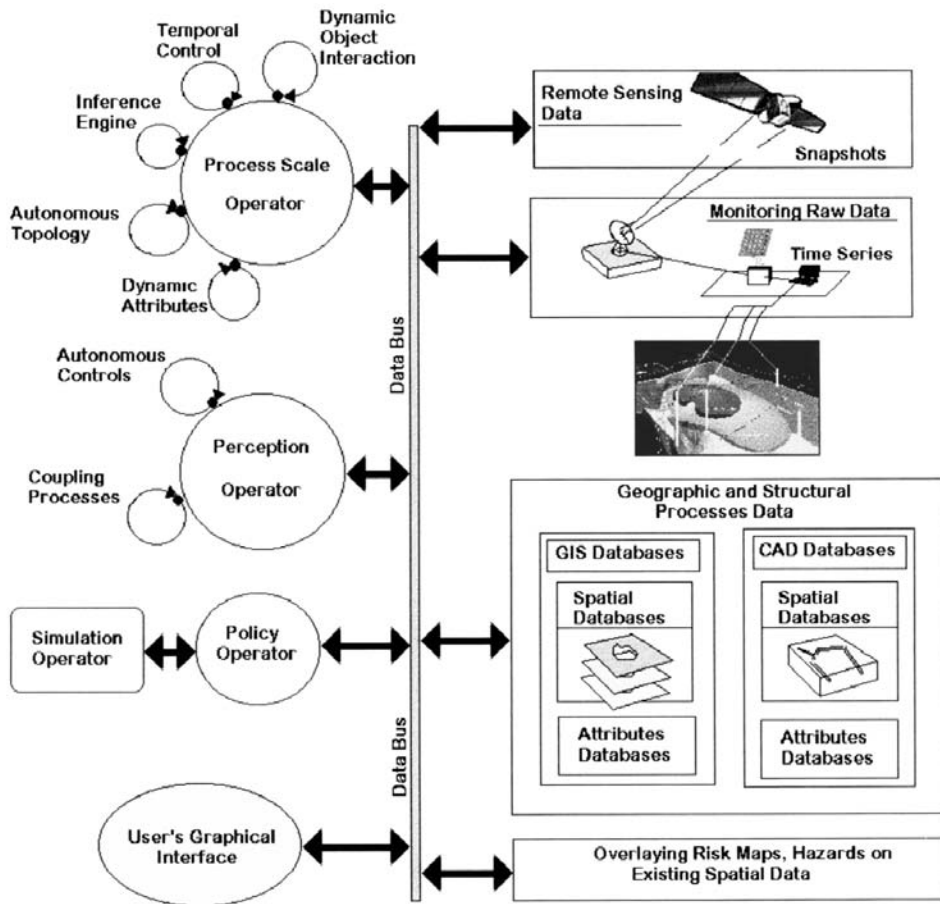
FIGURE 19.2 EIS software components.

(manipulations) to be performed on the data. The ability to handle complex interactions of hazardous wastes is one of the capabilities the new concepts bring to the existing systems. Complex feature support allows real-world features, such as contaminant plumes, which can be recursively split into component parts, to be more accurately represented to include all necessary subfeatures. Operations can be performed within and between these composite features (by-products) taking into account all the inherent natural complexities. This richer data modeling capability is one of the major benefits object-orientation and complex feature support brings to the user of the EIS platform.

The new platform requires the following components: (1) a processing unit for temporal control, (2) a processor for dynamic object interaction, (3) a processor to maintain an autonomous topology, (4) a data base for dynamic attributes, (5) autonomous or user-defined controls, (6) software interfaces, and (7) coupling process models. Figure 19.2 illustrates these basic components as they interrelate with the user's interface.

19.3 Defining the Mechanism of Interactions in the Contaminant Migration Problem

Chemical species interact with each other and their environment through a clearly defined hierarchy of natural processes. These processes can be grouped by category. Examples are: *mechanical*, i.e., the macroscopic balance of mass and momentum; *physical/thermodynamic*, i.e., the balance of energy and entropy

via the introduction of constitutive relations for the fluxes; *chemical*, i.e., interactions among chemical species, and partition, adsorption, precipitation/dissolution for interphase interaction; *electromagnetic;* and, *biological*, via microbial population dynamics (using the Monod Kinetic theory, for example). All these classes of natural processes affect the contaminant migration either in parallel (simultaneous interaction) or in series (sequential interaction). They also affect the soil structure, volume, shear strength, and microstructure interdependently, in response to appropriate relative gradients. In turn, each class of natural processes is represented by a series of processes as illustrated in Table 19.1.

TABLE 19.1 Natural Processes Affecting the Contaminant Interactions

Mechanical Processes	Biological Processes	Chemical Processes	Electrical Processes
Ground deformability	Tree, plant root	Acid-base	Electrokinetics
Surface water flow	Macro-animals	Precipitation	Ion exchange
Groundwater flow	Micro-animals	Volatilization	Electro/osmosis
Diffusion/dispersion	Bacteria	Complexation	
Air flow		Surface reaction	
		Oxidation-reduction	

19.3.1 Contaminant Interaction with Geologic Features

Basic geologic features encountered by a contaminant plume at shallow depths encompass the shallow soil medium, the different rock formations, surface waters in rivers, ponds, or lakes, and the groundwater (subsurface) in aquifers and aquitards. Air is also encountered in geomedia, in the unsaturated zone, and also in dissolved form in water. Where the contaminant is in bulk fluid form, additional phases must be included for the fluid and vapor phase of the contaminant. Typical materials found in the soil and rock media are shown in Table 19.2.

TABLE 19.2 Typical Materials of the Soil Medium

Soil and Rock Formations	Material Type
Soil medium	Gravel, coarse sand, medium sand, fine sand, silt, loess, till, clay, marine clay
Sedimentary rock	Karst and reef limestone, limestone, dolomite, sandstone, siltstone, salt, anhydrite, shale
Crystalline rock	Permeable basalt, fractured and unfractured igneous and metamorphic rock, weathered granite, weathered gabbro, basalt

The difficulty in topologically describing these features arises from their geometric diversity and their multiphase characteristics. The soil and rock formations are porous, fractured, or nonfractured solids of fractal nature (exhibiting noncontinuous characteristics, repetitive at different geometric scales), while the surface and subsurface waters are fluids contained and interacting with the solid phases. For example, in the presence of surface waters, depending on the difference between the piezometric head of the aquifer and the level of the surface water, the surface water may be seeping into the aquifer, or the aquifer may be discharging into the surface waters. Any dissolved species will migrate along this pathway, and additional processes specific to each feature will affect the fate of the chemical, such as benthic exchange processes, water column processes, photolysis, and others.

19.3.2 Interaction of Contaminant Species (Solid, Liquid, and Gas Phases)

A chemical species in the geomedia can be present in the following forms: (1) solid phase (metals and minerals), (2) Gas phase (CO_2 gas in the soil porous medium or the groundwater), (3) free liquid phase (crude oil and liquid organic contaminants also known as nonaqueous liquids or free products), and (4) aqueous solute phase (Na^+ and Cl^- or organics dissolved in the groundwater).

The occurrence of chemical species in different phases must be explicitly incorporated into the development of a holistic model simulating the chemical pathways. A list of the most commonly encountered chemicals is given in Table 19.3. Most of these chemicals can be found dissolved in the groundwater. They all react chemically to various degrees with one another, and with the mineral soil matrix, transferring mass between the liquid, gas, and solid phases. Classification of all possible chemical reactions allows the determination of the type and degree of interaction among contaminant species.

TABLE 19.3 Organic and Inorganic Chemical Species

Major Inorganic Constituents	Bicarbonate, calcium, chloride, magnesium, silicon, sodium, sulfate, carbonic acid
Minor Inorganic Constituents	Boron, carbonate, fluoride, iron, nitrate, potassium, strontium
Trace Inorganic Constituents	Aluminum, antimony, arsenic, barium, beryllium, bismuth, bromide, cadmium, cerium, cesium, chromium, cobalt, copper, gallium, germanium, gold, indium, iodide, lanthanum, lead, lithium, manganese, molybdenum, nickel, niobium, phosphate, platinum, radium, rubidium, ruthenium, scandium, selenium, silver, thallium, thorium, tin, titanium, tungsten, uranium, vanadium, ytterbium, yttrium, zinc, zirconium
Dissolved Organic Compounds	Humid acid, fulvic acid, carbohydrates, aminoacids, tannins, lignins, hydrocarbons, acetate, propionate
Nondissolved Organic Compounds	Halogenated hydrocarbons, phosphorous compounds, organometallic compounds, phenols, amines, ketones, aldehydes, alcohols, esters, ethers, polynuclear aromatic hydrocarbons, alkyne hydrocarbons

Typical chemical reactions are: *acid-base reactions, solution, volatilization, dissolution and precipitation, complexation reactions, surface reactions (sorption, adsorption), oxidation-reduction reactions, hydrolysis, and isotopic processes.* An example of the complexity of the chemical processes as they interact with the soil medium is shown in Figure 19.3. A chemical constituent originally in liquid form is deposited at the surface and migrates under gravity through the vadose zone of the soil medium. Several chemical processes now take place interacting with different phases of the soil medium, changing the fate and migration rates of the chemical constituent. Sorption deposits parts of the constituent on the grain surface, while volatilization and evapotranspiration further reduce the concentration of the chemical constituent in the liquid phase. Finally, under special thermal and ambient conditions the primary chemical constituent is transformed into daughter products. Clearly, a computational framework is needed to accommodate the plethora of processes, reactions, phases, and geologic, spatial, and temporal scales. Furthermore, the mathematical abstraction of the various modeling elements requires the development of a road map for their systematic (automated) intercommunication.

Macroengineering is the discipline that systematizes the development of an information network between components of large, complex, integrated systems. The contamination problem in geomedia fits this description and is ideally suited for resolution by the macroengineering approach.

19.4 Macroengineering Formulation for Contaminant Interaction

The complexity of the large number of physicochemical constituents interacting with each other in the geoenvironment necessitates a new framework in order to keep track of their interdependency. Furthermore, these interactions take place through a multitude of natural processes in a multiphase environment in which delicate balances exist between the solid, fluid, and gas phase of a large array of chemical species and a plethora of different soil conditions.

The macroengineering framework is needed to evaluate every constituent that is selected for analysis in terms of its potential interaction and impact on the geoenvironment. The basic concept behind the macroengineering framework is simple: an informational network keeps track of the simulation requirements, adapting itself to the degree of interaction between a given constituent and the characteristics of the geoenvironment and other chemicals.
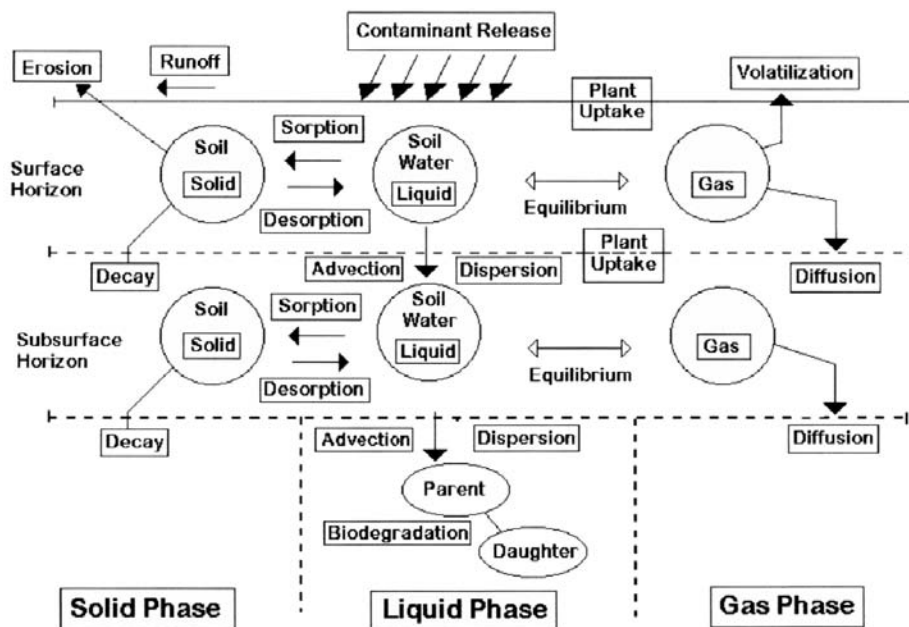
FIGURE 19.3 Interaction of multiple phases: solid, fluid, and gas.

The goal of the proposed macroengineering architecture is to create a homogeneous, functional aggregate by integrating diverse subsystems (algorithmic and cognitive strategies embedded in different network layers) governing the contaminant migration problem. Typically, these subsystems exhibit incompatible protocols and simulation algorithms that need to be interfaced in a unified system. This can be accomplished according to the following macroengineering principles:

- The overall simulation system must be layered and structured to accommodate conflicting scales and time marching strategies among different natural processes affecting the contaminant migration in the geomedia.
- One logical network must be able to service all communications requirements between layers and distributed databases such as text, data, and images.
- There must be ease of movement between various layers, making different simulation components appear similar to the user. This integration necessitates similar protocols, common files, and subroutines/functions that are independent of their layered characteristics.

In the adopted layered architecture, each layer is served through a protocol designed and implemented for that specific layer. Successive layers are interconnected through interfaces. The latter are software constructs which accomplish two important functions: message passing from layer to layer, and mapping of one protocol into the other using the appropriate knowledge banks and databases. Both the knowledge banks and databases coexist and cooperate within finite networks of loosely coupled, concurrent communicating layers. Furthermore, the overall system architecture accounts for a query-dominant computer environment. To avoid inconsistency, this architecture calls for synchronization at each step of a transaction or message exchange as follows. Transactions are given unrestricted read access to the database during the initial read phase. Then transactions enter a validation phase to check for conflicts using the appropriate knowledge base. If successful, the transactions enter a write phase to incorporate their updates into the available database. In fact, each layer represents all needed abstractions pertaining to a natural process characterizing the interaction of a contaminant with the soil medium. One layer, for example, represents the advection of the contaminant (necessitating the simulation of the groundwater flow), while another layer deals with the biodegradation of the same contaminant. Both processes take place within

the same time frame requiring synchronization. Messages between the two layers are controlled by an interface scheme, part of the knowledge bank of the macroengineering system which includes the following elements (see also Figure 19.2):

- A perception operator (an invariant reference frame) topologically connected to the different abstraction layers representing the different natural processes and databases. The perception operator includes an inference engine to determine the characteristics of the reference frame.
- A process scale operator to coordinate the information flow between simulation layers and between integration operations of different space and temporal scales.

These macroengineering components require the determination of a concise framework to construct the different interactions that occur at different space and time scales. Interactions can be classified in terms of the natural (mechanical, chemical, and biological) processes. The variables and parameters characterizing the physical processes govern these interactions. At one end of the scale, the engineer/scientist is interested in controlling the various processes at the site scale with a resolution of the order of 10 to 100 ft., while physicochemical processes may be operating at smaller scales. A hierarchy of scales and processes is built into the system such that integration is automatically operated to produce values of the parameters at the scale prescribed by the analyst.

Another aspect of the interaction is the parameter or variable with which the interaction takes place. For example, for surface water (pond, river) interaction with groundwater, the interaction parameter is the piezometric head. For the case of spills or a contaminant plume, the interaction takes place via concentrations. These features or processes can be thought of as subgrid scale features. That is, the model is built of blocks or control volumes at the scale of the resolution required by the analyst. All other features are isolated and addressed at the level of each macroelement separately.

The innovation consists of considering all possible interactions at the level of the macroelement, not the entire domain. This approach through its built-in hierarchy of competing processes allows one to go as deep as necessary in the microscale of the phenomenon, conceivably even the molecular scale. This conceptualization makes it possible to alter the scale of the control volume or macroelement without having to re-enter the geometry (topology) of the geologic and loading features.

These steps are illustrated in Figure 19.4. Natural processes take place at different scales but the analyst need only concentrate on determining their required level of resolution, that is the control volume or macroelement at the macroscale. The processes that may appear in the simulation may include geologic features such as geologic faults and mechanical processes at the mesoscale, or chemical and biological processes at the microscale. The level of interaction or density of the network of interactions grows geometrically with each additional scale introduced. But because that level of interaction is dealt with locally, the problem never grows out of proportion. This is true because at the control volume level the more minute interactions require intraelement integration. It is their integrated effect that is felt at the between macroelement interaction level. This is achieved by introducing and paying close attention to an *invariant reference frame*.

19.4.1 Invariant Reference Frame

The basis of the macroengineering approach is an adaptable Eulerian referential grid system that controls all the informational layers needed for a particular contamination simulation scenario. This referential grid is formed of macroelements as shown in Figure 19.5. At each macroelement (information node) are operated integrations at the levels of geologic features, natural processes governing the contamination problem, and corresponding loading conditions, taking into account their respective interactions.

In a sense, the adaptable Eulerian grid forms an informational network, whose topology (branches, nodes, intersections), unlike wire networks in conventional communication theory, can be changed during the course of a study. The information content of every node is automatically inferred from a database network that includes monitoring data, maps, images, soil characteristics, data related to the natural processes and their constitutive parameters, as illustrated in Figure 19.5. The databases and their
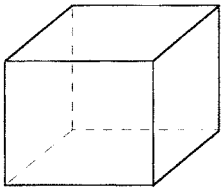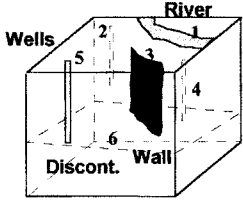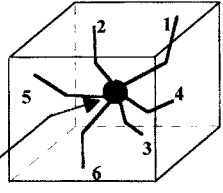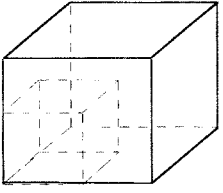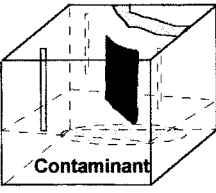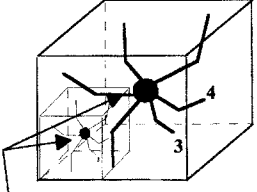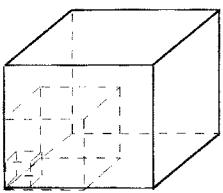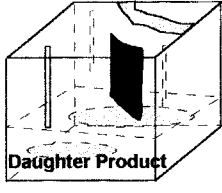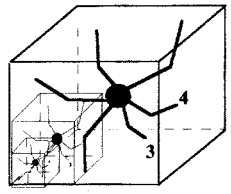
Control Volume for Interaction (Macroelement)	Modeling Features	Density of Interactions
Macroscale	**Geologic Features and Loadings**	**Low Number of Interactions**
Definition of Topology		Informational Node
Mesoscale		**Medium Number of Interactions**
Water Balance	Interaction via Mechanical Processes	Informational Nodes
Microscale	**Chemical Processes**	**High Number of Interactions**
Chemical Mass Balance	Interaction via Chemical Processes	Informational Nodes

FIGURE 19.4 Control volumes for modeling interactions.

temporal updates form the information system invariants: the information they contain remains unaltered; what changes is the inference to accommodate the characteristics of the updated informational network (Eulerian grid). Because of this network architecture, data can be automatically distributed to the adaptable Eulerian grid (along with the associated inference error), so as to achieve optimal simulation
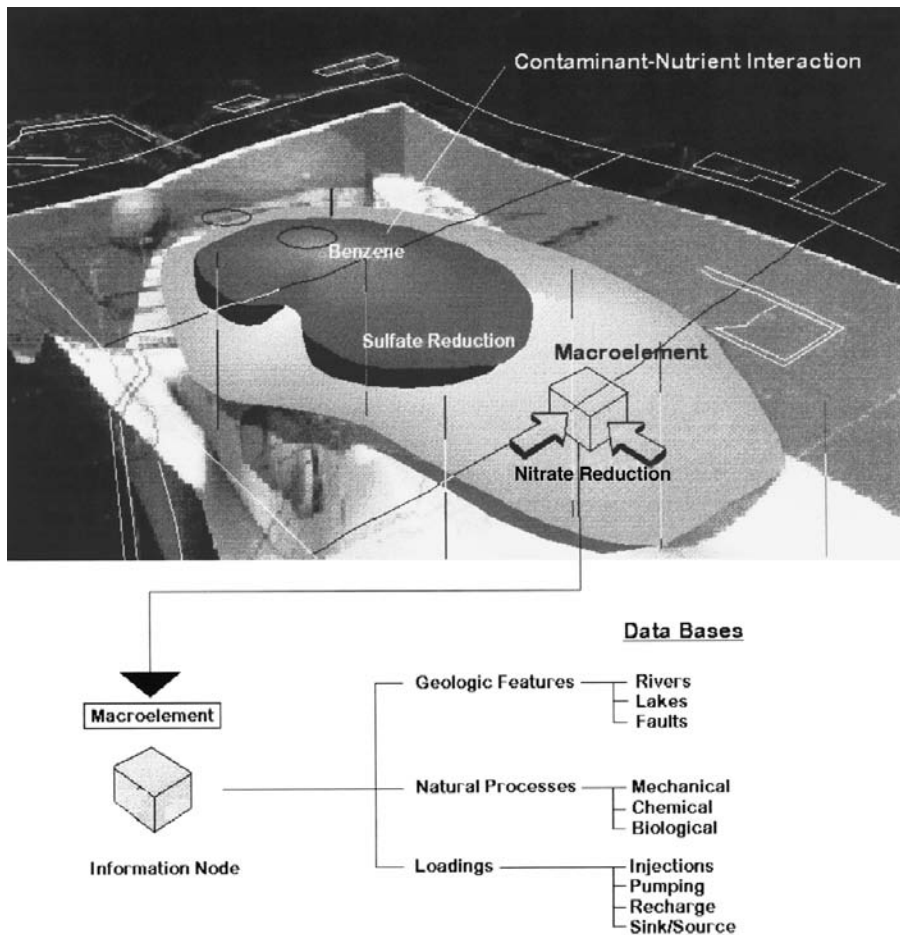
FIGURE 19.5 Eulerian referential system.

results. The density of the grid is directly proportional to the density and distribution of the monitoring network. Thus, this network architecture can also be used to optimally design monitoring strategies.

19.4.2 Inference Engine

An integral part of the macroengineering framework is an inference engine, which performs a mapping from the vertices of the monitoring network to the information nodes (macroelements) of the Eulerian grid. This mapping operates at two levels. At the log-point level, raw data (e.g., cone penetrometer signal, tip resistance, friction, pore pressure, contaminant concentration profile) are sorted and interpreted into information about soil system stratification, anisotropy, and other characteristics. At the grid level, the geomedia properties and the associated inference error are assigned to the grid information nodes by geostatistical means (generalized covariance kriging) or fractal mapping.

Monitoring technologies include electromagnetic tomographies, magnetic and seismic imaging. The macroengineering framework allows automatic implementation of data fusion principles. Data fusion pertains to the sorting and prioritizing of redundant (and sometimes conflicting) information. Multiple (disparate) sources of data are automatically (expertly) combined into one reliable information data bank, after taking into account all available information, sometimes qualitatively, and rejecting erroneous or biased information. These fused data banks form the invariants in the adaptive Eulerian grid mapping.

19.4.3 Process Scale Operator

At the heart of the macroengineering approach is a time and space synchronization mechanism under the control of the *process scale operator*. This operator controls the integration at the microscale and macroscale levels for each natural process retained for the simulation within the domain of the macro-element grid. The starting point is the determination of the microscale (control volume) in which the mechanical, chemical, and biological processes influence the contaminant migration in the geomedia as shown in Figure 19.6. These natural processes govern the fate (short- and long-term distribution) of the chemicals in the geomedia (air, soil, water, and biota).

Organic contaminants, for example, can reach the saturated zone of the geomedia either in a dissolved phase in water or as organic liquids that may be immiscible in water. Figure 19.6 shows typical interacting processes that affect the fate of a given nonaqueous phase liquid (NAPL) such as diffusion of chemicals in the NAPL, volatilization, dissolution from NAPL into macropore fluid (water), and abiotic and biotic transformations. These contaminants travel with the soil moisture and are influenced in their migration by various factors. For example, the rate of biodegradation varies for each chemical with time, microbial population characteristics, temperature, pH, and other reactants. Similar types of pathways are also
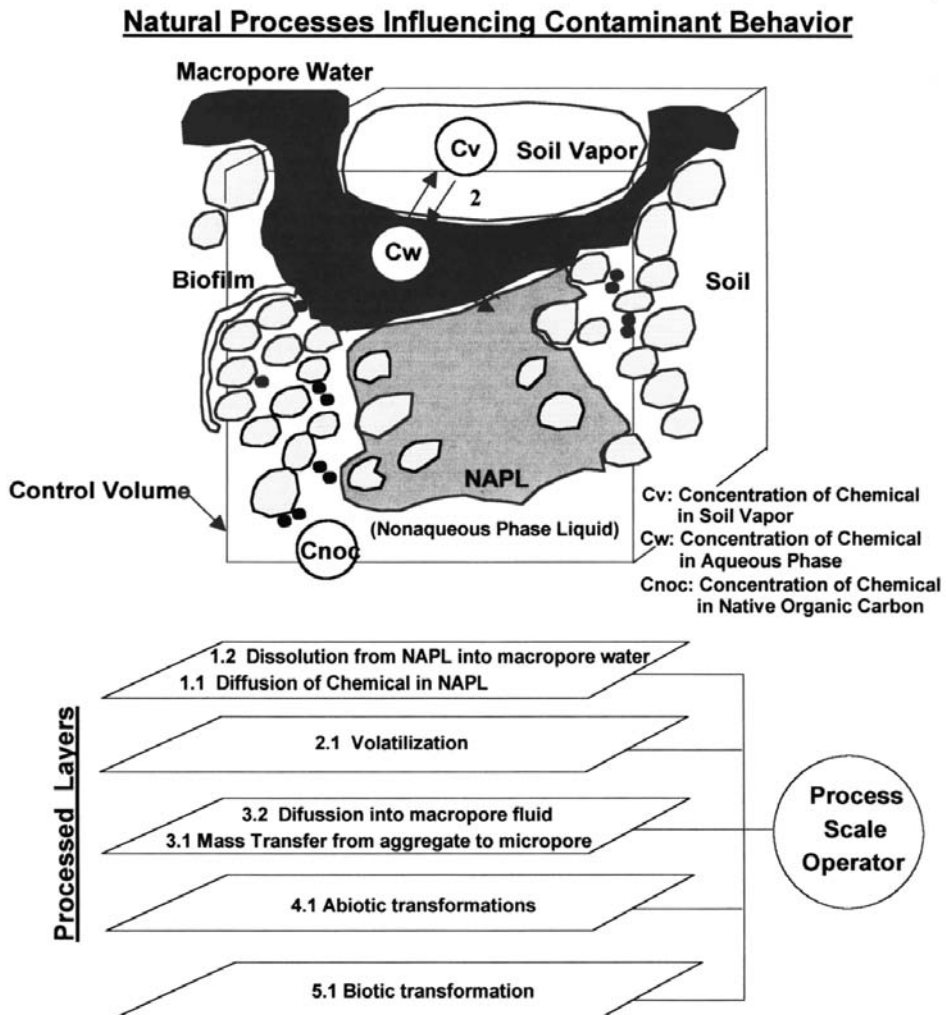
FIGURE 19.6 Natural processes influencing contaminant behavior.

observed for inorganic contaminants and include the following natural mechanisms: volatilization, solubility, fast aqueous reactions, slow aqueous reactions, speciation, soil interaction, precipitation, and bioaccumulation.

Managing all these components for both the diagnosis and prognosis phases of a study necessitates a macroengineering framework that exhibits the following capabilities:

- Classification of the natural processes and their characteristic properties affecting the contaminant migration during the diagnosis phase of the study (site investigation)
- Synchronization of loading parameters
- Synchronization of time scales and integration of the natural processes at the macroelement level during the prognosis (predictive) phase
- Interaction of different chemicals and their natural processes

19.4.4 Classification of Natural Processes

Here lies an inherent complexity associated with the determination of the scales, representative of a particular natural process. Generally, the physical properties characterizing these processes are dependent on the scale in which they are measured. For example, most *in situ* investigations are representative of a macroscale averaging scheme. However, when soil properties are averaged on a different scale, their values are drastically changed, as shown in the top panel of Figure 19.7 for longitudinal dispersivity. At the transition zone between various scales, it is difficult to define a representative value. Other processes such as fracture flow also share this tendency for dispersivity values to be scale-dependent. Therefore, an overall process scale operator is needed to adapt the simulation strategy taking scale transitions into account. This problem of scale transitions and how to account for them has implications for the prediction of the contaminant migration and fate. The approach that has been adopted uses data fusion principles relating, for example, dispersivities to geostatistical models of another physical parameter like the hydraulic conductivity. For the success of this approach the geostatistical description of properties like hydraulic conductivity, porosity, or cation exchange capacity, must capture the heterogeneity of the geomedia by a small number of statistical parameters such as the mean, the variance, and the correlation length scales. A generalized covariance kriging scheme provides the quantification of these statistical parameters based on the data collected at the site.

19.4.5 Synchronization of Loading Parameters

Loading parameters for a contaminant migration simulation can have one of two forms: they can be either derived from the natural climate cycle (hydrologic cycle) or be caused by a man-made activity, pollution, accident, or a remedial action. The time variation of the intensity of these loadings is of great importance to the study.

The macroengineering framework keeps track of the relative importance of these parameters in an automatic fashion. For example, the simulation of the fate of a chemical is different in the case of a dissolved trace-level contaminant from the case of a large-scale spill release. Both the release (source) mechanism and the evolution processes of the chemical (advection, dispersion, reactions) are affected by the time incrementation used in the simulation, and the model automatically identifies the proper time discretization satisfying all requisite criteria (e.g., Courant conditions).

19.4.6 Integration of the Natural Processes at the Macroelement Level

The inference engine as described in the previous section identifies the spatial distribution of the soil, fluid, and chemical properties, allowing the numerical integration of all necessary natural processes at the macroelement level. Several schemes are available to perform this numerical integration throughout the volume of each macroelement accounting for all possible modeling features such as vertical and horizontal wells, slurry walls, geosynthetic or clay liners, drains, rivers, lakes, and ponds. This scheme
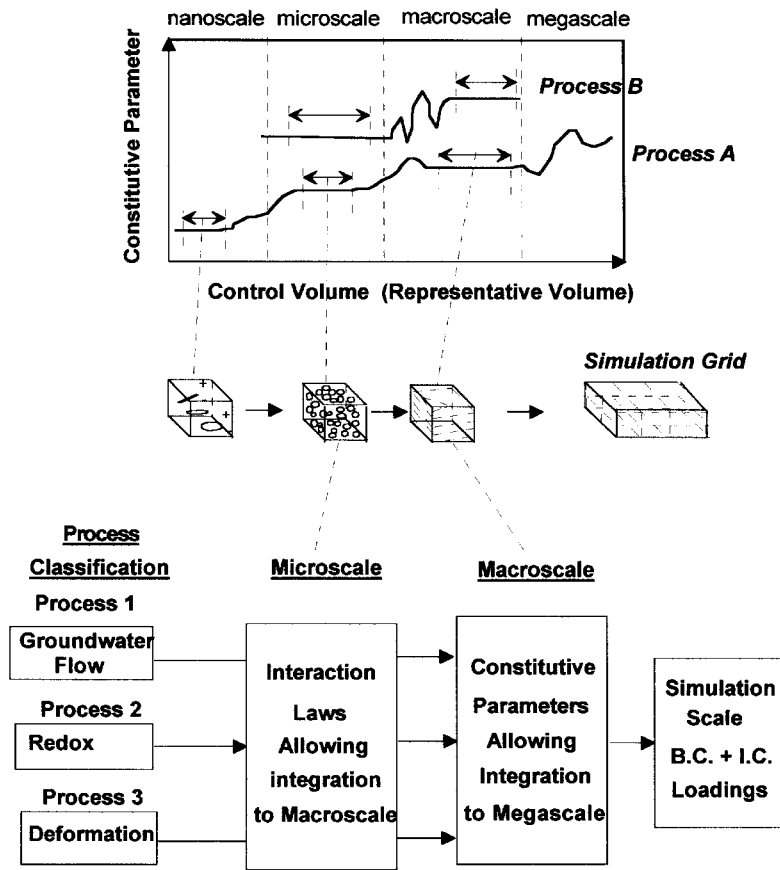
FIGURE 19.7 Determination of microscopic and macroscopic levels.

encompasses techniques like finite differences, finite elements, discrete elements, and boundary elements. Any one of these methods fits the requirements of the macroengineering approach, as long as an exhaustive library of macroelements exists to allow an automatic integration of all pertinent features and processes interacting within the confines of each macroelement. Compatibility of pertinent physical parameters between adjacent macroelements (for example, energy fluxes and chemical concentrations) is closely examined at regular intervals by the appropriate time marching algorithm, which is an integral part of the scale control operator. This approach assures a comprehensive solution to the contaminant migration problem.

The key ingredient, which always limits the accuracy of the solution algorithm, is uncertainty. There are many sources of uncertainty in the geomedia, such as spatial and time variations of geological conditions within the macroelement, soil-input parameters, and contaminant degradation rates. The degree of uncertainty is increased, in most cases, by lack of sufficient data for parameter estimation and model validation, and the margin for error is intimately related to the sensitivity of the receptors at or near the contaminated area.

The Process Scale Operator provides an estimate of this uncertainty on the basis of the kriging procedure of the inference engine and the geometric configuration of each macroelement. An error distribution is automatically estimated for the entire simulation domain. The operator also checks for overall conditions such as generating a denser mesh to decrease the uncertainty in relation to the spatial distribution of the input information (site investigation).

19.4.7 Interaction of Chemicals and Appropriate Time Marching Algorithms

The movement of fluids is due to gradients in hydraulic head, density, temperature, or voltage. Additionally, mass also moves in response to fluid flow (advection) while undergoing dispersion, chemical and biological reactions, and degradation. These interrelated processes are part of a large family of processes that the process scale operator must identify and simulate. Interacting processes (coupled) may not be described by a simple relationship between one flux and one driving force (mechanism). Such natural phenomena are, for example, chemical and electro- and thermo-osmotic processes. Each of these processes is embedded in its geometric multiphase domain. The degree of interdependence is estimated by relationships of the following form:

$$J_i = L_{ij} \cdot G_j \qquad (1)$$

where
J_i = interaction of component (i) (customarily a flux)
L_{ij} = interdependence coefficient (material property, conductance)
G_j = gradient of component (j) (normally some potential)

This, in fact, is a general rate equation. It is immediately recognized that Darcy's law, Fick's law, Fourier's law, and Ohm's law can all be represented by this equation shown as the diagonal terms of Table 19.4. This table also contains the interaction of the various rate equations. The driving mechanisms at the macroscale level are listed across the top; flows (or fluxes) are given along the side of the table. The off-diagonal terms are the interaction laws (osmotic laws). Many of these interaction phenomena are thought to influence the chemical migration. These interactions are embedded automatically in a simulation time frame.

TABLE 19.4 Interrelationship Between Driving Mechanism and Different Fluxes

Flux	Driving Mechanism			
	Head Gradient	Concentration Gradient	Voltage Gradient	Temperature Gradient
Fluid	Darcy's law	Chemical osmosis	Electro-osmosis	Thermo-osmosis
Mass		Fick's law		Soret effect
Current			Ohm's law	
Heat		Dufour effect		Fourier's law

Interactions also exist between different chemical species. They are defined according to equilibrium or chemical kinetic conditions, which provides a useful framework for describing reactions in relation to time and pathways using the same generalized rate law. It must also be recognized that these reactions take place between different phases (solid, fluid, and gas).

The process scale operator coordinates the time marching algorithms according to the nature of the simulated process. Figure 19.8 illustrates the reaction half-lives for many of the common reactions in aqueous systems. The time incrementation for the appropriate simulation is automatically selected so as to satisfy the prevailing convergence criteria (e.g., Courant conditions).

As a general rule, the fastest reactions are the solute-solute or solute-water reactions with half-lives of fractions of seconds to, at most, a few minutes. These reactions are examples of homogeneous reactions, reactions that occur in a single phase (examples of such reactions are acid-base and complexation reactions). The dissolution-precipitation reactions are examples of heterogeneous reactions involving more than one phase. These reactions have half-lives varying from days to several years. Heterogeneous solution or exsolution reactions involving a gas phase have relatively larger half-lives than do liquid-solid reactions. Reactions on surfaces such as adsorption-desorption have half-lives ranging from seconds to days. The larger half-lives describe surface reactions in the porous matrix of rock fragments. Redox reactions, on the other hand, are in general relatively slow because they are mediated by microorganisms (half-lives ranging from hours to a few years). Transformation reactions involving organic contaminants
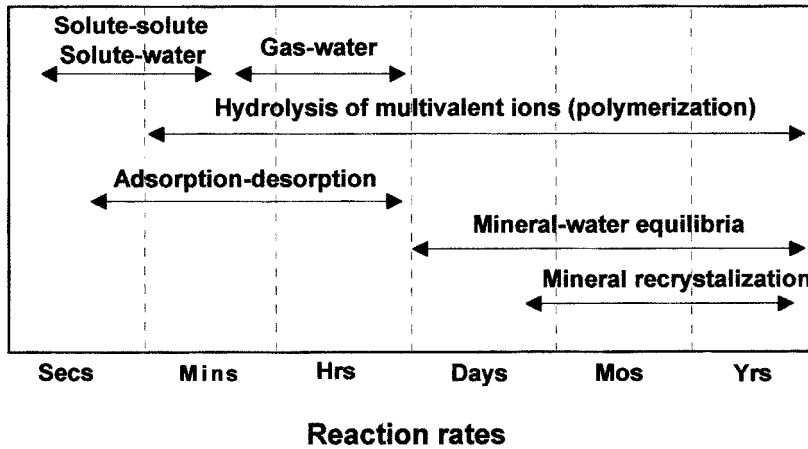
FIGURE 19.8 Reaction half-life for common reactions in aqueous systems.

TABLE 19.5 Compatibility Parameters of Different Natural Processes

Natural Process	Code #	Compatibility Parameters	Time Synchronization	Priority Index
Ground deformability	M101	Stress field	Years	2
Surface water flow	M102	Flux	Years	2
Advection (Groundwater flow)	M103	Flux	Years	1
Diffusion/dispersion	M104	Flux	Years	1
Thermal	M201	Flux		2
Air flow	M202	Flux	Hours	2
Acid-base	M301	Mass		2
Precipitation	M302	Mass	Hours	2
Volatilization	M303	Mass	Hours	2
Complexation	M304	Mass		2
Surface reaction	M305	Mass	Hours	3
Oxydation-reduction	M306	Mass	Minutes	2
Hydrolysis	M307	Mass	Years	3
Isotopic processes	M308	Mass		3
Bacteria	M401	Mass		2
Micro-animals	M402	Mass		3
Tree, plant roots	M403	Mass		4
Electroosmosis	M501	Flux		3
Electrophoresis	M502	Flux		4
Electromigration	M503	Flux		3
Magnetic separation	M504	Mass		4

in groundwater depend on the physical and chemical properties of the particular compound (for example, hydrolysis) and in addition, on the presence of a variety of microbes (causing biodegradation).

The macroengineering framework allows a hierarchization of all natural processes as they evolve with time in the geomedia. Processes that have a large impact on the contaminant migration problem are given the highest priority by the process scale operator that keeps an inventory of default values for time synchronization and scale compatibility, as illustrated in Table 19.5.

19.4.8 Simulated Processes Affecting *In Situ* Remediation Technologies

The ultimate goal is to define a macroengineering framework to help organize the development of *in situ* remediation technologies and reduce the overall development effort and cost.

TABLE 19.6 *In Situ* Remediation Technologies and Corresponding Natural Processes

In Situ Remediation Technology	Category	Natural Processes Affecting the Geomedia
		Containment/Immobilization
Bio-immobilization		Advection, diffusion/dispersion, deformability, redox
Grouting		Advection, diffusion/dispersion, deformability, hydraulic stability, soil permeation, grout reaction
Immobilization	Chemical	Advection, diffusion/dispersion, deformability, sorption, ferric reactions
	Physical	Advection, diffusion/dispersion, deformability
	Redox manipulation	Advection, diffusion/dispersion, deformability
Subsidence control		Advection, diffusion/dispersion, deformability
Barriers	Grout	Advection, diffusion/dispersion, deformability
	Chemical	Advection, diffusion/dispersion
	Cryogenic	Advection, diffusion/dispersion, deformability, heat
	Electrokinetic	Advection, diffusion/dispersion, electroosmosis, electrophoresis, electromigration
In situ vitrification		Advection, diffusion/dispersion, deformability, heat
Treatment		
Bioremediation	Microbes	Advection, diffusion/dispersion, deformability, redox, precipitation
	Biomass	Advection, diffusion/dispersion, redox
	Biofilters	Advection, diffusion/dispersion, redox
Permeable treatment	Barrier	Advection, diffusion/dispersion
Chemical oxidation		Advection, diffusion/dispersion
Nitrate destruction		Advection, diffusion/dispersion, deformability
In situ oxidation		Advection, diffusion/dispersion, deformability
Subsurface Process Control		
Electrokinetic		Advection, diffusion/dispersion, electroosmosis, electrophoresis, electromigration
In situ heating		Advection, diffusion/dispersion
Soil flushing		Advection, diffusion/dispersion, deformability
Bioleaching		Advection, diffusion/dispersion, deformability
Auger/Jet mixing		Advection, diffusion/dispersion, deformability

Technologies that will benefit from the proposed macroengineering approach are given in Table 19.6. This follows the In Situ Remediation Integrated Program (ISR-IP) classification by the U.S. Department of Energy.

19.5 Examples from Case Studies

The above "blueprint" was used in developing the prototype EIS platform that was used for performing emerging technology case studies. To better recognize the necessity and advantages of the new approach, a comparison is given in Table 19.7 between the requirements of a conventional and the new approach.

These macroengineering concepts were tested in three independent case studies that are presented hereafter.

Case Study No 1. Evaluation of Biodegradation Potential of BTEX Contamination at Several Air Force Bases Sites

The EIS platform provided the framework to determine the long-term monitoring (LTM) plans and to establish points of compliance (POC) for BTEX migration in shallow groundwater.

Innovations: In these projects for the first time the interaction of the BTEX contaminant with three available electron acceptors (oxygen, sulfate, nitrate) and two metabolic products (iron and methane)

TABLE 19.7 Comparison Between Conventional and EIS Studies

Computational Steps of Conventional Approach	Computational Steps of a Macro-engineering Approach
For each simulation scenario:	For a particular site:
1 Create maps and reports detailing existing features (land uses, soil associations, weather regions, roads, hydrography, basin and county boundaries, etc.).	1 Select a regional plan and spatial scale.
2 Calculate contaminant sources; contaminant assimilation along streams and canals, and the final contaminant loading to a defined receptor for a particular regional plan.	2 Automatically sort site data and activate appropriate simulation options in accordance with contaminant control practices within specified spatial scale. Several scenarios can be handled in parallel.
3 Create maps and reports detailing material and chemical properties of geomedia compiled from site investigations.	3 Display analysis results and the effects of the changes made to create the new regional plan.
4 Pass information to and call appropriate external simulation models, incorporate simulation results back into available GIS databases, and display results in reports and maps. It displays time-series output of simulation runs.	
5 Select groundwater quality, quantity, and weather stations in the studied basin and display time-series charts and graphs of monitoring data.	
Start the procedure all over for a new scenario analysis.	

was modeled in three dimensions (multilayer system). The reproduction of the field data for these chemicals demonstrated the importance of using a 3D framework to model the interacting natural processes, advection, dispersion, sorption, and biodegradation, that influence the contaminant migration. The three-dimensional model was much more accurate than the conventional two-dimensional models and allowed a five-year prediction of the contaminant migration, leading to the determination of a reliable long-term monitoring plan.

Sample Results: Figure 19.9 shows the simulation results of the benzene migration at an Air Force base in Florida, which include the interaction of the main contaminant with several electron acceptors (oxygen, sulfate, and nitrate). Of interest is the shape of the benzene plume as shown in the child window (Figure 19.9), which indicates a preferential migration path in the second soil layer.

The biodegradation model was successfully calibrated for a period of 365 days, considering the stoichiometry of the aerobic and anaerobic reactions shown in Table 19.8. The simulated oxygen depletion is illustrated in Figure 19.10, clearly indicating the extent of this depletion below the contaminant plume.

Based on these simulations six LTM (long-term monitoring) wells and three poc (point of compliance) wells were installed. The EIS platform allowed a close evaluation and identification of the wells where sampling can be safely discontinued.

Case Study No 2. **Regional Groundwater Modeling Initiative in Delaware**

The objective of the Regional Ground-Water Modeling Initiative (RGMI) was to assess the extent of groundwater contamination in an industrial zone in Delaware for waste minimization and pollution prevention. The industrial zone is clustered with a significant concentration of petrochemical and chemical facilities (superfund and RCRA sites). It was conjectured that the contaminated areas in, around, and under the different facilities were impacting each other, both in terms of spreading contamination beyond facility borders, and/or being influenced by remedial actions on adjacent sites. Remediation of this region on an individual site-by-site basis could result in one site having a negative impact on its neighboring sites. A holistic modeling approach was adopted to address the complexity of the site conditions considering the delicate interactions of all natural processes that may influence the ground-
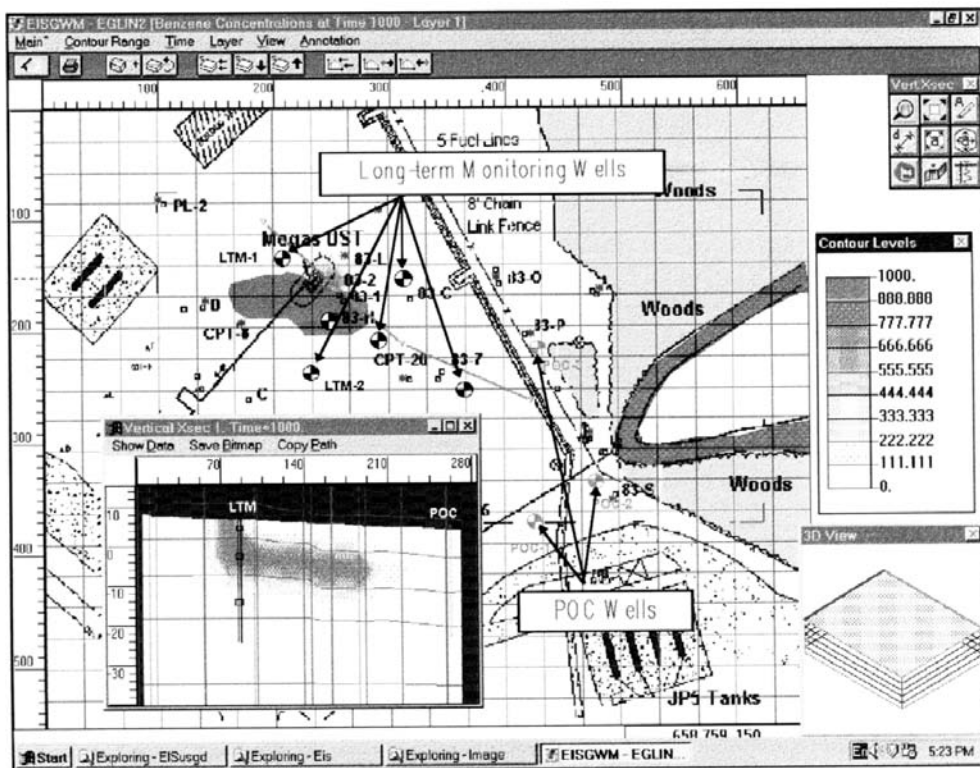
FIGURE 19.9 Location of LTM and POC wells.

TABLE 19.8 Biodegradation Reactions and Corresponding Stoichiometry

Reaction Type	Reaction Equation	Ratio Benzene/ Electron Acceptor
Benzene oxidation Aerobic respiration	$7.5O_2 + C_6H_6 = 6CO_{2,g} + 3H_2O$	3.14
Benzene oxidation Denitrification	$6NO_3 + 6H + C_6H_6 = 6CO_{2,g} + 6H_2O + 3N_{2,g}$	4.9
Benzene oxidation Manganese reduction	$30H + 15MnO_2 + C_6H_6 = 6CO_{2,g} + 15Mn^2 + 18H_2O$	
Benzene oxidation Iron reduction	$60H + 30Fe(OH)_{3,a} + C_6H_6 = 6CO_{2,g} + 30Fe^2 + 78H_2O$	21.8
Benzene oxidation Sulfate reduction	$7.5H + 375SO_4^2 + C_6H_6 = 6CO_{2,g} + 3.75H_2S + 3H_2O$	4.7
Benzene oxidation Methanogenesis	$4.5H_2O + C_6H_6 = 2.25CO_{2,g} + 3.75CH$	0.78

water contaminant migration. Processes at work include advection, sorption, dispersion, and biodegradation as they evolve in the three-dimensional regional scale domain (Figure 19.11).

Innovations: An automated subgrid and sublayer module was developed for the automatic passage from the regional scale (where most piezometric heads were), to a site scale where the water quality data were available. Because of the disparity and paucity of the existing data, a two-level plan was used, starting with a groundwater simulation at the regional scale, then proceeding with subgrid, site-scale level water quality simulations that made use of site level data. The compatibility between the different scale grids was enforced at the level of piezometric head gradients between site grids.
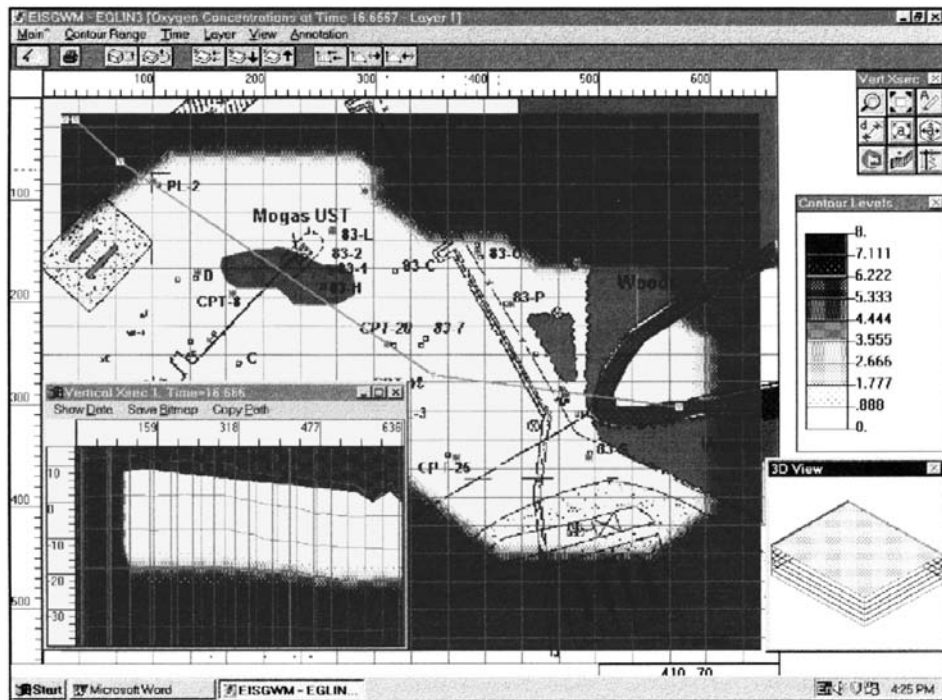
FIGURE 19.10 Oxygen depletion.
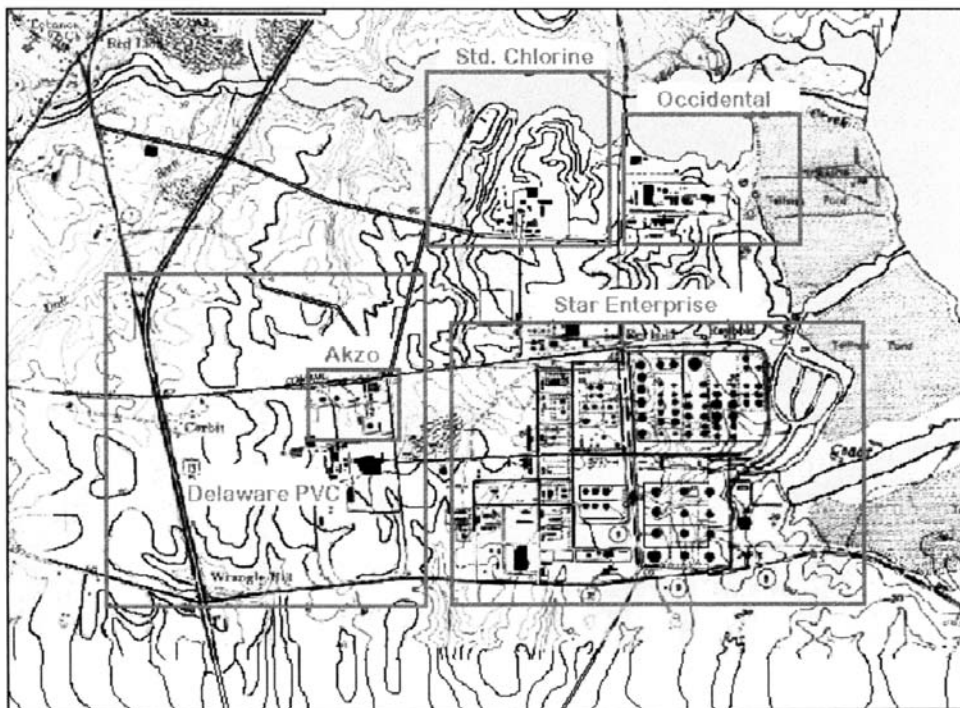
FIGURE 19.11 Regional Ground-Water Modeling Initiative.
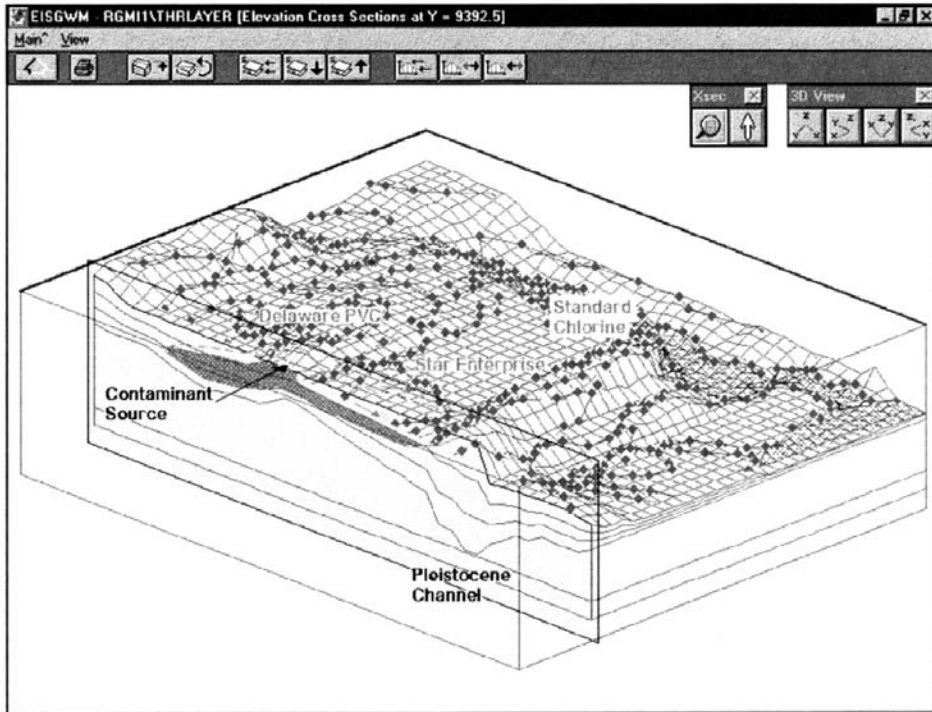
FIGURE 19.12 Migration of chlorinated VOCs after a 15-year simulation.

Sample Results: The study at the regional scale clearly showed the potential for contaminant migration between adjacent sites. In particular, we explained the presence of chlorinated contaminants underneath one site as migrating from another site. Of particular interest was that chlorinated VOCs (EDC, VCM, and TCE) formed a plume which initially spread eastward and then diverged north-south along a pleistocene paleochannel (Figure 19.12).

Similar findings were obtained for the other sites. Clearly the holistic approach adopted in this study was the only alternative that could help identify an optimum plan to contain and alleviate the existing pollution sources.

Case Study No 3. **Management of Raw Data at the National Geotechnical Experimentation Sites**

Under the auspices of the Federal Highway Administration's (FHWA) National Geotechnical Experimentation Sites (NGES) program, a new software module was developed using the EIS Platform to incorporate the following features:

1. An editor for dynamically assigning attributes to log points of the site investigation. This tool allows the user to freely work with any set of field data by scripting the properties of each sensor (measuring device).
2. An automatic procedure allowing the user to delineate the medium stratification on the basis of all profile data (CPT, CPTU tests).
3. An automatic procedure for developing spatial distributions of the measured parameters (properties) as integrated over the user-specified stratification, based on the EIS geostatistical (kriging) procedure, which also provides interpolation error estimates.
4. An automatic procedure to predict the anticipated probe signal at points which have not been sampled. This feature (fractal kriging) allows the site manager to complete the QA/QC loop on

the spot and develop a sense of reliability on his/her interpretation of the site geostratigraphy (e.g., look for missing clay lenses).

Innovations: The great challenge in this project was to develop an optimum inference engine capable of inferring data with depth (signatures of physical parameters) at any location inside the site domain from sparse site data collected using boring technologies such as drilling and CPT tests. A typical log point requires 3000 data entries per physical parameter for a 100-ft. borehole. Assuming that on the average we have 50 log points per site, then the inference algorithm to determine the spatial distribution of the measured physical parameter within the perimeter of the site manipulates 150,000 data entries. Conventional inference schemes make the problem intractable (see Table 19.9). A dramatic reduction in memory is obtained with the new algorithm, which is based on a fractal kriging scheme: only 10 to 20 data entries are stored per log point instead of 3000. The reduced data set is then processed with the EIS generalized kriging scheme. Figure 19.13 illustrates the key components of the new scheme, and Figure 19.14 illustrates fractally extrapolated CPT signatures.

TABLE 19.9 Comparison of Conventional and Fractal Kriging

Grid Size	Node Number for Conventional Kriging Scheme	Node Number for New Algorithm (Fractal Kriging)
10×10	30,000 nodes	500 nodes
100×100	30×10^6 nodes	50,000 nodes
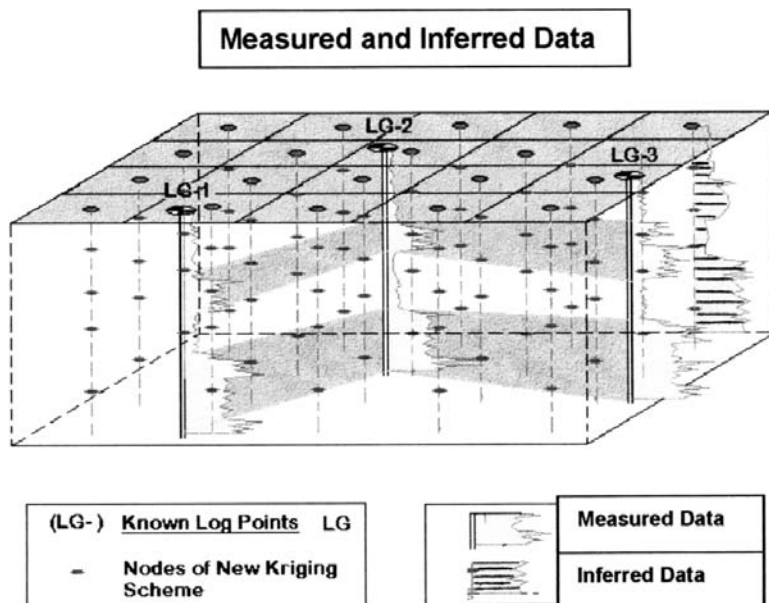
FIGURE 19.13 Grid for kriging procedures and log-points.

Sample Results: This software provides an automated management tool for the raw data of a site investigation. Most important, it allows the combination of various site signatures into modeling parameters that are automatically integrated into the simulation models. All information pertaining to the site is optimally stored in a database that is accessible at any time during the simulation activities (Figure 19.14), and a scenario analysis to identify an optimum remediation scheme becomes a simple and efficient task.
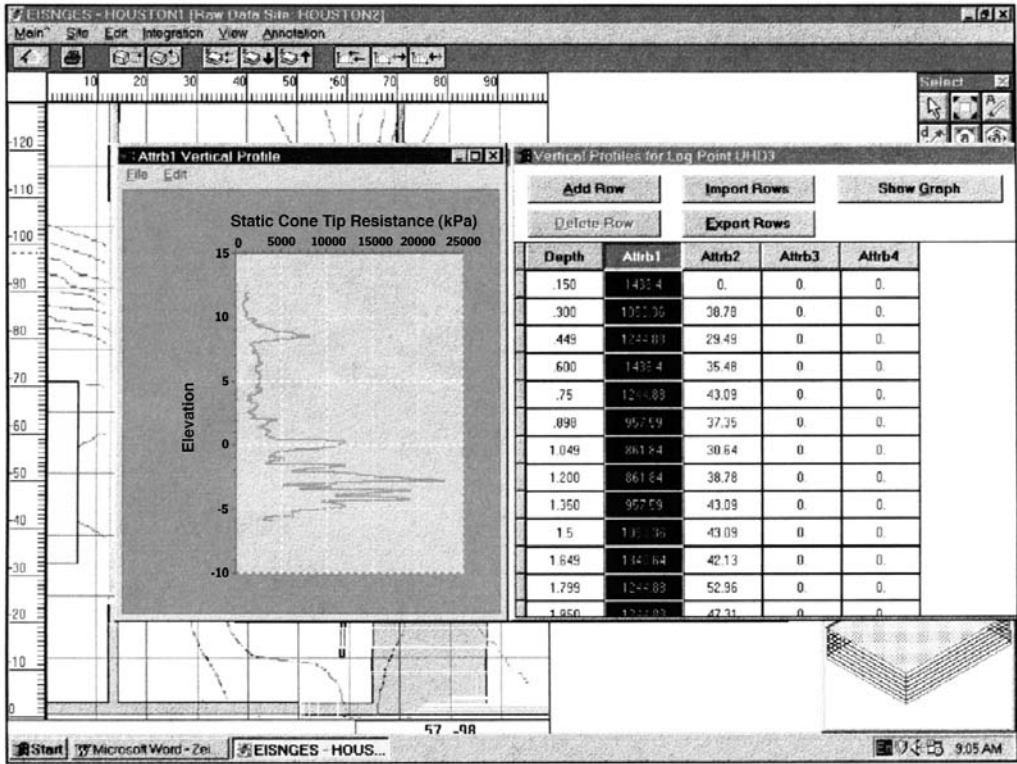
FIGURE 19.14 Viewing and manipulating raw site data.

19.6 Concluding Remarks

This chapter lays the foundation of a new approach in the field of numerical simulation of contaminant fate and migration in the geomedia in the form of a general framework. The general features of this approach are as follows. A general macroengineering framework is presented for the simulation of the interaction of competing natural processes affecting contaminant migration that can be accurately and easily implemented. This framework necessitates a Eulerian reference frame to connect different modeling abstraction layers and *in situ* databases, an inference engine to map the physical characteristics from known *in situ* locations to the simulation reference frame, and a scale operator to coordinate the different simulation algorithms.

These features are integrated in the prototype EIS/GWM platform. Details on specific modules of the platform can be found in the EIS materials listed in the references. Example applications of the macroengineering approach include bioremediation because of its importance in today's cleanup operations. It provides the test bench to simulate the interaction of different modeling components such as geologic and structural features (layered soils, rivers, slurry walls), chemical species (benzene, oxygen), and loadings (injection wells, piezometric head gradients). The same overall framework can also be used to simulate other remediation technologies, individually or in combination. In fact, the macroengineering framework is the *only* way to examine the potential of merging several remediation processes into a very effective combined *in situ* treatment system.

The holistic framework presented here offers the means to understand the delicate balance between competing physicochemical processes that evolve in the three-dimensional soil medium. This framework places particular attention on the scales of the available data and their use by the various simulation algorithms. This approach allows the evaluation of an error estimate at each predictive step of the time

marching algorithms. A correction can then be implemented in the next step depending on the deviation of the interacting variables. Furthermore, this simulation-based approach offers the only way to address the promising avenues of merging different treatment technologies into integrated treatment systems that achieve enhanced productivity and efficiency (treatment time, cost, and clean up level).

Specific steps for an efficient implementation of the macroengineering approach are:

- Better coordinate soil data collection with simulation of interaction processes for prediction needs.
- Hierarchically organize space and time scales of field measurements of interaction phenomena and corresponding media properties and attendant models
- Place more emphasis on the scientific understanding (prediction) of physicochemical processes rather than relying solely on data collection
- Use integrated simulation tools to better identify combinations of treatment processes as they interact with the soil media

References

ZEi, 1994-1997, Environmental Information System (EIS) Graphical Platform, Documentation:
    EIS Installation Guide, Mi-96-G001.
    EIS User's Guide, Mi-96-G002.
    EIS Theoretical Manual, Mi-96-M010.
    EIS Automated Generalized Kriging Procedures, Mi-96-M011.
    EIS-Biorem3D - Modeling Intrinsic Remediation, Mi-96-M012.
    EIS-BioQuick- Quick Model for IR, Mi-96-M013.
    EIS-Raw Data Base, Mi-96-M014.
Goodchild, M. F. 1991. Spatial analysis with GIS: Problems and Prospects. Proceedings, GIS/LIS '91, vol. 1, 40-48. Bethesda: American Congress on Surveying and Mapping, Bethesda, MD.
Egenhofer, M. J. and Frank, A. U. 1992. Object-oriented modeling for GIS. *URISA Journal* 4(2):3-19, London.
Konikow, L. F. and Bredehoeft, J. D. 1978. Computer model of two-dimensional solute transport and dispersion in groundwater, *Techniques of Water Resources Investigations*, Book 7, Chap. C2, U.S. Geological Survey, Reston, Virginia.
Properties of Soil-water, Soil-Gas, Soil-Contaminant under EIS, Mi-96-D001.

20
Groundwater Modeling

Leonard F. Konikow and
Thomas E. Reilly
U.S. Geological Survey

20.1 Introduction

Effective management of groundwater requires the ability to predict subsurface flow and transport of solutes, and the response of fluid and solute flux to changes in natural or human-induced stresses. One popular type of tool that has been evolving since the mid-1960s is the deterministic, distributed-parameter, computer simulation model for analyzing flow and solute-transport in groundwater systems. The development of the computer simulation model has somewhat paralleled the development and increasing availability of faster, larger memory, more capable, yet less expensive computer systems.

The purpose of this chapter is to review the state of the art in deterministic modeling of groundwater flow and transport processes. This chapter, based largely on Konikow (1996), is aimed at practitioners and is intended to describe the types of models that are available and how they may be applied to complex field problems. It will discuss the philosophy and theoretical basis of deterministic modeling, the advantages and limitations of models, the use and misuse of models, how to select a model, and how to calibrate and evaluate a model. However, as this chapter is only a review, it cannot offer comprehensive and in-depth coverage of this complex topic; instead, it guides the reader to references that provide more details.

20.2 Models

The word *model* has so many definitions and is so overused that it is sometimes difficult to discern its meaning (Konikow and Bredehoeft, 1992). A model is perhaps most simply defined as a representation of a real system or process. A *conceptual model* is a hypothesis for how a system or process operates. This hypothesis can be expressed quantitatively as a mathematical model. *Mathematical models* are abstractions that represent processes as equations, physical properties as constants or coefficients in the equations, and measures of state or potential in the system as variables.

Most groundwater models in use today are deterministic mathematical models. *Deterministic models* are based on conservation of mass, momentum, and energy and describe cause and effect relations. The underlying assumption is that given a high degree of understanding of the processes by which stresses on a system produce subsequent responses in that system, the system's response to any set of stresses can be predetermined, even if the magnitude of the new stresses falls outside the range of historically observed stresses.

Deterministic groundwater models generally require the solution of partial differential equations. Exact solutions can often be obtained analytically, but *analytical models* require that the parameters and boundaries be highly idealized. Some deterministic models treat the properties of porous media as lumped parameters (essentially, as a black box), but this precludes the representation of heterogeneous hydraulic properties in the model. Heterogeneity, or variability in aquifer properties, is characteristic of all geologic systems and is now recognized as playing a key role in influencing groundwater flow and solute transport. Thus, it is often preferable to apply distributed-parameter models, which allow the representation of more realistic distributions of system properties. Numerical methods yield approximate solutions to the governing equation (or equations) through the discretization of space and time. Within the discretized problem domain, the variable internal properties, boundaries, and stresses of the system are approximated. Deterministic, distributed-parameter, *numerical models* can relax the rigid idealized conditions of analytical models or lumped-parameter models, and they can therefore be more realistic and flexible for simulating field conditions (if applied properly).

The number and types of equations to be solved are determined by the concepts of the dominant governing processes. The coefficients of the equations are the parameters that are measures of the properties, boundaries, and stresses of the system; the dependent variables of the equations are the measures of the state of the system and are mathematically determined by the solution of the equations. When a numerical algorithm is implemented in a computer code to solve one or more partial differential equations, the resulting computer code can be considered a *generic model.* When the grid dimensions, boundary conditions, and other parameters (such as hydraulic conductivity and storativity), are specified in an application of a generic model to represent a particular geographic area, the resulting computer program is a *site-specific model.* The ability of generic models to solve the governing equations accurately is typically demonstrated by example applications to simplified problems. This does not guarantee a similar level of accuracy when the model is applied to a complex field problem.

If the user of a model is unaware of or ignores the details of the numerical method, including the derivative approximations, the scale of discretization, and the matrix solution techniques, significant errors can be introduced and remain undetected. For example, if the groundwater flow equation is solved iteratively, but the convergence criterion is relatively too coarse, then the numerical solution may converge, but to a poor solution. The inaccuracy of the solution may or may not be reflected in the mass-balance error. The mass-balance error itself may not be readily observed by inexperienced model users. Unrecognized errors in numerical groundwater models are becoming more possible as user-friendly graphic interfaces make it easier for models to be used (and misused). These interfaces effectively place more distance between the modeler and the numerical method that lies at the core of the model.

20.3 Flow and Transport Processes

The process of groundwater flow is generally assumed to be governed by the relations expressed in Darcy's law (see Chapter 2) and the conservation of mass. However, Darcy's law does have limits on its range of applicability, and these limits must be evaluated in any application.

The purpose of a model that simulates solute transport in groundwater is to compute the concentration of a dissolved chemical species in an aquifer at any specified time and place. The theoretical basis for the equation describing solute transport has been well documented in the literature (e.g., Bear, 1979; Domenico and Schwartz, 1990). Reilly et al. (1987) provide a conceptual framework for analyzing and modeling physical solute-transport processes in groundwater. Changes in chemical concentration occur within a dynamic groundwater system primarily due to four distinct processes: (1) advective transport, in which dissolved chemicals are moving with the flowing groundwater; (2) hydrodynamic dispersion, in which molecular and ionic diffusion and small-scale variations in the flow velocity through the porous media cause the paths of dissolved molecules and ions to diverge or spread from the average direction of groundwater flow; (3) fluid sources, where water of one composition is introduced into and mixed with water of a different composition; and (4) reactions, in which some amount of a particular dissolved chemical species may be added to or removed from the groundwater as a result of chemical, biological, and physical reactions in the water or between the water and the solid aquifer materials or other separate liquid phases.

The subsurface environment constitutes a complex, three-dimensional, heterogeneous hydrogeologic setting. This variability strongly influences groundwater flow and transport, and such a reality can be described accurately only through careful hydrogeologic practice in the field. However, regardless of how much data are collected, uncertainty always remains about the properties and boundaries of the groundwater system of interest. Stochastic approaches have resulted in many significant advances in characterizing subsurface heterogeneity and dealing with uncertainty (see Gelhar, 1993).

20.4 Governing Equations

The mathematical equations that describe groundwater flow (see Chapter 3) and transport processes (see Chapters 14 and 15 and the summary in Chapter 2) may be developed from the fundamental principle of conservation of mass of fluid or of solute. Given a representative elementary volume (REV) of porous medium, a general equation for conservation of mass for the volume may be expressed as:

rate of mass inflow – rate of mass outflow + rate of mass production/consumption

$\quad$ = rate of mass accumulation $\hfill$ (1)

This statement of conservation of mass (or continuity equation) may be combined with a mathematical expression of the relevant process to obtain a differential equation that describes flow or transport.

20.4.1 Groundwater Flow Equation

The rate of flow of water through a porous media is related to the properties of the water, the properties of the porous media, and the gradient of the hydraulic head, as represented by Darcy's law, which can be written as:

$$q_i = -K_{ij} \frac{\partial h}{\partial x_j} \tag{2}$$

where q_i is the specific discharge, LT^{-1}; K_{ij} is the hydraulic conductivity of the porous medium (a second-order tensor), LT^{-1}; and h is the hydraulic head, L.

A general form of the equation describing the transient flow of a compressible fluid in a nonhomogeneous anisotropic aquifer may be derived by combining Darcy's law with the continuity equation. A general groundwater flow equation may be written in Cartesian tensor notation as:

$$\frac{\partial}{\partial x_i}\left(K_{ij}\frac{\partial h}{\partial x_i}\right) = S_S\frac{\partial h}{\partial t} + W^* \tag{3}$$

where S_S is the specific storage, L^{-1}; t is time, T; W^* is the volumetric flux per unit volume (positive for outflow and negative for inflow), T^{-1}; and x_i are the Cartesian coordinates, L. The summation convention of Cartesian tensor analysis is implied in Equations (2) and (3). Equation (3) can generally be applied if isothermal conditions prevail, the porous medium only deforms vertically, the volume of individual grains remains constant during deformation, Darcy's law applies (and gradients of hydraulic head are the only driving force), and fluid properties (density and viscosity) are homogeneous and constant. Aquifer properties can vary spatially, and fluid stresses (W^*) can vary in space and time.

If the aquifer is relatively thin compared to its lateral extent, it may be appropriate to assume that groundwater flow is areally two-dimensional. This allows the three-dimensional flow equation to be reduced to the case of two-dimensional areal flow, for which several additional simplifications are possible. The advantages of reducing the dimensionality of the equation include less stringent data requirements, smaller computer memory requirements, and shorter computer execution times to achieve numerical solutions.

An expression similar to Equation (3) may be derived for the two-dimensional areal flow of a homogeneous fluid in a confined aquifer and written as:

$$\frac{\partial}{\partial x_i}\left(T_{ij}\frac{\partial h}{\partial x_i}\right) = S\frac{\partial h}{\partial t} + W \tag{4}$$

where T_{ij} is the transmissivity, L^2T^{-1}; and $T_{ij} = K_{ij}\,b$; b is the saturated thickness of the aquifer, L; S is the storage coefficient (dimensionless); and $W = W^*b$ is the volume flux per unit area, LT^{-1}.

When Equation (4) is applied to an unconfined (water-table) aquifer system, it must be assumed that flow is horizontal and equipotential lines are vertical, that the horizontal hydraulic gradient equals the slope of the water table, and that the storage coefficient is equal to the specific yield (S_y) (Anderson and Woessner, 1992). Note that in an unconfined system, the saturated thickness changes as the water-table elevation (or head) changes. Thus, the transmissivity also can change over space and time (that is, $T_{ij} = K_{ij}b$, where $b(x,y,t) = h - h_b$, and h_b is the elevation of the bottom of the aquifer).

The cross-product terms of the hydraulic conductivity tensor drop out when the coordinate axes are aligned with the principal axes of the tensor; that is, $K_{ij} = 0$ when $i \neq j$. Therefore, the only hydraulic conductivity terms with possible nonzero values are K_{xx} and K_{yy}. Under this assumption, Equation (4) may be simplified to:

$$\frac{\partial}{\partial x}\left(T_{xx}\frac{\partial h}{\partial x}\right) + \frac{\partial}{\partial y}\left(T_{yy}\frac{\partial h}{\partial y}\right) = S\frac{\partial h}{\partial t} + W \tag{5}$$

for two-dimensional flow.

In some field situations, fluid properties such as density and viscosity may vary significantly in space or time. This may occur where water temperature or dissolved-solids concentration changes significantly. When the water properties are heterogeneous and (or) transient, the relations among water levels, hydraulic heads, fluid pressures, and flow velocities are neither simple nor straightforward. In such cases, the flow equation is written and solved in terms of fluid pressures, fluid densities, and the intrinsic permeability of the porous media (see Konikow and Grove, 1977).

20.4.2 Seepage Velocity

The migration and mixing of chemicals dissolved in groundwater will obviously be affected by the velocity of the flowing groundwater. The specific discharge calculated from Equation (2) is sometimes called the Darcy velocity. However, this nomenclature can be misleading because q_i does not actually represent the speed of water movement. Rather, q_i represents a volumetric flux per unit cross-sectional area. Thus, to calculate the actual seepage velocity of groundwater, one must account for the actual cross-sectional area through which flow is occurring, as follows:

$$V_i = \frac{q_i}{\varepsilon} = -\frac{K_{ij}}{\varepsilon}\frac{\partial h}{\partial x_j} \tag{6}$$

where V_i is the seepage velocity (also commonly called average linear velocity or average interstitial velocity), LT^{-1}; and ε is the effective porosity of the porous medium.

20.4.3 Solute Transport Equation

An equation describing the transport and dispersion of a dissolved chemical in flowing groundwater may be derived from the principle of conservation of mass by considering all fluxes into and out of a representative elementary volume (REV), as described by Bear (1979, p. 29). A generalized form of the solute-transport equation is presented by Grove (1976), in which terms are incorporated to represent chemical reactions and solute concentration both in the pore fluid and on the solid surface, as:

$$\frac{\partial(\varepsilon C)}{\partial t} = \frac{\partial}{\partial x_i}\left(\varepsilon D_{ij}\frac{\partial C}{\partial x_j}\right) - \frac{\partial}{\partial x_i}\left(\varepsilon C V_i\right) - C'W^* + CHEM \tag{7}$$

where *CHEM* equals one or more of the following:

$-\rho_b\dfrac{\partial \overline{C}}{\partial t}$ for linear equilibrium controlled sorption or ion-exchange reactions,

$\displaystyle\sum_{k=1}^{s} R_k$ for s chemical rate-controlled reactions, and (or)

$-\lambda\left(\varepsilon C + \rho_b \overline{C}\right)$ for decay,

and where D_{ij} is the coefficient of hydrodynamic dispersion (a second-order tensor), $L^2 T^{-1}$, C' is the concentration of the solute in the source or sink fluid, $\overline{C}$ is the concentration of the species adsorbed on the solid (mass of solute/mass of solid), ρ_b is the bulk density of the sediment, ML^{-3}, R_k is the rate of production of the solute in reaction k, $ML^{-3}T^{-1}$, and λ is the decay constant (equal to ln2/half life), T^{-1} (Grove, 1976).

The first term on the right side of Equation (7) represents the change in concentration due to hydrodynamic dispersion. This expression is analogous to Fick's law describing diffusive flux. This Fickian model assumes that the driving force is the concentration gradient and that the dispersive flux occurs in a direction from higher toward lower concentrations. However, this assumption is not always consistent with field observations and is the subject of much ongoing research and field study (see, for example, Gelhar et al., 1992). The coefficient of hydrodynamic dispersion is defined as the sum of mechanical dispersion and molecular diffusion (Bear, 1979). The mechanical dispersion is a function both of the

intrinsic properties of the porous medium (such as heterogeneities in hydraulic conductivity and porosity) and of the fluid flow. Molecular diffusion in a porous medium will differ from that in free water because of the effects of tortuous paths of fluid connectivity in porous media. These relations are commonly expressed as:

$$D_{ij} = \alpha_{ijmn} \frac{V_m V_n}{|V|} + D_m \qquad i, j, m, n = 1, 2, 3 \qquad (8)$$

where α_{ijmn} is the dispersivity of the porous medium (a fourth-order tensor), L; V_m and V_n are the components of the flow velocity of the fluid in the m and n directions, respectively, LT^{-1}; D_m is the effective coefficient of molecular diffusion, $L^2 T^{-1}$; and $|V|$ is the magnitude of the velocity vector, LT^{-1}, defined as $|V| = \sqrt{V_x^2 + V_y^2 + V_z^2}$ (Bear, 1979; Domenico and Schwartz, 1990). The dispersivity of an isotropic porous medium can be defined by two constants. These are the longitudinal dispersivity of the medium, α_L, and the transverse dispersivity of the medium, α_T. These are related to the longitudinal and transverse dispersion coefficients by $D_L = \alpha_L |V|$ and $D_T = \alpha_T |V|$. Most documented applications of transport models to groundwater problems have been based on this conventional formulation, even for cases in which the hydraulic conductivity is assumed to be anisotropic (despite the conceptual inconsistency). However, some models (for example, Voss, 1984) incorporate an additional level of complexity by allowing α_L and (or) α_T to vary with direction.

Although conventional theory holds that α_L is generally an intrinsic property of the aquifer, it is found in practice to be dependent on and proportional to the scale of the measurement. Most reported values of α_L fall in a range from 0.01 to 1.0 times the scale of the measurement, although the ratio of α_L to scale of measurement tends to decrease at larger scales (see Anderson, 1984; Gelhar et al., 1992). Field-scale dispersion (commonly called macrodispersion) results from large-scale spatial variations in hydraulic properties. Consequently, the use of relatively large values of dispersivity together with uniform hydraulic properties (K_{ij} and ε) is inappropriate for describing transport in geological systems (Smith and Schwartz, 1980). Part of the scale dependence of dispersivity may be explained as an artifact of the models used, in that a scaling up of dispersivity will occur whenever an (n-1)-dimensional model is calibrated or used to describe an n-dimensional system (Domenico and Robbins, 1984). Furthermore, if a model applied to a system having variable hydraulic conductivity uses mean values and thereby does not explicitly represent the variability, the model calibration will likely yield values for the dispersivity coefficients that are larger than would be measured locally in the field area. Similarly, representing a transient flow field by a mean steady-state flow field, as is commonly done, inherently ignores some of the variability in velocity and must be compensated for by using increased values of dispersivity (primarily transverse dispersivity) (Goode and Konikow, 1990). Overall, the more accurately a model can represent or simulate the true velocity distribution in space and time, the less of a problem will be the uncertainty concerning representation of dispersion processes.

The mathematical solute-transport model requires at least two partial differential equations. One is the equation of flow, from which groundwater flow velocities are obtained, and the second is the solute-transport equation, whose solution gives the chemical concentration in groundwater. If the properties of the water are affected significantly by changes in solute concentration, as in a saltwater intrusion problem, then the flow and transport equations should be solved simultaneously (or at least iteratively). If the properties of the water remain constant, then the flow and transport equations can be decoupled and solved sequentially, which is simpler numerically.

20.5 Numerical Methods To Solve Equations

The partial differential equations describing groundwater flow and transport can be solved mathematically using either analytical solutions or numerical solutions. The advantages of an analytical solution, when it is possible to apply one, are that it usually provides an exact solution to the governing equation and is often relatively simple and efficient to use. Many analytical solutions have been developed for the flow equation; however, most applications are limited to well hydraulics problems involving radial symmetry. The familiar Theis type curve represents the solution of one such analytical model. Analytical solutions are also available to solve the solute-transport equation (e.g., Bear, 1979; Javandel et al., 1984; Wexler, 1992). In general, obtaining the exact analytical solution to the partial differential equation requires that the properties and boundaries of the flow system be highly and perhaps unrealistically idealized. For simulating most field problems, the mathematical benefits of obtaining an exact analytical solution are probably outweighed by the errors introduced by the simplifying assumptions about the complex field environment that are required to apply the analytical approach.

Alternatively, for problems where the simplified analytical models no longer describe the physics of the situation, the partial differential equations can be approximated numerically. In so doing, the continuous variables are replaced with discrete variables that are defined at grid blocks or nodes. Thus, the continuous differential equation, which defines hydraulic head or solute concentration everywhere in the system, is replaced by a finite number of algebraic equations that defines the hydraulic head or concentration at specific points. This system of algebraic equations generally is solved using matrix techniques. This approach constitutes a numerical model.

Two major classes of numerical methods have come to be well accepted for solving the groundwater flow equation. These are the finite-difference methods and the finite-element methods. Each of these two major classes of numerical methods includes a variety of subclasses and implementation alternatives. Comprehensive treatments of the application of these numerical methods to groundwater problems are presented by Remson et al. (1971) and Wang and Anderson (1982). Both of these numerical approaches require that the area of interest be subdivided by a grid into a number of smaller subareas (cells or elements) that are associated with nodal points (either at the centers or peripheries of the subareas).

In addition to finite-difference and finite-element methods, boundary integral equation methods and analytical element methods can also be applied to solve the flow equation (for example, see Haitjema, 1995). Their main advantage is that, for homogeneous regions, they can provide precise solutions without discretization. Thus, if a system's heterogeneity can be adequately represented by using only a few very large elements, the methods can be very efficient in terms of computer time. If heterogeneities are such that a large number of elements are required to describe them adequately, then finite-difference or finite-element methods may be preferable. To date, finite-difference and finite-element methods have been more widely used than other numerical methods in simulating groundwater flow problems.

Finite-difference methods approximate the first derivatives in the partial differential equations as difference quotients (the differences between values of the independent variable at adjacent nodes with respect to the distance between the nodes, and at two successive time levels with respect to the duration of the time-step increment). Finite-element methods use assumed functions of the dependent variable and parameters to evaluate equivalent integral formulations of the partial differential equations. Huyakorn and Pinder (1983) present a comprehensive analysis of the application of finite-element methods to groundwater problems. In both numerical approaches, the discretization of the space and time dimensions allows the continuous boundary-value problem for the solution of the partial differential equation to be reduced to the simultaneous solution of a set of algebraic equations. These equations can then be solved using either iterative or direct matrix methods.

Each approach has advantages and disadvantages, but there are very few groundwater problems for which either is clearly superior. In general, the finite-difference methods are simpler conceptually and mathematically, and are easier to program. They are typically keyed to a relatively simple, rectangular

grid, which also eases data entry. Finite-element methods generally require the use of more sophisticated mathematics but, for some problems, may be more accurate numerically than standard finite-difference methods. A major advantage of the finite-element methods is the flexibility of the finite-element grid, which allows a close spatial approximation of irregular boundaries of the aquifer and (or) of parameter zones within the aquifer when they are considered. However, the construction and specification of an input data set are much more difficult for an irregular finite-element grid than for a regular rectangular finite-difference grid. Thus, the use of a model preprocessor, which includes a mesh generator and a scheme to number the nodes and elements of the mesh and to specify the spatial coordinates of each node, is recommended. Figure 20.1 illustrates a hypothetical aquifer system, which has impermeable boundaries and a well field (Figure 20.1A), which has been discretized using finite-difference (Figure 20.1B) and finite-element (Figure 20.1C) grids. Figures 20.1B and 20.1C illustrate conceptually how their respective grids can be adjusted to use a finer mesh spacing in selected areas of interest. The rectangular finite-difference grid approximates the aquifer boundaries in a stepwise manner, resulting in some nodes or cells outside the aquifer, whereas sides of the triangular elements of the finite-element grid can closely follow the outer boundary using a minimal number of nodes.
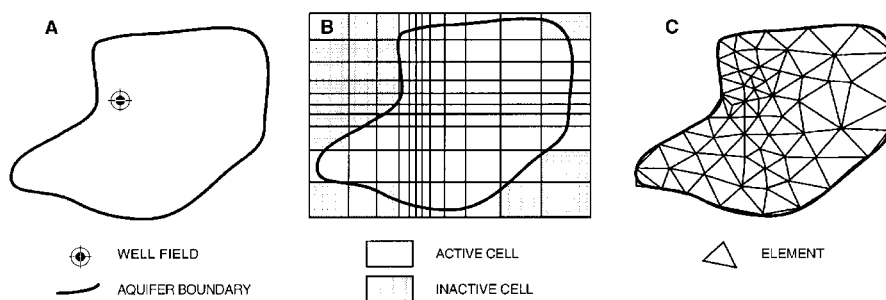
FIGURE 20.1 Hypothetical application to (A) an irregularly bounded aquifer of (B) finite-difference and (C) finite-element grids. (From Konikow, L. F. 1996. Numerical models of groundwater flow and transport, in *Manual on Mathematical Models in Isotope Hydrogeology.* International Atomic Energy Agency, Vienna.)

The solute-transport equation is more difficult to solve numerically than the groundwater flow equation, largely because the mathematical properties of the transport equation vary depending upon which terms in the equation are dominant in a particular situation. When solute transport is dominated by advective transport, as is common in many field problems, then Equation (7) approximates a hyperbolic type of equation (similar to equations describing the propagation of a wave or of a shock front). But if a system is dominated by dispersive fluxes, such as might occur where fluid velocities are relatively low and aquifer dispersivities are relatively high, then Equation (7) becomes more parabolic in nature (similar to the transient groundwater flow equation).

The numerical methods that work best for parabolic partial differential equations are not best for solving hyperbolic equations, and vice versa. Thus, no one numerical method or simulation model will be ideal for the entire spectrum of groundwater transport problems likely to be encountered in the field. Further compounding this difficulty is the fact that in the field, the seepage velocity of groundwater is highly variable, even if aquifer properties are relatively homogeneous because of the effects of complex boundary conditions. Thus, in low permeability zones or near stagnation points, the velocity may be close to zero and the transport processes will be dominated by dispersion processes; in high permeability zones or near stress points (such as pumping wells), the velocity may be several meters per day and the transport processes will be advection dominated. In other words, for the same system, the governing equation may be more hyperbolic in one area (or at one time) and more parabolic in another area (or at another time). Therefore, regardless of which numerical method is chosen as the basis for a simulation model, it will not be ideal or optimal over the entire domain of the problem, and significant numerical

errors may be introduced somewhere in the solution. The transport modeling effort must recognize this inherent difficulty and strive to minimize and control the numerical errors.

Additional complications arise when the solutes of interest are reactive. The reaction terms included in Equation (7) are mathematically simple ones. They do not necessarily represent the true complexities of many reactions. Also, particularly difficult numerical problems arise when reaction terms are highly nonlinear, or if the concentration of the solute of interest is strongly dependent on the concentration of other chemical constituents. In reality, isotherms may not be linear and may not be equilibrium controlled. For field problems in which reactions significantly affect solute concentrations, simulation accuracy is less limited by mathematical constraints than by data constraints. That is, the types and rates of reactions for the specific solutes and minerals in the particular groundwater system of interest are rarely known and require an extensive amount of data to assess accurately.

Finite-difference and finite-element methods also can be applied to solve the transport equation, particularly when dispersive transport is large compared to advective transport. However, numerical errors, such as numerical dispersion and oscillations, may be significant for some problems. The numerical errors can generally be reduced by using a finer discretization (either shorter time steps or finer spatial grid). An example of a documented three-dimensional, transient, finite-difference model that simultaneously solves the fluid pressure, energy-transport, and solute-transport equations for nonhomogeneous miscible fluids is HST3D (Kipp, 1987). An example of a two-dimensional finite-element transport model is SUTRA, documented by Voss (1984).

Although finite-difference and finite-element models are commonly applied to transport problems, other types of numerical methods have also been applied to transport problems, including the method of characteristics, random walk, Eulerian-Lagrangian methods, and adaptive grid methods. All of these methods have the ability to track sharp fronts accurately with a minimum of numerical dispersion. Documented models based on variants of these approaches include Konikow and Bredehoeft (1978), Sanford and Konikow (1985), Prickett et al. (1981), and Zheng (1990).

No single one of the standard numerical methods is ideal for a wide range of transport problems and conditions. Thus, there is currently still much research on developing better mixed or adaptive methods that aim to minimize numerical errors and combine the best features of alternative standard numerical approaches.

20.5.1 Basics of Finite-Difference Methods

The partial differential equations describing the flow and transport processes in groundwater include terms representing derivatives of continuous variables in space and time. Finite-difference methods are based on the approximation of these derivatives (or slopes of curves) by discrete linear changes over discrete intervals of space or time. If the intervals are sufficiently small, then all of the linear increments will represent a good approximation of the true curvilinear surface or hydrograph.

If we consider the observation wells in a confined aquifer, as illustrated in Figure 20.2A, Bennett (1976) shows that a reasonable approximation for the derivative of head, $\partial h/\partial x$, at a point (d) midway between wells 1 and 0 is:

$$\left(\frac{\partial h}{\partial x}\right)_d \approx \frac{h_0 - h_1}{\Delta x} \tag{9}$$

Note that the observation wells are spaced an equal distance apart. Similarly, a reasonable approximation for the second derivative, $\partial^2 h/\partial x^2$, at point 0 (the location of the center well) can be given as:

$$\left(\frac{\partial^2 h}{\partial x^2}\right)_0 \approx \frac{\left(\dfrac{\partial h}{\partial x}\right)_e - \left(\dfrac{\partial h}{\partial x}\right)_d}{\Delta x} \approx \frac{\dfrac{h_2 - h_0}{\Delta x} - \dfrac{h_0 - h_1}{\Delta x}}{\Delta x} = \frac{h_1 + h_2 - 2h_0}{\left(\Delta x\right)^2} \tag{10}$$
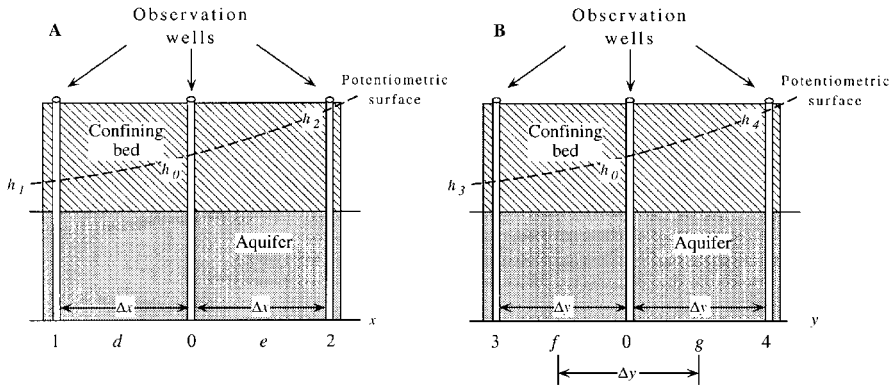
FIGURE 20.2 Schematic cross section through confined aquifer to illustrate numerical approximation to derivatives of head, (A) ∂h/∂x and (B) ∂h/∂y. (Adapted from Bennett, G. D. 1976. *Introduction to Ground-Water Hydraulics: A Programmed Text for Self-Instruction.* Techniques of Water-Res. Invests. of the U.S. Geol. Survey, Book 3, Ch. B2.)

If we also consider wells 3 and 4 shown in Figure 20.2B, located on a line parallel to the y-axis, we can similarly approximate $\partial^2 h/\partial y^2$ at point 0 (the same point 0 as in Figure 20.2A) as (Bennett, 1976):

$$\left(\frac{\partial^2 h}{\partial y^2}\right)_0 \approx \frac{h_3 + h_4 - 2h_0}{\left(\Delta y\right)^2} \tag{11}$$

If the spacing of the wells in Figure 20.2B is uniform (that is, $\Delta x = \Delta y = a$), then we can develop the following approximation:

$$\frac{\partial^2 h}{\partial x^2} + \frac{\partial^2 h}{\partial y^2} \approx \frac{h_1 + h_2 + h_3 + h_4 - 4h_0}{a^2} \tag{12}$$

These approximations can also be obtained through the use of Taylor series expansions. A certain error is involved in approximating the derivatives by finite-differences, but this error will generally decrease as a (or Δx and Δy) is given smaller and smaller values. This error is called a "truncation error" because the replacement of a derivative by a difference quotient is equivalent to using a truncated Taylor series, so that the exact solution of a difference equation differs from the solution of the corresponding differential equation (Peaceman, 1977). Also, it may not be possible to achieve an "exact" solution of the difference equation because of limits of precision in storing numbers in a digital computer. In solving a large set of difference equations, many arithmetic operations are performed, and round-off errors may sometimes accumulate.

Next consider the construction of a rectangular finite-difference grid. Two possible modes of grid construction are illustrated in two dimensions in Figures 20.3A and 3B. In Figure 20.3A, the calculation points (or nodes) are located at the centers of the blocks (or cells) formed by the grid lines. This type of grid is commonly called a block-centered grid. In the second type (Figure 20.3B), the nodes are considered to be located at the intersections of the grid lines. This type has been variously called a point-centered, node-centered, mesh-centered, or lattice-centered grid. Although there is no overall inherent advantage of one type over the other, there will be some operational differences between the two approaches in the treatment of boundaries and in areas of influence around nodes. Most, but not all, finite-difference groundwater models are based on the use of block-centered grids. Double indexing is normally used to identify functions and variables within the two-dimensional region. For example, $h_{i,j}$
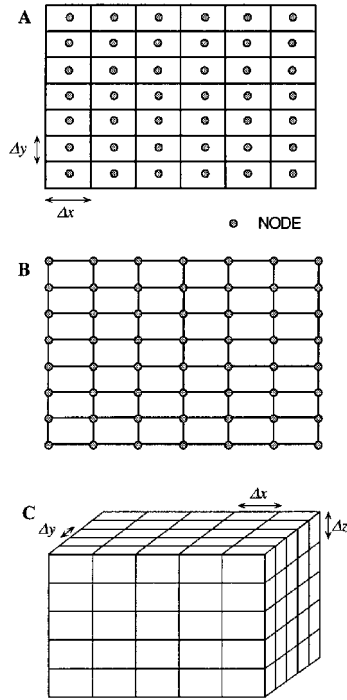
FIGURE 20.3 Examples of finite-difference grids: (A) two-dimensional block-centered grid, (B) two-dimensional node-centered grid, and (C) three-dimensional block-centered grid. (A and B from Konikow, L. F. 1996. Numerical models of groundwater flow and transport, in *Manual on Mathematical Models in Isotope Hydrogeology*. International Atomic Energy Agency, Vienna. C from Konikow, L. F., Goode, D. J., and Hornberger, G. Z. 1996. A three-dimensional method-of-characteristics solute-transport model (MOC3D). U.S. Geol. Survey Water-Res. Inv. Rept. 96-4267).

is the head at node i,j, where i and j are the row and column locations in the finite-difference grid. This procedure is easily extended to three dimensions, as illustrated in Figure 20.3C. Here the vertical dimension (or z-direction) is indexed by the subscript k and $h_{i,j,k}$ would represent the head at node i,j,k.

We must also consider the discretization of time, which may be viewed as another dimension, and hence represented by another index. If we consider a representative segment of a hydrograph (see Figure 20.4), in which head is plotted against time for a transient flow system, n is the index or subscript used to denote the time at which a given head value is observed. The slope of the hydrograph at any point is the derivative of head with respect to time, and it can be approximated as $\partial h/\partial t = \Delta h/\Delta t$. In terms of the heads calculated at specific time increments (or time nodes), the slope of the hydrograph at time n can be approximated by:

$$\left(\frac{\partial h}{\partial t}\right)_{n\Delta t} \approx \frac{h_{n+1} - h_n}{\Delta t} \tag{13}$$

or

$$\left(\frac{\partial h}{\partial t}\right)_{n\Delta t} \approx \frac{h_n - h_{n-1}}{\Delta t} \tag{14}$$

We are calculating the derivative at t = nΔt in Equation (13) by taking a "forward difference" from time *n* to time *n*+1, and by taking a "backward difference" in Equation (14). In terms of solving the
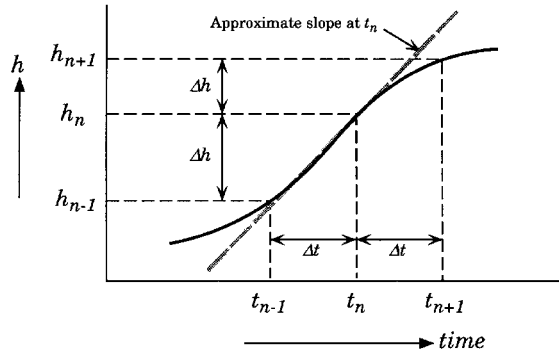
FIGURE 20.4 Part of a hydrograph showing that the derivative (or slope, $\partial h/\partial t$) at time node t_n may be approximated by $\Delta h/\Delta t$. (From Konikow, L. F. 1996. Numerical models of groundwater flow and transport, in *Manual on Mathematical Models in Isotope Hydrogeology*. International Atomic Energy Agency, Vienna.)
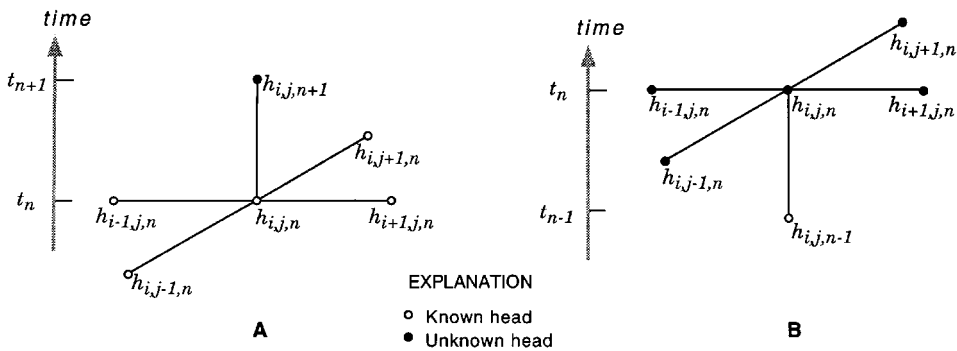
FIGURE 20.5 Grid stencil showing discretization of time at node (i,j) in two-dimensional finite-difference grid: (A) explicit (forward-difference) formulation and (B) implicit (backward-difference) formulation. (From Konikow, L. F. 1996. Numerical models of groundwater flow and transport, in *Manual on Mathematical Models in Isotope Hydrogeology*. International Atomic Energy Agency, Vienna.)

groundwater flow equation for a node (*i,j*) of a finite-difference grid, we have to consider heads at five nodes and at two time levels, as illustrated in Figure 20.5. In Figure 20.5A, we have expressed the spatial derivatives of head at time level n, where all values are known, and the time derivative as a forward difference to the unknown head at time step n+1. Then for every node of the grid we will have a separate difference equation, each of which contains only one unknown variable. Thus, these equations can be solved explicitly. Explicit finite-difference equations are thus simple and straightforward to solve, but they may have associated stability criteria. That is, if time increments are too large, small numerical errors or perturbations may propagate into larger errors at later stages of the computations.

In Figure 20.5B, we have expressed the time derivative as a backward difference from the heads at time level n, which are thereby the unknown heads, whereas the heads at the previous time level, n-1, are known (either from specified initial conditions for the first time step or from subsequent solutions at later time steps). The spatial derivatives of head are written at time level n, where all values are unknown, so for every node of the grid we will have one difference equation that contains five unknowns, which cannot be solved directly. However, for the entire grid, which contains N nodes, we would have a system of N equations containing a total of N unknowns. Such a system of simultaneous equations, together with specified boundary conditions, can be solved implicitly. Although implicit solutions are more complicated, they also have the advantage of generally being unconditionally stable. This implies that a solution will be obtained, although not necessarily that the estimate of the derivative that is calculated

will be accurate, if the time steps are large relative to the rate of change of head. Most available groundwater flow models solve an implicit finite-difference approximation to the flow equation.

We may next consider a two-dimensional groundwater flow equation for a heterogeneous, anisotropic aquifer (Equation [5]), in which the coordinate system is aligned with the major axes of the transmissivity tensor. This may be approximated by the following finite-difference equation for representative node (i,j) as:

$$
T_{xx\left[i-\frac{1}{2}j\right]}\left(\frac{h_{i-1,j,n}-h_{i,j,n}}{\left(\varDelta x\right)^2}\right)+T_{xx\left[i+\frac{1}{2}j\right]}\left(\frac{h_{i+1,j,n}-h_{i,j,n}}{\left(\varDelta x\right)^2}\right)+T_{yy\left[i,j-\frac{1}{2}\right]}\left(\frac{h_{i,j-1,n}-h_{i,j,n}}{\left(\varDelta y\right)^2}\right)
$$

$$
+T_{yy\left[i,j+\frac{1}{2}\right]}\left(\frac{h_{i,j+1,n}-h_{i,j,n}}{\left(\varDelta y\right)^2}\right)=S\left(\frac{h_{i,j,n}-h_{i,j,n-1}}{\varDelta t}\right)-\frac{q_{i,j}}{\varDelta x \varDelta y}-\frac{K_z}{m}\left(H_{s[i,j]}-h_{i,j,n}\right)
$$

(15)

where $q_{i,j}$ is the volumetric rate of withdrawal (negative in sign) or recharge (positive) at the i,j node, L^3T^{-1}. This formulation inherently assumes that any stresses, such as represented by $q_{i,j}$, are applied over the entire surface area of cell i,j rather than at a point (or at node i,j). This implies that if a pumping well is represented at node i,j, then the head will be calculated as if it were being withdrawn from a well that had a horizontal surface area for the borehole equal to $\varDelta x \varDelta y$ rather than its actual value. In Equation (15), the transmissivity terms represent the harmonic means of the transmissivity of the two adjacent cells. The harmonic mean can be shown to be appropriate and consistent with the assumption that transmissivity is constant and uniform within each cell but may be different between cells. Other types of means for interblock transmissivity may be more appropriate for other assumptions about the transmissivity distribution, such as smoothly varying transmissivity (Goode and Appel, 1992).

20.5.2 Basics of Finite-Element Methods

The finite-element method (FEM) is a numerical analysis technique for obtaining approximate solutions to a wide variety of problems in physics and engineering. The method was originally applied to structural mechanics but is now used in all fields of continuum mechanics. Huebner (1975) describes four different approaches to formulate the finite-element method for a problem, which are: the direct approach, the variational approach, the weighted residual approach, and the energy balance approach. In groundwater problems, the approach frequently used is either the weighted residual or variational approach.

The finite-element method (FEM) uses a concept of "piecewise approximation." The domain of the problem, that is the extent of the aquifer to be simulated, is divided into a set of elements or pieces. In theory, the elements can be of different shapes and sizes. Most FEM computer programs use one shape element, most commonly either triangular or quadrilateral elements. In the groundwater model MODFE (Torak, 1993; Cooley, 1992) triangular elements are used, whereas in the groundwater model SUTRA (Voss, 1984) quadrilateral elements are used. Point values of the dependent variable (for example, head, pressure, or concentration) are calculated at nodes, which are the corners or vertices of the elements, and a simple equation is used to describe the value of the dependent variable within the element. This simple equation is called a basis function and each node that is part of an element has an associated basis function. The simplest basis functions that are usually used are linear functions. The solution to the differential equation for flow (Equation [3]) or transport (Equation [7]) is approximated by a set of elements in which the dependent variable only varies linearly within the element, but the entire set of elements approximates the complex distribution of head or concentration. Figure 20.6 shows the approximate modeled hydraulic head distribution (Figure 20.6C) comprised of a set of triangular elements (Figure 20.6A) having a linear approximation of head variation within each element (Figure 20.6B).
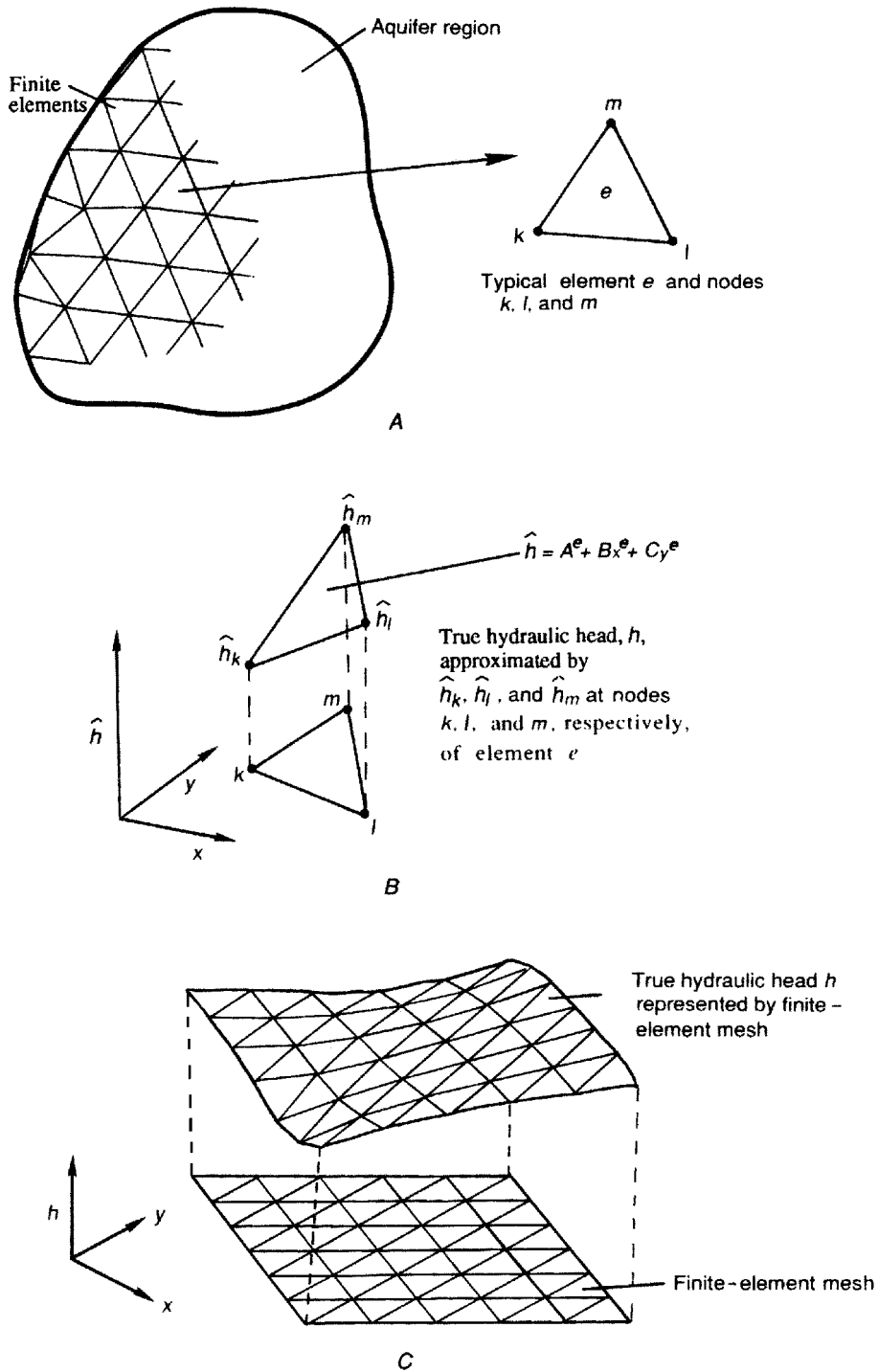
FIGURE 20.6 Diagram showing (A) aquifer region partially subdivided by finite elements and typical element e, (B) finite-element representation of hydraulic head $\hat{h}$, and (C) finite-element mesh configuration for approximating true hydraulic head. (From Torak, L. J. 1993. A modular finite-element model (MODFE) for areal and axisymmetric ground-water-flow problems, Part 1: Model description and user's manual. Techniques of Water-Res. Invests. of the U.S. Geol. Survey, Book 6, Ch. A3.)

In the method of weighted residuals, the piecewise continuous surface is obtained by minimizing the difference between the approximate surface and the continuous surface. The method of weighted residuals is summarized by Huyakorn and Pinder (1983, p. 39) as follows. Any differential equation L(h), such as the steady-state form of Equation (3) (the groundwater flow equation) can be written:

$$L(h) = 0 \qquad (16)$$

over the domain of the problem R. The first step in obtaining the approximate solution is to define the approximate solution as the sum of all the simple basis functions as:

$$\hat{h} = \sum_{i=1}^{n} N_i Z_i \qquad (17)$$

where $\hat{h}$ is the approximate solution, n is the number of linearly independent basis functions, N_i are the linearly independent basis functions defined over the entire domain, and Z_i are the unknown coefficients to be determined (there is one coefficient for each node in the finite-element mesh). The trial function $\hat{h}$ is an approximation, so that when it is substituted into Equation (16) there will be some error, ξ, defined as:

$$\xi = L(\hat{h}) \qquad (18)$$

The method of weighted residuals determines the unknown coefficients by minimizing the error. This is accomplished by weighting the error, integrating the error, and setting the error equal to zero over the entire domain. A weighting function, W_i, can be specified for each basis function and the resulting integration is:

$$\int_R W_i \xi \, dR = \int_R W_i L(\hat{h}) \, dR = 0 \qquad i = 1, 2 \ldots n \qquad (19)$$

Equation (17) is substituted into Equation (19), and weighting functions are specified. There are then n equations and n unknowns. The selection of the weighting functions and the simplification of the integral in Equation (19) into a linear algebraic equation is mathematically straightforward, but not intuitive. In the Galerkin method, the weighting functions are chosen to be identical to the basis functions, and Equation (19) is simplified by using integration by parts. Because the basis functions and weighting functions are defined to be of a specific algebraic form (for example, linear basis functions), the modified integral is straightforward to solve and becomes a set of n simultaneous algebraic equations.

After Equation (19) is mathematically evaluated into a set of n simultaneous equations, they are solved using matrix solution techniques for the n unknown coefficients Z_i, and the approximate solution $\hat{h}$ is determined at each node. The time derivative is frequently approximated by finite differences as discussed in the previous section. Huyakorn and Pinder (1983), Huebner (1975), Zienkiewicz (1971), Wang and Anderson (1982), and Cooley (1992) provide more comprehensive explanations of the method.

20.5.3 Basics of Method-of-Characteristics Methods

The method of characteristics was developed to solve hyperbolic differential equations (advectively dominated transport equations). A major advantage is that the method minimizes numerical dispersion

(Reddell and Sunada, 1970; Garder et al., 1964; Zheng and Bennett, 1995). The approach taken by the method of characteristics is not to solve Equation (7) directly, but rather to solve an equivalent system of ordinary differential equations. A form of Equation (7), accounting for equilibrium-controlled sorption or exchange and first-order irreversible rate reactions, can be further modified for improved compatibility with this method by expanding the advection term, substituting relations from Darcy's law and the flow equation, and rearranging terms to obtain:

$$\frac{\partial C}{\partial t} = \frac{1}{R_f}\frac{\partial}{\partial x_i}\left(D_{ij}\frac{\partial C}{\partial x_j}\right) - \frac{V_i}{R_f}\frac{\partial C}{\partial x_i} + \frac{W^*\left(C - C'\right)}{\varepsilon R_f} - \lambda C \tag{20}$$

where R_f is defined as a dimensionless retardation factor, $R_f = 1 + \dfrac{\rho_b K_d}{\varepsilon}$, and K_d is the distribution coefficient, $L^3 M^{-1}$. If we consider the material derivative of concentration with respect to time, dC/dt, as describing the change in concentration of a parcel of water moving at the seepage velocity of water, it may be defined for a two-dimensional system as:

$$\frac{dC}{dt} = \frac{\partial C}{\partial t} + \frac{\partial C}{\partial x}\frac{dx}{dt} + \frac{\partial C}{\partial y}\frac{dy}{dt} \tag{21}$$

The second and third terms on the right side include the material derivatives of position, which are defined by the velocity in the x and y directions. We then have:

$$\frac{dx}{dt} = \frac{V_x}{R_f} \tag{22}$$

$$\frac{dy}{dt} = \frac{V_y}{R_f} \tag{23}$$

and

$$\frac{dC}{dt} = \frac{1}{R_f}\frac{\partial}{\partial x_i}\left(D_{ij}\frac{\partial C}{\partial x_j}\right) + \frac{W^*\left(C - C'\right)}{\varepsilon R_f} - \lambda C \tag{24}$$

The solutions of the system of equations comprising Equations (22) through (24) may be given as $x = x(t)$, $y = y(t)$, and $C = C(t)$, and are called the characteristic curves of Equation (20). Given solutions to Equations (22) through (24), a solution to the partial differential equation may be obtained by following the characteristic curves, which are defined by the particle pathlines. This may be accomplished by introducing a set of moving points (or reference particles) that can be traced within the stationary coordinates of a finite-difference grid. Each particle corresponds to one characteristic curve, and values of x, y, and C are obtained as functions of t for each characteristic (Garder et al., 1964). Each point has a concentration and position associated with it and is moved through the flow field in proportion to the flow velocity at its location (see Figure 20.7). The concentrations at the nodes of the fixed finite-difference grid may then be estimated as an arithmetic or weighted mean of the concentrations of all particles contained within the cell area for that node.
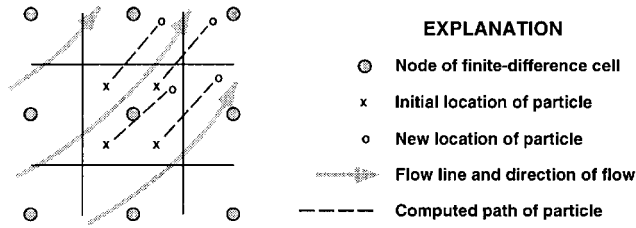
FIGURE 20.7 Part of a hypothetical finite-difference grid showing relation of flow field to movement of points (or particles) in method-of-characteristics model for simulating solute transport. (Adapted from Konikow, L. F. and Bredehoeft, J. D. 1978. *Computer Model of Two-Dimensional Solute Transport and Dispersion in Ground Water.* Techniques of Water-Res. Invests. of the U.S. Geol. Survey, Book 7, Ch. C2.)

20.5.4 Matrix Solution Techniques

As indicated, the finite-difference and finite-element approximations lead to an algebraic equation for each node point. The set of algebraic equations may be solved numerically by one of two basic methods: direct or iterative. In direct methods, a sequence of operations is performed only once to solve the matrix equation, providing a solution that is exact, except for machine round-off error. Iterative methods arrive at a solution by a process of successive approximation. They involve making an initial guess at the solution, then improving this guess by some iterative process until an error criterion is satisfied. Therefore, in these techniques, convergence and the rate of convergence are of concern.

Direct methods can be further subdivided into: (1) solution by determinants, (2) solution by successive elimination of the unknowns, and (3) solution by matrix inversion. Direct methods have two main disadvantages. The first problem is one of computer resource requirements, including large storage (memory) requirements and long computation times for large problems. The matrix is sparse (contains many zero values) and to minimize computational effort, several techniques have been proposed. However, for finite-difference and finite-element methods, storage requirements may still prove to be unavoidably large for three-dimensional problems. The second problem with direct methods is round-off error. Because many arithmetic operations are performed, round-off errors can accumulate for certain types of matrices.

Iterative schemes avoid the need for storing large matrices, which make them attractive for solving problems with many unknowns. Numerous schemes have been developed; a few of the more commonly used ones include successive over-relaxation methods, iterative alternating-direction implicit procedure, and the strongly implicit procedure.

Because iterative methods start with an initial estimate for the solution, the efficiency of the method depends somewhat on this initial guess. To speed up the iterative process, relaxation and acceleration factors are used. Unfortunately, the definition of best values for these factors commonly is problem dependent. In addition, iterative approaches require that an error tolerance be specified to stop the iterative process. An optimal value for the tolerance, which is used to evaluate when the iterative calculations have converged on a solution, may also be problem dependent. If the tolerance is set too large, then the iterations may stop before adequate numerical accuracy is achieved. If the tolerance is set too small, then the iterative process may consume excessive computational resources in striving for numerical precision that may be orders of magnitude smaller than the precision of the field data, or the iterative process may even fail to converge.

More recently, a semi-iterative method, or class of methods, known as conjugate-gradient methods, has gained popularity. One advantage of the conjugate-gradient method is that it does not require the use or specification of iteration parameters, thereby eliminating this partly subjective procedure.

20.5.5 Boundary and Initial Conditions

To obtain a unique solution of a partial differential equation corresponding to a given physical process, additional information about the physical state of the process is required. This information is supplied by boundary and initial conditions. For steady-state problems, only boundary conditions are required, whereas for transient problems, boundary and initial conditions must be specified.

Mathematically, the boundary conditions include the geometry of the boundary and the values of the dependent variable or its derivative normal to the boundary. In physical terms, for groundwater model applications, the boundary conditions are generally of three types: (1) specified value (head or concentration), (2) specified flux (corresponding to a specified gradient of head or concentration), or (3) value-dependent flux (or mixed boundary condition, in which the flux across a boundary is related to both the normal derivative and the value) (Mercer and Faust, 1981; Franke et al., 1987). The third type of boundary condition might be used, for example, to represent leakage or exchange between a stream and an adjacent aquifer, in which the leakage may change over time as the head in the aquifer changes, even though the head in the stream might remain fixed. A no-flow boundary is a special case of the second type of boundary condition. The types of boundaries appropriate to a particular field problem require careful consideration.

The initial conditions are simply the values of the dependent variable specified everywhere inside the boundary at the start of the simulation. Normally, the initial conditions are specified to be a steady-state solution. If, however, initial conditions are specified so that transient flow is occurring in the system at the start of the simulation, it should be recognized that heads will change during the simulation, not only in response to the new pumping stress, but also due to the initial conditions (Franke et al., 1987).

20.6 Model Design, Development, and Application

The first step in model design and application is to define the nature of the problem and the purpose of the model. Although this may seem obvious, it is an important first step that is sometimes overlooked in a hasty effort to take action. This step is closely linked with the formulation of a conceptual model, which again is required prior to development of a mathematical model. A possible outcome of such a preliminary assessment might even be that a deterministic simulation model is not needed. In formulating a conceptual model, the analyst must evaluate which processes are significant in the system being investigated for the particular problem at hand. Some processes may be important to consider at one scale of study, but negligible or irrelevant at another scale of investigation. The analyst must similarly decide on the appropriate dimensionality for the numerical model. Good judgment is required to evaluate and balance the trade-offs between accuracy and cost, with respect to model development, model use, and data requirements. The key to efficiency and accuracy in modeling a system probably is more affected by the formulation of a proper and appropriate conceptual model than by the choice of a particular numerical method or code.

Once a decision to develop a model has been made, a code (or generic model) must be selected (or modified or constructed) that is appropriate for the given problem. Next, the generic code must be adapted to the specific site or region being simulated. Development of a numerical deterministic, distributed-parameter, simulation model involves selecting or designing spatial grids and time increments that will yield an accurate solution for the given system and problem. The analyst must then specify the properties of the system (and their distributions), stresses on the system (such as recharge and pumping rates), boundary conditions, initial conditions (for transient problems), and geochemical processes/reactions (if appropriate). All of the parameter specifications and boundary conditions are really part of the overall conceptual model of the system, and the initial numerical model reflects the analyst's conceptual model of the system.

It must always be remembered that a model is an approximation of a very complex reality, and a model is used to simplify that reality in a manner that captures or represents the essential features and processes relative to the problem at hand. In the development of a deterministic groundwater model for a specific

area and purpose, an appropriate level of model complexity (or, rather, simplicity) must be selected. One may be inclined to believe that finer resolution in a model will yield greater accuracy, and there is a legitimate basis for this. However, there also exists the practical constraint that even when appropriate data are available, a finely discretized three-dimensional numerical model may be too large to run on available computers, especially if transport processes are included. The selection of the appropriate model and appropriate level of model complexity remains subjective and dependent on the judgment and experience of the analysts, the objectives of the study, the level of prior information available for the system of interest, and the complexity of the system being modeled. The trade-off between model accuracy and model cost will always be a difficult one to resolve, but will always have to be made. In any case, water managers and other users of model results must be made aware that these trade-offs and judgments have been made and may affect the reliability of the model.

In general, it is more difficult to calibrate a solute-transport model of an aquifer than it is to calibrate a groundwater flow model. Fewer parameters need to be defined to compute the head distribution with a flow model than are required to compute concentration changes with similar confidence using a solute-transport model. Also, in typical field problems, defining the source term for a solute-transport model is especially difficult for point-source contamination problems because the timing and strength of releases of solute mass into an aquifer system are rarely known or reported accurately (and, in fact, are commonly the very point of contention in litigation).

Because the groundwater seepage velocity is determined from the head distribution, and because both advective transport and hydrodynamic dispersion are functions of the seepage velocity, a model of groundwater flow is typically calibrated before a solute-transport model is developed. In fact, in a field environment perhaps the single most important key to understanding a solute-transport problem is the development of an accurate definition (or model) of the flow system. This is particularly relevant to transport in fractured rocks, where simulation is commonly based on porous-media concepts. In highly heterogeneous systems, the potential (or head) field can often be simulated fairly accurately, whereas the calculated velocity field may still be greatly in error, resulting in considerable errors in simulations of transport.

20.6.1 Generic Model Verification

One of the first things that must be demonstrated is that the generic model accurately solves the governing equations for various boundary value problems, an evaluation that is often called model "verification." This is checked by demonstrating that the code gives good results for problems having known solutions. This test is usually done by comparing the numerical model results to that of an analytical solution. Numerical accuracy is rarely a problem for the solution to the flow equation, but may sometimes be a significant problem in transport modeling.

It must be remembered that numerical solutions are sensitive to spatial and temporal discretization. Therefore, even a perfect agreement for test cases only proves that the numerical code can accurately solve the governing equations, not that it will accurately solve problems under any and all circumstances.

Analytical solutions generally require simple geometry, uniform properties, and idealized boundary and initial conditions. The power of the numerical methods is that they are not constrained by the simplification imposed by analytical methods and allow the introduction of nonhomogeneous, aniso-tropic parameter sets, irregular geometry, mixed boundary conditions, and even nonlinearities into the boundary value problems. Usually, analytical solutions approximating these complexities are unavailable for comparison. Therefore, once these complexities are introduced there is no definitive basis for verifying the numerical model.

One approach that improves confidence for complex heterogeneous problems is to compare the model results to experimental data, to results of other well-accepted models, or to some other accepted standard. Such evaluations might best be termed benchmarking. The HYDROCOIN Project used standardized problem definitions as a basis for intercode comparisons (Swedish Nuclear Power Inspectorate, 1987). While this type of benchmarking helps assure consistency, it does not guarantee or measure accuracy. A

collection and detailed discussion of a number of classical groundwater problems that have been used historically as a basis of model evaluation are presented and documented by Ségol (1994).

20.6.2 Grid Design

The dimensionality of the model (i.e., one, two, or three dimensions) should be selected during the formulation of the conceptual model. If a one- or two-dimensional model is selected, then it is important that the grid be aligned with the flow system so that there is no unaccounted flux into or out of the line or plane of the grid. For example, if a two-dimensional areal model is applied, then there should be no significant vertical components of flow and any vertical leakage or flux must be accounted for by boundary conditions; if a two-dimensional profile model is applied, then the line of the cross section should be aligned with an areal streamline, and there should not be any significant lateral flow into or out of the plane of the cross section.

To minimize a variety of sources of numerical errors, the model grid should be designed using the finest mesh spacing and time steps that are possible, given limitations on computer memory and computational time. To the extent possible, the grid should be aligned with the fabric of the rock and with the average direction of groundwater flow. The boundaries of the grid also should be aligned, to the extent possible, with natural hydrologic and geologic boundaries of the system of interest. Where it is impractical to extend the grid to a natural boundary, then an appropriate boundary condition should be imposed at the edge of the grid to represent the net effects of the continuation of the system beyond the grid. This can typically be accomplished using head-dependent leakage (third type) boundary conditions. However, this would preclude calculating (and accounting for) any storage changes outside the active grid. These boundaries should also be placed as far as possible away from the area of interest and areas of stresses on the system, so as to minimize any impact of conceptual errors associated with these artificial boundary conditions. Note that it is possible for certain types of hydraulic boundaries, such as a groundwater divide, to change location over time if they are located near a major hydraulic stress. If this is anticipated, it might be preferable to extend the boundary of the grid some distance beyond the location of such a natural boundary.

In designing the grid, the length-to-width ratio (or aspect ratio) of cells or elements should be kept as close to one as possible. Long linear cells or elements can lead to numerical instabilities or errors, and should be avoided, particularly if the aspect ratio is greater than about five (Bear and Verruijt, 1987). However, this is a loose guideline as aspect ratios exceeding 100:1 are often used without introducing significant error. In applying this guideline to triangular finite-element methods, Torak (1993) recommends that angles less than 22.5° in a triangle should be avoided.

In specifying boundary conditions for a particular problem and grid design, care must be taken not to overconstrain the solution. That is, if dependent values are fixed at too many boundary nodes, at either internal or external nodes of a grid, the model may have too little freedom to calculate a meaningful solution. At the extreme, by manipulating boundary conditions, one can force any desired solution at any given node. While a forced solution may assure a perfect match to observed data used for calibration, such a match is, of course, not an indicator of model accuracy or reliability and, in fact, can be meaningless (Franke and Reilly, 1987).

To optimize computational resources in a model, it is sometimes advisable to use an irregular (or variably spaced) mesh in which the grid is finest in areas of point stresses, where gradients are steepest, where data are most dense, where the problem is most critical, and (or) where greatest numerical accuracy is desired. It is generally advisable to increase the mesh spacing by a factor no greater than about two between adjacent cells or elements. Similarly, time steps can often be increased geometrically during a transient simulation. At the initial times or after a change in the stress regime, very small time steps should be imposed, because that is when changes in the dependent variable over time are the greatest. As elapsed time increases, the rate of change in head typically decreases, so time steps can often be safely increased by a factor of two or more.

Because transmissivity is a property of the porous media, the cross-product terms of the transmissivity tensor drop out of the governing flow equation that is solved in a model by aligning the model grid with the major axes of the transmissivity tensor (as represented in Equation [5]). This makes the code simpler and more efficient, and, in fact, is a required assumption for most finite-difference models. However, this same simplification typically is not possible for the dispersion tensor in the transport equation because it is also related to, and depends on, the flow direction, which changes orientation over space and time. In general, it is not possible to design a fixed grid that will always be aligned with a changing flow field.

20.6.3 Model Calibration

Deterministic groundwater simulation models impose large requirements for data to define all of the parameters at all of the nodes of a grid. To determine uniquely the parameter distribution for a field problem, so much expensive field testing would be required that it is seldom feasible either economically or technically. Therefore, the model typically represents an attempt, in effect, to solve a large set of simultaneous equations having more unknowns than equations. It is inherently impossible to obtain a unique solution to such a problem.

Uncertainty in parameters logically leads to a lack of confidence in the interpretations and predictions that are based on a model analysis, unless the model can be demonstrated to be a reasonably accurate representation of the real system. To demonstrate that a deterministic groundwater simulation model is realistic, usually field observations of aquifer responses (such as changes in water levels for flow problems or changes in concentration for transport problems) are compared to corresponding values calculated by the model. The objective of this calibration procedure is to minimize differences between the observed data and calculated values. Usually, the model is considered calibrated when it reproduces historical data within some acceptable level of accuracy. The level of acceptability is, of course, determined subjectively. Although a poor match provides evidence of errors in the model, a good match in itself does not prove the validity or adequacy of the model (Konikow and Bredehoeft, 1992).

Because of the large number of variables in the set of simultaneous equations represented in a model, calibration will not yield a unique set of parameters. Where the match is poor, it suggests (1) an error in the conceptual model, (2) an error in the numerical solution, or (3) a poor set of parameter values. Even when the match to historical data is good, the model may still fail to predict future responses accurately, especially under a newer or more extended set of stresses than were experienced during the calibration period.

The calibration of a deterministic groundwater model is often accomplished through a trial and error adjustment of the model's input data (aquifer properties, sources and sinks, and boundary and initial conditions) to modify the model's output. Because a large number of interrelated factors affect the output, trial and error adjustment may become a highly subjective and inefficient procedure. Advances in parameter estimation procedures help to eliminate some of the subjectivity inherent in model calibration (Yeh, 1986). The newer approaches generally treat model calibration as a statistical procedure using multiple regression approaches. Parameter estimation procedures allow the simultaneous construction, application, and calibration of a model using uncertain data, so that the uncertainties in model parameters and in predictions and assessments can be quantified.

However, even with regression modeling, the hydrologic experience and judgment of the modeler continues to be a major factor in calibrating a model both accurately and efficiently. In any case, the modeler should be very familiar with the specific field area being studied in order to ensure that both the data base and the numerical model adequately represent prevailing field conditions. The modeler must also recognize that uncertainty in specification of sources, sinks, and boundary and initial conditions should be evaluated during the calibration procedure in the same manner as uncertainty in aquifer properties. Failure to recognize the uncertainty inherent both in the input data and in the calibration data may lead to "fine-tuning" of the model through unjustifiably precise parameter adjustments strictly to improve the match between observed and calculated variables. This may serve only to provide a false

confidence in the model without producing an equivalent (or any) increase in the predictive accuracy of the model or any improved conceptual understanding of the real system. Freyberg (1988) illustrated this in an exercise in which several groups were given the task of modeling a particular hypothetical groundwater problem. The group that achieved the best calibration, as measured by the minimum root mean square error, was not the group that developed the model that yielded the best prediction (measured by the same criterion). Freyberg (1988, p. 360) concluded that "simple measures of the goodness of a calibrated fit to head data are inadequate to evaluate the true worth of a calibrated parameter set."

Figure 20.8 illustrates in a general manner the use and role of deterministic models in the analysis of groundwater problems. The value of the modeling approach is its capability to integrate site-specific data with equations describing the relevant processes as a quantitative basis for predicting changes or responses in a groundwater system. There must be allowances for feedback from the stage of interpreting model output both to the data collection and analysis phase and to the conceptualization and mathematical definition of the relevant governing processes. One objective of model calibration should be to improve the conceptual model of the system. Because the model quantitatively integrates the effects of the many factors that affect groundwater flow or solute transport, the calculated results should be internally consistent with all input data, and it can be determined if any element of the conceptual model should be revised. In fact, prior concepts or interpretations of aquifer parameters or variables, such as represented by potentiometric maps or the specification of boundary conditions, may be revised during the calibration procedure as a result of feedback from the model's output. In a sense, any adjustment of input data constitutes a modification of the conceptual model.

Automated parameter-estimation techniques improve the efficiency of model calibration and have two general components — one part that calculates the best fit (sometimes called automatic history matching) and a second part that evaluates the statistical properties of the fit. The objective of automatic history matching is to obtain the estimates of system parameters that yield the closest match (minimize deviations) between observed data and model calculations. Least squares deviation is usually chosen as a criterion. The minimization procedure uses sensitivity coefficients that are based on the change in calculated value divided by the change in the parameter. For groundwater flow, for example, this may
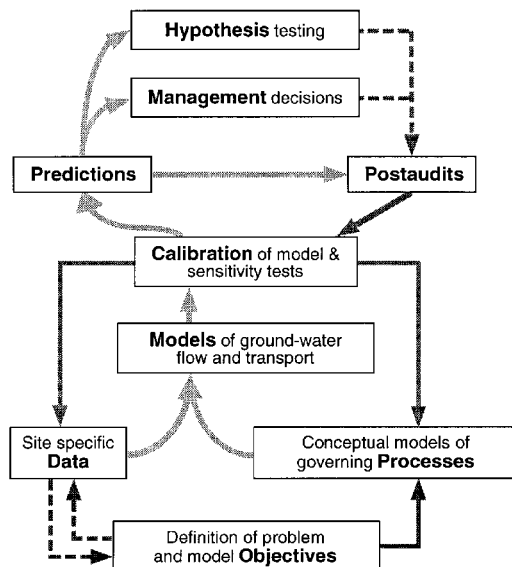
FIGURE 20.8 The use and role of models in the analysis of groundwater problems. (Adapted from Konikow, L. F. 1996. Numerical models of groundwater flow and transport, in *Manual on Mathematical Models in Isotope Hydrogeology*. International Atomic Energy Agency, Vienna.)

take the specific form of $\partial h/\partial T$; that is, the change in head with changing transmissivity. The sensitivity coefficients themselves may be useful in the consideration of additional data collection.

Parameter uncertainty is commonly addressed using a sensitivity analysis. A major objective of sensitivity analysis of simulation models is to determine the change in model results as a result of changes in the model input or system parameters. Conventional sensitivity analysis uses direct parameter sampling in which parameters are perturbed one by one and the complete set of system equations are resolved (Konikow and Mercer, 1988). Sensitivity coefficients for each of these perturbed parameters may be derived by a finite-difference approximation.

20.6.4 Model Error

Discrepancies between observed and calculated responses of a system are the manifestation of errors in the conceptual or mathematical model. In applying groundwater models to field problems, there are three sources of error, and it may not be possible to distinguish among them (Konikow and Bredehoeft, 1992). One source is conceptual errors — that is, misconceptions about the basic processes that are incorporated in the model. Conceptual errors include both neglecting relevant processes as well as inappropriate representation of processes. Examples of such errors include the use of a two-dimensional model where significant flow or transport occurs in the third dimension, or the application of a model based upon Darcy's law to media or environments where Darcy's law is inappropriate. A second source of error involves numerical errors arising in the equation-solving algorithm. These include truncation errors, round-off errors, and numerical dispersion. A third source of error arises from uncertainties and inadequacies in the input data that reflect our inability to describe comprehensively and uniquely the aquifer properties, stresses, and boundaries. In most model applications, conceptualization problems and uncertainty concerning the input data are the most common sources of error.

Numerical methods in general yield approximate solutions to the governing equations. There are a number of possible sources of numerical error in the solution. If model users are aware of the source and nature of these errors, they can control them and interpret the results in light of the presence of error. In solving advection-dominated transport problems in which a relatively sharp front (or steep concentration gradient) is moving through a system, it is numerically difficult to preserve the sharpness of the front.

Obviously, if the width of the front is narrower than the node spacing, then it is inherently impossible to calculate the correct values of concentration in the vicinity of the sharp front. However, even in situations where a front is less sharp, the numerical solution technique can calculate a greater dispersive flux than would occur by physical dispersion alone or would be indicated by an exact solution of the governing equation. That part of the calculated dispersion (or spreading of solute about the center of mass) introduced solely by the numerical solution algorithm is called numerical dispersion.

Figure 20.9 illustrates calculated breakthrough curves for a hypothetical problem of uniform flow and transport to the right, at some time and distance after a tracer having a relative concentration of 1.0 was injected at some point upstream. Curve A represents the breakthrough curve and position of a sharp front for a case having no dispersion (plug flow). Curve B represents an exact analytical solution for a nonzero dispersivity. Curve C illustrates the breakthrough curve calculated for the same conditions as B, but using a numerical method that introduces numerical dispersion. Significant differences exist between the analytical solution (B) and the numerical solution (C) in parts of the domain. Therefore, care must be taken to assess and minimize such numerical errors that would artificially add "numerical" spreading or mixing to the calculated dispersion attributable to physical and chemical processes.

Numerical dispersion can be controlled by reducing the grid spacing (Δx and Δy). However, reduction to a tolerable level may require an excessive number of nodes and render the computational costs unacceptably high. It may also be controlled in finite-element methods by using higher order basis functions or by adjusting the formulation of the equations (using different combinations of forward, backward, or centered in time and/or space, or using different weighting functions). Unfortunately, many approaches that eliminate or minimize numerical dispersion introduce oscillatory behavior, causing
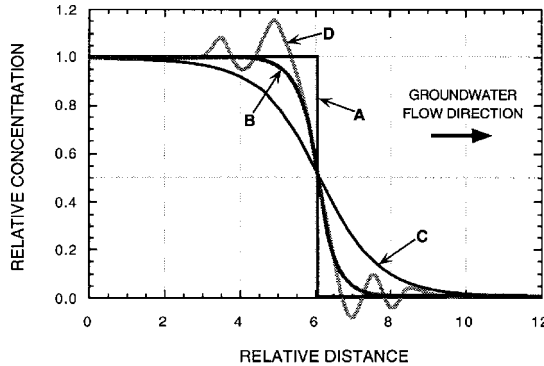
FIGURE 20.9 Representative breakthrough curves for a simple flow and transport problem to illustrate types of numerical errors that may occur in numerical solution to transport equation: (A) plug flow having no dispersion, (B) "exact" solution for transport with dispersion, (C) numerical solution for case B that exhibits effects of numerical dispersion, and (D) numerical solution for case B that exhibits oscillatory behavior. (Adapted from Konikow, L. F. 1996. Numerical models of groundwater flow and transport, in *Manual on Mathematical Models in Isotope Hydrogeology.* International Atomic Energy Agency, Vienna.)

overshoot behind a moving front and possibly undershoot ahead of the front (see curve D in Figure 20.9), and vice versa. Undershoot can result in the calculation of negative concentrations, which are obviously unrealistic. Overshoot can introduce errors of equal magnitude that may go unnoticed because the value is positive in sign (although greater than the source concentration, so still unrealistic). Oscillations generally do not introduce any mass balance errors, and often dampen out over simulation time. However, in some cases, oscillatory behavior can become unbounded, yielding an unstable solution or failure to converge numerically.

In solving the advective-dispersive transport equation, some numerical errors (mainly oscillations) can be related to two dimensionless parameter groups (or numbers). One is the Peclet number, P_e, which may be defined as $P_e = \Delta l/\alpha$, where Δl is a characteristic nodal spacing (although it should be noted that there are several alternative, though essentially equivalent, ways to define P_e). Anderson and Woessner (1992) recommend that the grid be designed so that $\Delta l < 4\alpha$ (or $P_e < 4$); Ségol (1994) recommends a criterion of $P_e \leq 2$. Similarly, time discretization can be related to the Courant number, C_o, which may be defined as $C_o = V\Delta t/\Delta l$ (Anderson and Woessner, 1992). Anderson and Woessner (1992) recommend that time steps be specified so that $\Delta t < \Delta l/V$ (or $C_o < 1.0$), which is equivalent to requiring that no solute be displaced by advection more than the distance across one grid cell or element during one time increment. Numerical error associated with the deviation of curves *C* or *D* (Figure 20.9) from the exact solution can be significant in some locations within the problem domain, although such errors tend to be minimal at the center of a front (relative concentration of 0.5).

In transport models, there may also be a grid-orientation effect in which the solute distribution, calculated for the same properties and boundary conditions, will vary somewhat depending on the angle of the flow relative to the grid. This phenomenon is largely related to the cross-product terms in the governing equation, and generally is not a serious source of error, but the model user should be aware of it.

20.6.5 Mass Balance

One measure of numerical accuracy is how well the model conserves mass. This can be measured by comparing the net fluxes calculated or specified in the model (e.g., inflow and sources minus outflow and sinks) with changes in storage (accumulation or depletion). Mass-balance calculations should always be performed and checked during the calibration procedure to help assess the numerical accuracy of the solution.

As part of these calculations, the hydraulic and chemical fluxes contributed by each distinct hydrologic component of the flow and transport model should be itemized separately to form hydrologic and chemical budgets for the system being modeled. The budgets are valuable assessment tools because they provide a measure of the relative importance of each component to the total budget.

Errors in the mass balance for flow models should generally be less than 0.1%. However, because the solute-transport equation is more difficult to solve numerically, the acceptable mass-balance error for a solute may be greater than for the fluid, but this will depend also on the nature of the numerical method implemented. Finite-difference and finite-element methods are inherently mass conservative, while some implementations of the method of characteristics and particle tracking approaches may not be (or their mass balance calculations themselves are only approximations). It must also be remembered that while a large mass-balance error provides evidence of a poor numerical solution, a perfect mass balance in itself does not and cannot prove that a true or accurate solution has been achieved or that the overall model is valid. That is, a perfect mass balance can be achieved if the model includes compensating errors. For example, the solutions C and D in Figure 20.9 that exhibit significant numerical dispersion or oscillatory behavior arise from solutions that show a near-perfect mass balance, but they are still wrong.

20.6.6 Sensitivity Tests

Assuming various values for given parameters also helps to achieve another objective of the calibration procedure, namely to determine the sensitivity of the model to factors that affect groundwater flow and transport and to errors and uncertainty in the data. Evaluating the relative importance of each factor helps determine which data must be defined most accurately and which data are already adequate or require only minimal further definition. If additional field data can be collected, such a sensitivity analysis helps in deciding which types of data are most critical and how to get the best information return on the costs of additional data collection. If additional data cannot be collected, then the sensitivity tests can help to assess the reliability of the model by demonstrating the effect of a given range of uncertainty or error in the input data on the output of the model. The relative sensitivities of the parameters that affect flow and transport will vary from problem to problem. Furthermore, the sensitivities may change over time as the stress regime imposed on a system evolves. Thus, one generalization is that a sensitivity analysis should be performed during the early stages of a model study.

The sensitivity of the solution to the grid design (or spacing), time-step criterion, nature and placement of boundary conditions, and other numerical parameters also should be evaluated, even if an inverse or regression modeling approach has been used. This step is frequently overlooked, but failure to do so may cause critical design flaws to remain undetected. For example, parameter-estimation models cannot evaluate the sensitivity to grid spacing or certain boundary conditions that are fixed in the model by the user. It is generally recommended that after a preliminary calibration has been achieved, the model should be rerun for the same stresses and properties using a finer grid, smaller time steps, and perhaps alternative boundary conditions. If such a test yields significantly different results, then the model should be recalibrated using design criteria that yield a more accurate numerical solution. If such a test yields no significant differences, then the coarser design is probably adequate for that particular problem.

20.6.7 Calibration Criteria

Model calibration may be viewed as an evolutionary process in which successive adjustments and modifications to the model are based on the results of previous simulations. The modeler must decide when sufficient adjustments have been made to the representation of parameters and processes and at some time accept the model as being adequately calibrated (or perhaps reject the model as being inadequate and seek alternative approaches). This decision is often based on a mix of subjective and objective criteria. The achievement of a best fit between values of observed and computed variables is a regression procedure and can be evaluated as such. That is, the residual errors should have a mean that approaches zero and the deviations should be minimized. Cooley (1977) discusses several statistical measures that can be used

to assess the reliability and "goodness of fit" of groundwater flow models. The accuracy tests should be applied to as many dependent variables as possible. The types of observed data that are most valuable for model calibration include head and concentration changes over space and time, and the quantity and quality of groundwater discharges from the aquifer.

While it is necessary to evaluate the accuracy of the model quantitatively, it is equally important to assure that the dependent variables that serve as a basis for the accuracy tests are reliable indicators of the computational power and accuracy of the model. For example, if a particular dependent variable was relatively insensitive to the governing parameters, then the existence of a high correlation between its observed and computed values would not necessarily be a reflection of a high level of accuracy in the overall model.

Similarly, caution must be exercised when the "observed data" contain an element of subjective interpretation. For example, matching an observed potentiometric surface or concentration distribution is sometimes used as a basis for calibrating groundwater models. However, a contoured surface is itself interpretive and can be a weak basis for model calibration because it includes a variability or error introduced by the contouring process, in addition to measurement errors present in the observed data at the specific points.

20.6.8 Predictions and Postaudits

As model calibration and parameter estimation are keyed to a set of historical data, the confidence in and reliability of the calibration process is proportional to the quality and comprehensiveness of the historical record. The time over which predictions are made with a calibrated model should also be related to, and limited by, the length of the historical record. A reasonable guideline is to predict only for a time comparable to the period that was matched.

The accuracy of a model's predictions is the best measure of its reliability. However, predictive accuracy can be evaluated only after the fact. Anderson and Woessner (1992) summarize several published studies in which the predictive accuracy of a deterministic groundwater model was evaluated several years after the prediction had been made. The results suggest that extrapolations into the future were rarely very accurate. Predictive errors often were related to having used a time period for history matching that was too short to capture an important element of the model or of the system, or to having an incomplete conceptual model. For example, processes and boundary conditions that are negligible or insignificant under the past and present stress regime may become nontrivial or even dominant under a different set of imposed stresses. Thus, a conceptual model founded on observed behavior of a groundwater system may prove to be inadequate in the future, when existing stresses are increased or new stresses are added. A major source of predictive error is sometimes attributable primarily to the uncertainty of future stresses, which is often controlled by demographic, political, economic, and (or) social factors. But if the range or probability of future stresses can be estimated, then the range or probability of future responses can be predicted. An encouraging trend is that many analysts are now attempting to place confidence bounds on predictions arising out of the uncertainty in parameter estimates. However, these confidence limits still would not bound errors arising from the selection of a wrong conceptual model or from problems in the numerical solution algorithms (Bredehoeft and Konikow, 1993).

If a model is to be used for prediction relating to a problem or system that is of continuing interest or significance to society, then field monitoring should continue and the model should be periodically postaudited, or recalibrated, to incorporate new information, such as changes in imposed stresses or revisions in the assumed conceptual model. A postaudit offers a means to evaluate the nature and magnitude of predictive errors, which may itself lead to a large increase in the understanding of the system and in the value of a subsequently revised model. Revised predictions can then be made with greater reliability.

20.6.9 Model Validation

It is natural for people who apply groundwater models, as well as those who make decisions based on model results, to want assurance that the model is valid. Groundwater models are embodiments of various scientific theories and hypotheses. Karl Popper (1959) argues that "as scientists we can never validate a hypothesis, only invalidate it." The same philosophy has been applied specifically to groundwater models (Konikow and Bredehoeft, 1992; Oreskes et al., 1994).

The criteria for labeling a model as validated are inherently subjective. In practice, validation is attempted through the same process that is typically and more correctly identified as calibration — that is, by comparing calculations with field or laboratory measurements. However, the nonuniqueness of model solutions means that a good comparison can be achieved with an inadequate or erroneous model. Also, because the definition of "good" is subjective, under the common operational definitions of validation, one competent and reasonable scientist may declare a model as validated while another may use the same data to demonstrate that the model is invalid. To the general public, proclaiming that a groundwater model is validated carries with it an aura of correctness that many modelers would not claim (Bredehoeft and Konikow, 1993). Because labeling a model as having been validated has very little objective or scientific meaning, such "certification" does little beyond instilling a false sense of confidence in such models. Konikow and Bredehoeft (1992) recommend that the term "validated" not be applied to groundwater models.

20.7 Overview of Representative Generic Models

A large number and variety of generic groundwater models are documented and available at the present time. Two widely used public domain models are explained in more detail as illustrative examples.

20.7.1 MODFLOW

One of the most popular and comprehensive deterministic groundwater models available today is the MODFLOW code of McDonald and Harbaugh (1988) and Harbaugh and McDonald (1996). This is actually a family of compatible codes that centers on an implicit finite-difference solution to the three-dimensional flow equation that was coded in FORTRAN in a modular style to allow and encourage the development of additional packages or modules that can be added on or linked to the original code. The basic model uses a block-centered finite-difference grid that allows variable spacing of the grid in three dimensions. Flow can be steady or transient. Layers can be simulated as confined, unconfined, or a combination of both. Aquifer properties can vary spatially and hydraulic conductivity (or transmissivity) can be anisotropic. Flow associated with external stresses, such as wells, areally distributed recharge, evapotranspiration, drains, and streams, can also be simulated through the use of specified head, specified flux, or head-dependent flux boundary conditions. The implicit finite-difference equations can be solved using either the strongly implicit procedure (SIP) or slice-successive overrelaxation (SSOR) methods. Newer packages offer several additional solution algorithms, including a preconditioned conjugate-gradient solver (Hill, 1990) and a direct solver (Harbaugh, 1995). Although the input and output systems of the program were designed to permit maximum flexibility, usability and ease of interpretation of model results can be enhanced by using one of several commercially available preprocessing and post-processing packages; some of these operate independently of MODFLOW, whereas others are directly integrated into reprogrammed and (or) recompiled versions of the MODFLOW code.

The pathline program MODPATH (Pollock, 1989, 1994) uses the results of the MODFLOW model and determines paths and travel times of water movement under steady-state and transient conditions. MODPATH uses a semianalytical particle-tracking scheme. The method assumes that each directional velocity component varies linearly within a grid cell in its own coordinate direction. MODPATH-PLOT is a graphics interface package that visually displays the results of MODPATH (Pollock, 1994).

The parameter-estimation package, MODFLOWP, can be used to estimate parameters (such as transmissivity, storage coefficient, leakance coefficients, recharge rates, evapotranspiration, and hydraulic head at constant-head boundaries) using nonlinear regression (Hill, 1992). Parameters are estimated by minimizing a weighted least-squares objective function by either the modified Gauss-Newton method or a conjugate-direction method. Data used to estimate parameters can include independent estimates of parameter values, observed heads or drawdowns, and observed gains or losses in streamflow. The MODFLOWP output includes statistics for analyzing the reliability of the estimated parameters and of the model.

A variety of other MODFLOW accessory codes, packages, and features are available. Most of these were developed by the U.S. Geological Survey (USGS) and are summarized by Appel and Reilly (1994); examples include coupled surface-water and groundwater flow, aquifer compaction, transient leakage from confining units, rewetting of dry cells, horizontal flow barriers, alternative interblock transmissivity conceptualizations, cylindrical flow to a well, a statistical processor, a data input program, and a program that calculates water budgets. Other packages have been developed by non-USGS sources to work with MODFLOW; one example is the advective-dispersive solute-transport model MT3D (Zheng, 1990).

20.7.2 MOC

The Method of Characteristics (MOC) model developed by Konikow and Bredehoeft (1978) simulates solute transport in flowing groundwater in two dimensions. The model has been extensively used since the mid-1970s and has been evolving through updates and improvements. The model computes changes in concentration over time caused by the processes of advective transport, hydrodynamic dispersion, mixing or dilution from fluid sources, and the following types of chemical reactions: first-order irreversible-rate reaction, such as radioactive decay; reversible equilibrium-controlled sorption with linear, Freundlich, or Langmuir isotherms; and reversible equilibrium-controlled ion exchange for monovalent or divalent ions. The model couples the groundwater flow equation with the solute-transport equation. The model uses a finite-difference approximation to the groundwater flow equation and the method of characteristics to solve the solute-transport equation. The model uses a particle tracking procedure to represent advective transport and an explicit finite-difference procedure to calculate concentration changes due to hydrodynamic dispersion.

The original model of Konikow and Bredehoeft (1978) was later revised by Goode and Konikow (1989) and Konikow et al. (1994). The model has also been used as the foundation for MOCDENSE (Sanford and Konikow, 1985), a model that can simulate two constituents in a density-dependent flow system. There are public-domain and commercial preprocessors available, including PREMOC (Granato et al., 1993). A three-dimensional version of the model (MOC3D) uses MODFLOW to simulate the flow system (Konikow et al., 1996).

20.8 Case Histories

A large number of documented examples of the application of groundwater models to a variety of hydrogeologic problems are available in the literature. Two case studies have been selected to help illustrate modeling philosophy and practice, including aspects of model conceptualization, model implementation, and interpretation of results.

20.8.1 Regional-Scale Flow in a Deep Confined Aquifer

The Powder River Basin of northeastern Wyoming and southeastern Montana contains large coal reserves that have not yet been fully developed. The future development of such energy resources in the Powder River Basin will be accompanied by increased demands for water, which is not abundantly available in this semiarid area. One plan had been formulated to construct a coal-slurry pipeline to transport coal out of the area; it would have required about 0.6 to 0.8 m^3/s of water. In the mid-1970s, a plan was

proposed to supply this water from up to 40 wells drilled about 1000 m into the Mississippian age Madison Limestone in Niobrara County, Wyoming. The Madison aquifer is an areally extensive carbonate rock system that underlies an area exceeding 260,000 km² in the northern Great Plains.

Concern that such relatively large groundwater withdrawals might cause significant water-level declines in the Madison aquifer, perhaps extending into adjacent states, as well as possibly causing decreases in streamflow and spring discharge in or near the outcrop areas, resulted in the need to predict the effects of the proposed large groundwater withdrawals on potentiometric levels, recharge, and discharge. Because the Madison aquifer lies at such great depths (from 300 to 5000 m) in most of the area, it is relatively undeveloped, and sufficient data are not available to define the head distribution and the hydraulic properties of the aquifer accurately and precisely. In light of this uncertainty, and as a prelude to a planned subsequent 5-year hydrogeologic investigation of the Madison aquifer, a preliminary, two-dimensional, finite-difference model of the aquifer was developed (Konikow, 1976). The objectives of the preliminary model study were to: (1) improve the conceptual model of groundwater flow in the aquifer system; (2) determine deficiencies in existing data, and help set priorities for future data collection by identifying the most sensitive parameters, assuming the model is appropriate; and (3) make a preliminary estimate of the regional hydrologic effects of the proposed well field (Konikow, 1976).

The results indicated that the aquifer can probably sustain the increased groundwater withdrawals, but that they probably would significantly lower the potentiometric surface in the Madison aquifer in a large part of the basin. Because of the great uncertainty in most of the parameters needed to represent the flow system, the model study and predictions were framed in terms of a sensitivity analysis. For example, Figure 20.10 shows drawdown predictions made for an area near the proposed well field for an assumed reasonable range of values for the storage and leakance coefficients (K_Z/m), where K_Z and m are the vertical hydraulic conductivity and the thickness, respectively, of the confining layer. The curves show that the range in plausible drawdowns, even after 1 year, is extremely large. The solutions also illustrate that sensitivities vary with time. At late times (about 100 years), there is no significant difference in drawdown for different values of S (simulations A, B, and C), and at early times (up to about 0.1 years) the drawdown is about the same for all values of leakance at a given value of S (simulations B, D, E, and F).

This preliminary model analysis helped in formulating an improved conceptual model of the Madison aquifer. For example, the important influences of temperature differences and aquifer discontinuities on groundwater flow were recognized and documented (see Konikow, 1976). Because the discrepancies between observed heads and those calculated with the earliest preliminary models did not appear to be
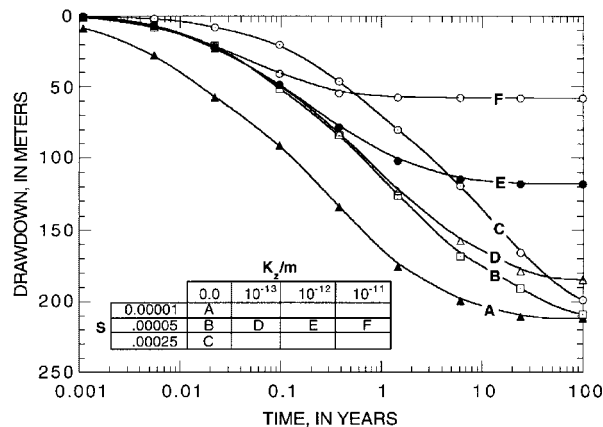
FIGURE 20.10 Time-drawdown curves for model node located near proposed well field to pump groundwater from the Madison Limestone aquifer. (Adapted from Konikow, L. F. 1976. Preliminary digital model of ground-water flow in the Madison group, Powder River basin and adjacent areas, Wyoming, Montana, South Dakota, North Dakota, and Nebraska. U.S. Geol. Survey Water-Res. Inv. 63-75.)

distributed randomly, it was thought that data uncertainty was not the only source of error. Although it could be argued that the importance of these influences could have been (or should have been) recognized on the basis of hydrogeologic principles without the use of a simulation model, the fact is that none of the earlier published studies of this aquifer system indicated that these factors were of major significance. The improvement over earlier studies arose from the quantitative hypothesis-testing role of the model; the nature of the inconsistencies between observed head distributions and those calculated using the initial estimates of model parameters helped direct the investigators toward testing hypotheses that would resolve or minimize the inconsistencies with only a small increment of added complexity. The demonstrated high sensitivity of drawdown to the leakance coefficient emphasized the need to reevaluate the system in a true three-dimensional framework so as to represent vertical components of flow more accurately, which was done in several subsequent studies (for example, see Downey and Weiss, 1980; Woodward-Clyde Consultants, 1981).

Cooley et al. (1986) applied a nonlinear-regression groundwater flow model to this same aquifer system. Their two-dimensional model was based on a Galerkin finite-element discretization scheme. The finite-element grid and boundary conditions are shown in Figure 20.11. The grid was designed to be finer where more data were available and (or) where hydraulic gradients are relatively steep. Regression analysis was used to estimate parameters, including intrinsic permeabilities of the main aquifer and separate lineament zones, discharges from eight major springs, and specified heads on the model boundaries. The regression approach also yielded statistical measures of the reliability of those parameter estimates. Analysis by Cooley et al. (1986) tends to confirm the existence of lineament zones, which appear to exert a strong influence upon the flow and head distribution in the Madison aquifer.

Thus, results from a variety of models were used to understand the sensitivity of the response of the conceptualized Madison aquifer to changes in simulated aquifer parameters. From these sensitivity analyses, improved predictions of aquifer responses were made, and the confidence in the predictions were assessed.
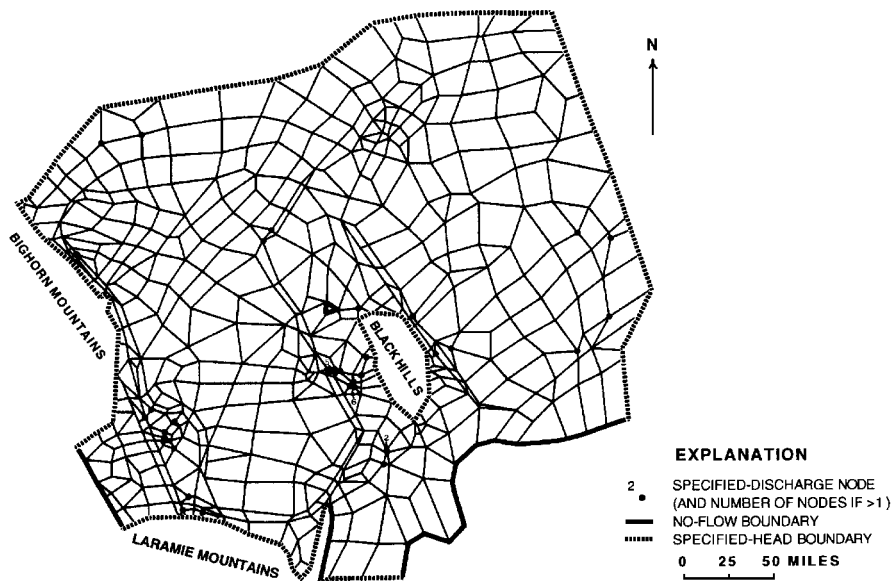
FIGURE 20.11 Finite-element grid showing boundary conditions and locations of specified-discharge points. (Adapted from Cooley, R. L., Konikow, L. F., and Naff, R. L. 1986. Nonlinear-regression groundwater flow modeling of a deep regional aquifer system. *Water Resour. Res.* 10(3):546-562.)

20.8.2 Local-Scale Flow and Transport in a Shallow Unconfined Aquifer

Reilly et al. (1994) combined the application of environmental tracers and deterministic numerical modeling to analyze and estimate recharge rates, flow rates, flow paths, and mixing properties of a shallow groundwater system near Locust Grove, in eastern Maryland. The study was undertaken as part of the U.S. Geological Survey's National Water Quality Assessment Program to provide flow paths and travel time estimates to be used in understanding and interpreting water-quality trends in monitoring wells and stream base flows. The study area encompassed about 2.6×10^7 m^2 of mostly agricultural land on the Delmarva Peninsula. The surficial aquifer includes unconsolidated permeable sands and gravel that range in thickness from less than 6 m to more than 20 m. This surficial aquifer is underlain by relatively impermeable silt and clay deposits, which form a confining unit.

In this study, chlorofluorocarbons (CFCs) and tritium were analyzed from a number of water samples collected from observation wells to estimate the age of groundwater at each sampling location and depth. Because errors and uncertainty are associated with estimates of age based on environmental tracers, just as errors and uncertainty are associated with deterministic models of groundwater flow and transport, the authors applied a feedback or iterative process based on comparisons of independent estimates of travel time. Their approach is summarized and outlined in Figure 20.12. Each task shown was designed to improve either the estimates of parameters or the conceptualization of the system.
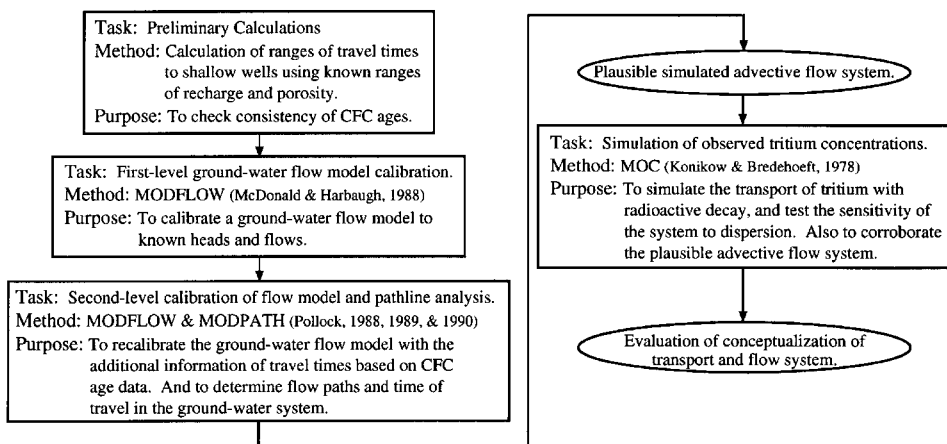
FIGURE 20.12 Flow diagram of the steps taken to quantify the flow paths in the Locust Grove, Maryland, groundwater flow system. (Adapted from Reilly, T. E., Plummer, L. N., Phillips, P. J., and Busenberg, E. 1994. The use of simulation and multiple environmental tracers to quantify groundwater flow in a shallow aquifer. *Water Resour. Res.* 30(2):421-433.)

The preliminary calculations (first task) were used to set bounds on the plausibility of the results of the more complex simulations and chemical analyses. The first-level calibration of a groundwater flow model (second task) provided the initial system conceptualization. The third task was a second-level calibration and analysis involving simulation of advective transport, which provided quantitative estimates of flow paths and time of travel to compare with those obtained from the CFC analyses. The fourth task involved the application of a solute-transport model to simulate tritium concentrations in the groundwater flow system as influenced by the processes of advection, dispersion, radioactive decay, and time-varying input (source concentration) functions.

The sampling wells were located approximately along an areal flow line, and a two-dimensional cross-sectional model was developed for the simulation of processes occurring along this flow line. The MODFLOW model (McDonald and Harbaugh, 1988) was used to simulate groundwater flow and
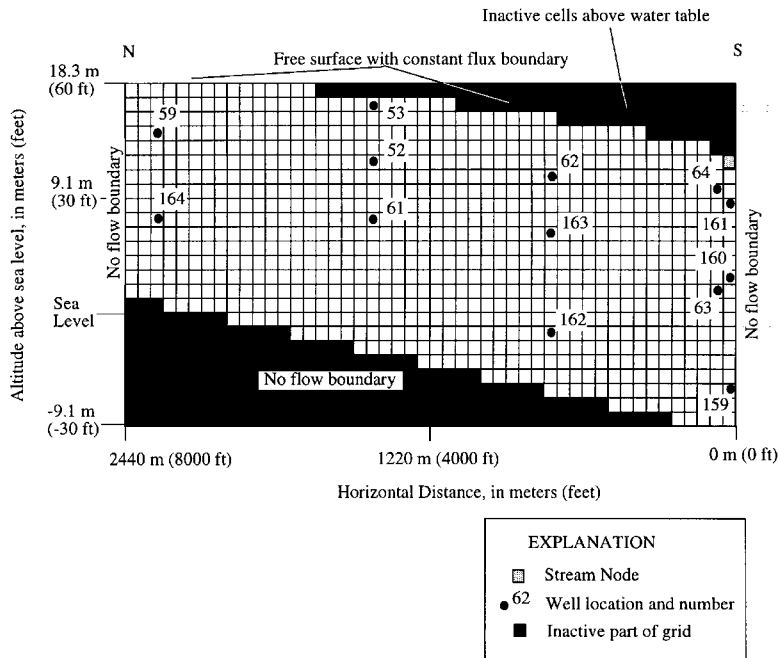
FIGURE 20.13 Model grid used to simulate Locust Grove cross section, showing well locations. (Adapted from Reilly, T. E., Plummer, L. N., Phillips, P. J., and Busenberg, E. 1994. The use of simulation and multiple environmental tracers to quantify groundwater flow in a shallow aquifer. *Water Resour. Res.* 30(2):421-433.)

advective transport. The finite-difference grid consisted of 24 layers and 48 columns of nodes, with each cell having dimensions of 1.14 by 50.80 m, as shown in Figure 20.13, which also shows the wells that lie in the cross section. The simulation was designed to represent average steady-state flow conditions.

After the flow model was calibrated, pathline and travel time analysis was undertaken and comparisons to CFC age estimates were made. Figure 20.14 shows the pathlines calculated using MODPATH (Pollock, 1989) after the second-level calibration with MODFLOW. The comparison with CFC estimates was generally good. However, Reilly et al. (1994) note that close to the stream, many flow lines converge, and the convergence of pathlines representing the entire range of travel times present in the aquifer causes waters of different ages to be relatively near each other. Thus, at the scale and grid spacing of the model, in the area near the stream the convergent flow lines cannot be readily differentiated in the model and the locations of individual well screens cannot be accurately represented directly under the stream. After the second-level calibration, the root mean squared error between the simulated ages and the CFC ages for the 10 wells farthest from the stream (i.e., excluding wells 159, 160, and 161) was 3.4 years.

Tritium concentrations of recharge waters have varied considerably over the last 40 years. Thus, the time of travel would not always be readily apparent from the tritium concentration in a water sample. Also, mixing of waters recharged during periods of these relatively sharp changes of input concentrations can make the interpretation of time of travel from tritium concentrations even more uncertain. Thus, the investigators simulated solute transport of tritium within the system using a model that accounts for mixing (dispersion), radioactive decay, and transient input functions, which also allowed a further evaluation of consistency with the results of the previous flow and advective transport model. They applied the MOC solute-transport model of Konikow and Bredehoeft (1978) and Goode and Konikow (1989) for this purpose.

The results of the simulation of the tritium distribution assuming (1) no dispersion and (2) α_L of 0.15 m and α_T of 0.015 m are shown in Figure 20.15. The limiting case simulation of no dispersion yielded acceptable results and was used as the best estimate of the tritium distribution in November
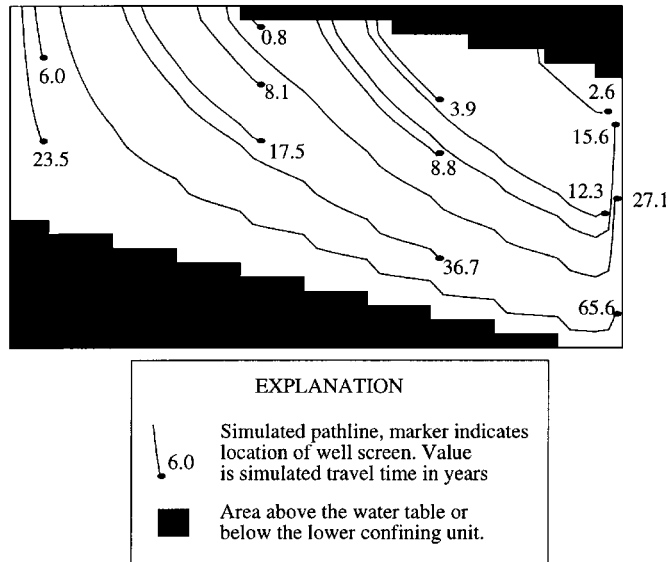
FIGURE 20.14 Pathlines (calculated using MODPATH after second-level calibration) in Locust Grove cross section to observation wells showing time of travel (in years) from the water table. (Adapted from Reilly, T. E., Plummer, L. N., Phillips, P. J., and Busenberg, E. 1994. The use of simulation and multiple environmental tracers to quantify groundwater flow in a shallow aquifer. *Water Resour. Res.* 30(2):421-433.)

1990 (Reilly et al., 1994). This case reproduces the sharp concentration gradients required to reproduce the low tritium values that were observed. The MOC model was advantageous for this problem because it minimizes numerical dispersion and it can solve the governing equations for α_L of 0.0, which transport models based on finite-difference or finite-element methods generally cannot do. The results of the solute-transport simulation are consistent with the advective flow system determined by the second-level calibration and thus strengthen the case for the conceptual model. The coupling of the tritium analyses and the transport model indicates where discrepancies between the measured and simulated concentrations occur, where additional data collection would be most useful, and where refinement of the conceptual model may be warranted.

This case study illustrates that environmental tracers and numerical simulation methods in combination are effective tools that complement each other and provide a means to estimate the flow rate and path of water moving through a groundwater system. Reilly et al. (1994) found that the environmental tracers and numerical simulation methods also provide a "feedback" that allows a more objective estimate of the uncertainties in the estimated rates and paths of movement. Together the two methods enabled a coherent explanation of the flow paths and rates of movement while identifying weaknesses in the understanding of the system that require additional data collection and refinement of conceptual models of the groundwater system.

20.9 Available Groundwater Models

A large number of generic deterministic groundwater models, based on a variety of numerical methods and a variety of conceptual models, are available. The selection of a numerical method or generic model for a particular field problem depends on several factors, including accuracy, efficiency/cost, and usability. The first two factors are related primarily to the nature of the field problem, availability of data, and scope or intensity of the investigation. The usability of a method may depend partly on the mathematical background of the modeler, as it is preferable for the model user to understand the nature of the numerical methods implemented in a code. It may be necessary to modify and adapt the program to the specific
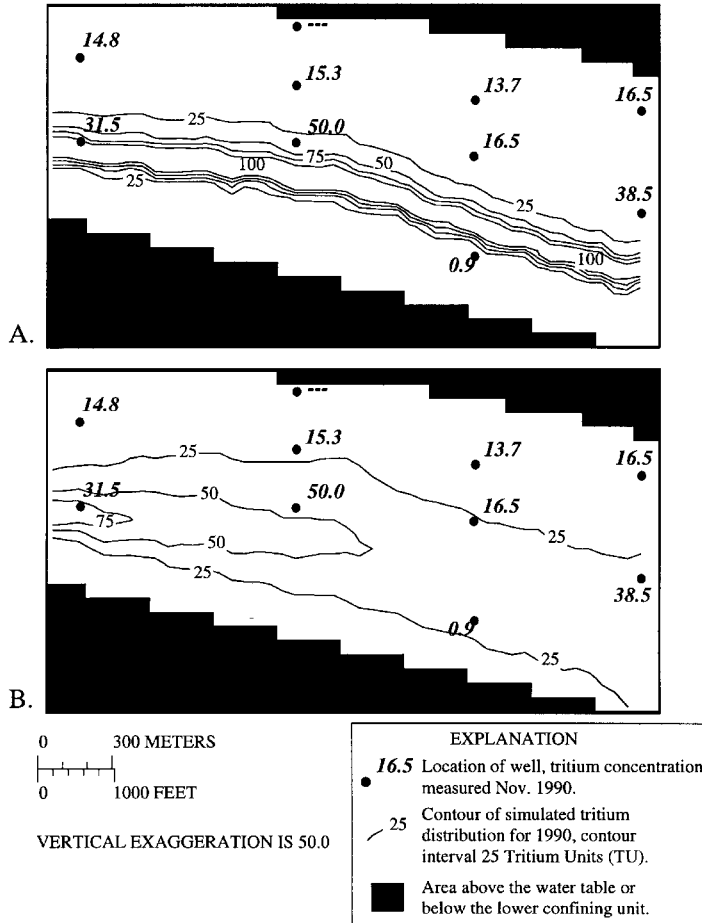
FIGURE 20.15 Simulated tritium distribution at the end of 1990: (A) with dispersivity $\alpha_L = 0.0$ m and $\alpha_T = 0.0$ m, and (B) with dispersivity $\alpha_L = 0.15$ m and $\alpha_T = 0.015$ m. Contour interval 25 tritium units (TU). Measured concentrations from samples obtained from wells in November 1990 are given for their location in bold italics. (Adapted from Reilly, T. E., Plummer, L. N., Phillips, P. J., and Busenberg, E. 1994. The use of simulation and multiple environmental tracers to quantify groundwater flow in a shallow aquifer. *Water Resour. Res.* 30(2):421-433.)

problem of interest, and this can sometimes require modifications to the source code. In selecting a model that is appropriate for a particular application, it is most important to choose one that incorporates the proper conceptual model; one must avoid force fitting an inappropriate model to a field situation solely because of the model's convenience, availability, or familiarity to the user. Usability is also enhanced by the availability of preprocessing and postprocessing programs or features, and by the availability of comprehensive yet understandable documentation.

A number of surveys of available models have been published in recent years (Appel and Reilly, 1994; Van der Heijde et al., 1985). Van der Heijde et al. (1985) report on an international survey of 399 models, of which 206 had been documented at that time. This was a significant increase from the 245 models available for a similar review 5 years earlier. Appel and Reilly (1994) summarize the nature and availability of 89 groundwater flow and quality models produced by and available from the U.S. Geological Survey. Anderson et al. (1992), in their review of groundwater models, list 19 separate software distributors and provide brief descriptions of several codes. The International Ground Water Modeling Center, Golden, CO, (see internet address in *For Further Information*) maintains a clearinghouse and distribution center for groundwater simulation models.

A large number of public and private organizations distribute public domain and (or) proprietary software for groundwater modeling. A growing availability of models is also occurring on the internet (see *For Further Information* for some examples). Some internet sites allow computer codes to be downloaded at no cost, while other sites provide catalog information, demonstrations, and pricing information.

Acknowledgment

The authors appreciate the helpful review comments provided by Robert Nicholson, Michael Planert, and Cliff Voss, all of the USGS, and by Mary P. Anderson of the University of Wisconsin–Madison. We also acknowledge Dr. Y. Yurtsever and the International Atomic Energy Agency for supporting the development of the report upon which part of this chapter is based.

For Further Information

Textbooks and examples of good reports on site-specific models are provided as starting points for readers who would like to obtain more information or study representative applications.

Textbooks

Anderson and Woessner (1992) present an overview of applied groundwater flow and advective transport modeling.

Zheng and Bennett (1995) present an overview of the theory and practice of contaminant transport modeling.

Examples of Reports on Site-Specific Models

Comprehensive reports on site-specific models provide insight into applied groundwater simulation. A few examples from the work of the U.S. Geological Survey are provided below. Obviously, this list is not inclusive, and many other reports could have been listed.

Regional Flow Models

Kernodle et al. (1995) describe a three-dimensional flow model of the Albuquerque Basin in New Mexico.

Fleck and Vroblesky (1996) describe the application of a three-dimensional groundwater flow model to a coastal plain system in the northeastern U.S.

Local Flow Model

Masterson and Barlow (1994) used a two-step approach to simulate a saltwater–freshwater system.

Local Radial-Flow Model

Lindner and Reilly (1983) used a finite-element radial flow model to analyze aquifer tests on Long Island, New York.

Local Advective-Transport Model

Barlow (1994) examined contributing areas to public-supply wells at Cape Cod, Massachusetts.

Solute-Transport Model

LeBlanc (1984) documented a two-dimensional simulation of a 6-km-long sewage plume at Cape Cod, Massachusetts.

Lambert (1996) used a three-dimensional model to simulate a contaminant plume in an approximately 480-km^2 area in Utah.

Model Calibration

Masterson et al. (1996) used particle tracking and contaminant plumes to improve calibration of a three-dimensional flow model.

Yager (1997) used a parameter-estimation model (MODFLOWP) to help calibrate a three-dimensional flow model for a fractured dolomite aquifer system.

Internet

A number of sites on the World Wide Web provide compendia of codes and sources of information about groundwater modeling, as well as providing links to other websites related to groundwater modeling. Many of these sites allow codes to be downloaded. Examples of several groundwater-oriented home page locations are: **http://www.ems.psu.edu/Hydrogeologist/**, **http://www.et.byu.edu/~asce-gw/**, **http://www.mines.edu/igwmc/**, and **http://www.ibmpcug.co.uk/~bedrock/gsd/**. Also, many of the U.S. Geological Survey public domain codes are available from the "USGS Water Resources Applications Software" link on the USGS Water Resources Information Home page at: **http://water.usgs.gov/**.

References

Anderson, M. P. 1984. Movement of contaminants in groundwater: Groundwater transport — Advection and dispersion, in *Groundwater Contamination*. National Academy Press, Washington, D.C., 37-45.

Anderson, M. P., Ward, D. S., Lappala, E. G., and Prickett, T. A. 1992. Computer models for subsurface water, in *Handbook of Hydrology*. Ed. D.R. Maidment, McGraw-Hill, New York, 22.1-22.34.

Anderson, M. P. and Woessner, W. W. 1992. *Applied Groundwater Modeling*. Academic Press, San Diego.

Appel, C. A. and Reilly, T. E. 1994. *Summary of Computer Programs Produced by the U.S. Geological Survey for Simulation of Ground-Water Flow and Quality — 1994*. U.S. Geol. Survey Circular 1104.

Barlow, P. M. 1994. Particle-tracking analysis of contributing areas of public-supply wells in simple and complex flow systems, Cape Cod, Massachusetts. U.S. Geol. Survey Open-File Rept. 93-159.

Bear, J. 1979. *Hydraulics of Groundwater*. McGraw-Hill, New York.

Bear, J. and Verruijt, A. 1987. *Modeling Groundwater Flow and Pollution*. Reidel Publishing Co., Dordrecht, The Netherlands.

Bennett, G. D. 1976. *Introduction to Ground-Water Hydraulics: A Programmed Text for Self-Instruction*. Techniques of Water-Res. Invests. of the U.S. Geol. Survey, Book 3, Ch. B2.

Bredehoeft, J. D. and Konikow, L. F. 1993. Ground-water models: Validate or invalidate. *Ground Water*. 31(2):178-179.

Cooley, R. L. 1977. A method of estimating parameters and assessing reliability for models of steady state groundwater flow, 1., Theory and numerical properties. *Water Resour. Res.* 13(2):318-324.

Cooley, R. L. 1992. *A Modular Finite-Element Model (MODFE) for Areal and Axisymmetric Ground-water-Flow Problems, Part 2: Derivation of Finite-Element Equations and Comparisons with Analytical Solutions*, Techniques of Water-Res. Invests. of the U.S. Geol. Survey, Book 6, Ch. A4.

Cooley, R. L., Konikow, L. F., and Naff, R. L. 1986. Nonlinear-regression groundwater flow modeling of a deep regional aquifer system. *Water Resour. Res.* 10(3):546-562.

Domenico, P. A. and Robbins, G. A. 1984. A dispersion scale effect in model calibrations and field tracer experiments, *J. Hydrol.* 70:123-132.

Domenico, P. A. and Schwartz, F. W. 1990. *Physical and Chemical Hydrogeology*. John Wiley & Sons, New York.

Downey, J. S. and Weiss, E. J. 1980. Preliminary data set for three-dimensional digital model of the Red River and Madison aquifers. U.S. Geol. Survey Open-File Rept. 80-756.

Fleck, W. B. and Vroblesky, D. A. 1996. Simulation of ground-water flow of the coastal plain aquifers in parts of Maryland, Delaware, and the District of Columbia. U.S. Geol. Survey Prof. Paper 1404-J.

Franke, O. L. and Reilly, T. E. 1987. The effects of boundary conditions on the steady-state response of three hypothetical groundwater systems — Results and implications of numerical experiments. U.S. Geol. Survey Water-Supply Paper 2315.

Franke, O. L., Reilly, T. E., and Bennett, G. D. 1987. *Definition of Boundary and Initial Conditions in the Analysis of Saturated Ground-Water Flow Systems — An Introduction*. Techniques of Water-Res. Invests. of the U.S. Geol. Survey, Book 3, Ch. B5.

Freyberg, D. L. 1988. An exercise in ground-water model calibration and prediction. *Ground Water.* 26(3):350-360.

Garder, A. O., Peaceman, D. W., and Pozzi, A. L. 1964. Numerical calculation of multidimensional miscible displacement by the method of characteristics. *Soc. Petroleum Eng. J.* 4(1):26-36.

Gelhar, L. W. 1993. *Stochastic Subsurface Hydrology*. Prentice Hall, Englewood Cliffs, NJ.

Gelhar, L. W., Welty, C., and Rehfeldt, K. R. 1992. A critical review of data on field-scale dispersion in aquifers. *Water Resour. Res.* 28(7):1955-1974.

Goode, D. J. and Appel, C. A. 1992. Finite-difference interblock transmissivity for unconfined aquifers and for aquifers having smoothly varying transmissivity. U.S. Geol. Survey Water-Res. Inv. Rept. 92-4124.

Goode, D. J. and Konikow, L. F. 1989. Modification of a method-of-characteristics solute-transport model to incorporate decay and equilibrium-controlled sorption or ion exchange. U.S. Geol. Survey Water-Res. Inv. Rept. 89-4030.

Goode, D. J. and Konikow, L. F. 1990. Apparent dispersion in transient groundwater flow. *Water Resour. Res.* 26(10):2339-2351.

Granato, G. E., Konikow, L. F., and Srinivasan, P. 1993. PREMOC version 4.0, A preprocessor for the two-dimensional method of characteristics (MOC) solute-transport model. International Ground Water Modeling Center, Golden, CO, Report IGWMC — FOS 23.

Grove, D. B. 1976. Ion exchange reactions important in groundwater quality models, in *Advances in Groundwater Hydrology*, Am. Water Res. Assoc., Minneapolis, 409-436.

Haitjema, H. M. 1995. *Analytic Element Modeling of Groundwater Flow*. Academic Press, San Diego.

Harbaugh, A. W. 1995. Direct solution package based on alternating diagonal ordering for the U.S. Geological Survey modular finite-difference ground-water flow model. U.S. Geol. Survey Open-File Rept. 95-288.

Harbaugh, A. W. and McDonald, M. G. 1996. User's documentation for MODFLOW-96, an update to the U.S. Geological Survey modular finite-difference ground-water flow model. U.S. Geol. Survey Open-File Report 96-485.

Hill, M. C. 1990. Preconditioned Conjugate-Gradient 2 (PCG2) — A computer program for solving ground-water flow equations. U.S. Geol. Survey Water-Res. Inv. Rept. 90-4048.

Hill, M. C. 1992. A computer program (MODFLOWP) for estimating parameters of a transient, three-dimensional, ground-water flow model using nonlinear regression. U.S. Geol. Survey Open-File Rept. 91-484.

Huebner, K. H. 1975. *The Finite Element Method for Engineers*. John Wiley & Sons, New York.

Huyakorn, P. S. and Pinder, G. F. 1983. *Computational Methods in Subsurface Flow*. Academic Press, New York.

Javandel, I., Doughty, D., and Tsang, C. -F. 1984. *Groundwater Transport: Handbook of Mathematical Models*. Am. Geophysical Union, Water Res. Monograph 10.

Kernodle, J. M., McAda, D. P., and Thorn, C. R. 1995. Simulation of ground-water flow in the Albuquerque Basin, central New Mexico, 1901-1994, with projections to 2020. U.S. Geol. Survey Water-Res. Inv. Rept. 94-4251.

Kipp, K. L. Jr. 1987. HST3D: A computer code for simulation of heat and solute transport in three-dimensional ground-water flow systems. U.S. Geol. Survey Water-Res. Inv. Rept. 86-4095.

Konikow, L. F. 1976. Preliminary digital model of ground-water flow in the Madison group, Powder River basin and adjacent areas, Wyoming, Montana, South Dakota, North Dakota, and Nebraska. U.S. Geol. Survey Water-Res. Inv. 63-75.

Konikow, L. F. 1996. Numerical models of groundwater flow and transport, in *Manual on Mathematical Models in Isotope Hydrogeology*. International Atomic Energy Agency, Vienna.

Konikow, L. F. and Bredehoeft, J. D. 1978. *Computer Model of Two-Dimensional Solute Transport and Dispersion in Ground Water.* Techniques of Water-Res. Invests. of the U.S. Geol. Survey, Book 7, Ch. C2.

Konikow, L. F. and Bredehoeft, J. D. 1992. Ground-water models cannot be validated. *Advances in Water Resources.* 15(1):75-83.

Konikow, L. F. and Grove, D. B. 1977. Derivation of equations describing solute transport in ground water. U.S. Geological Survey Water-Res. Inv. Rept. 77-19.

Konikow, L. F., Goode, D. J., and Hornberger, G. Z. 1996. A three-dimensional method-of-characteristics solute-transport model (MOC3D). U.S. Geol. Survey Water-Res. Inv. Rept. 96-4267.

Konikow, L. F., Granato, G. E., and Hornberger, G. Z. 1994. User's guide to revised method-of-characteristics solute-transport model (MOC—Version 3.1). U.S. Geol. Survey Water-Res. Inv. Rept. 94-4115.

Konikow, L. F. and Mercer, J. M. 1988. Groundwater flow and transport modeling, *J. Hydrol.* 100(2):379-409.

Lambert, P. M. 1996. Numerical simulation of the movement of sulfate in ground water in southwestern Salt Lake Valley, Utah. Tech. Pub. No. 110-D, Utah Dept. of Natural Resources, Salt Lake City, UT.

LeBlanc, D. R. 1984. Digital modeling of solute transport in a plume of sewage-contaminated ground water, in *Movement and Fate of Solutes in a Plume of Sewage Contaminated Ground Water, Cape Cod, Massachusetts.* U.S. Geological Survey Toxic Waste Ground-Water Contamination Program, ed. D.R. LeBlanc. P. 11-45. U.S. Geol. Survey Open-File Rept. 84-475.

Lindner, J. B. and Reilly, T. E. 1983. Analysis of three tests of the unconfined aquifer in southern Nassau County, Long Island, New York. U.S. Geol. Survey Water-Res. Inv. Rept. 82-4021.

Masterson, J. P. and Barlow, P. M. 1994. Effects of simulated ground-water pumping and recharge on ground-water flow in Cape Cod, Martha's Vineyard, and Nantucket Island Basins, Massachusetts. U.S. Geol. Survey Open-File Rept. 94-316.

Masterson, J. P., Walter, D. A., and Savoie, J. 1996. Use of particle tracking to improve numerical model calibration and to analyze ground-water flow and contaminant migration, Massachusetts Military Reservation, western Cape Cod, Massachusetts. U.S. Geol. Survey Open-File Rept. 96-214.

McDonald, M. G. and Harbaugh, A. W. 1988. *A Modular Three-Dimensional Finite-Difference Ground-Water Flow Model.* Techniques of Water-Res. Invests. of the U.S. Geol. Survey, Book 6, Ch. A1.

Mercer, J. W. and Faust, C. R. 1981. *Ground-Water Modeling.* Natl. Water Well Assoc., Worthington, Ohio.

Oreskes, N., Shrader-Frechette, K., and Belitz, K. 1994. Verification, validation, and confirmation of numerical models in the earth sciences. *Science.* 263:641-646.

Peaceman, D. W. 1977. *Fundamentals of Numerical Reservoir Simulation.* Elsevier, Amsterdam.

Pollock, D. W. 1988. Semianalytical computation of path lines for finite-difference models. *Ground Water.* 26(6):743-750.

Pollock, D. W. 1989. Documentation of computer programs to compute and display pathlines using results from the U.S. Geological Survey modular three-dimensional finite-difference ground-water flow model. U.S. Geol. Survey Open-File Rept. 89-381.

Pollock, D. W. 1994. User's guide for MODPATH/MODPATH-PLOT, version 3: A particle tracking post processing package for MODFLOW, the U.S. Geological Survey finite-difference ground-water flow model. U.S. Geol. Survey Open-File Rept. 94-464.

Popper, Sir Karl. 1959. *The Logic of Scientific Discovery.* Harper and Row, New York.

Prickett, T. A., Naymik, T. G., and Lonnquist, C. G. 1981. A "random-walk" solute transport model for selected groundwater quality evaluations. Ill. State Water Survey Bulletin 65.

Reddell, D. L. and Sunada, D. K. 1970. Numerical simulation of dispersion in groundwater aquifers. Colorado State University, Ft. Collins, Hydrology Paper 41.

Reilly, T. E., Franke, O. L., Buxton, H. T., and Bennett, G. D. 1987. A conceptual framework for ground-water solute-transport studies with emphasis on physical mechanisms of solute movement. U.S. Geol. Survey Water-Res. Inv. Rept. 87-4191.

Reilly, T. E., Plummer, L. N., Phillips, P. J., and Busenberg, E. 1994. The use of simulation and multiple environmental tracers to quantify groundwater flow in a shallow aquifer. *Water Resour. Res.* 30(2):421-433.

Remson, I., Hornberger, G. M., and Molz, F. J. 1971. *Numerical Methods in Subsurface Hydrology*, John Wiley & Sons, New York.

Sanford, W. E. and Konikow, L. F. 1985. A two-constituent solute-transport model for ground water having variable density. U.S. Geol. Survey Water-Res. Inv. Rept. 85-4279.

Ségol, G. 1994. *Classic Groundwater Simulations: Proving and Improving Numerical Models.* PTR Prentice Hall, Englewood Cliffs, NJ.

Smith, L. and Schwartz, F. W. 1980. Mass transport, 1, A stochastic analysis of macroscopic dispersion. *Water Resour. Res.* 16(2):303-313.

Swedish Nuclear Power Inspectorate. 1987. The International HYDROCOIN Project—Background and Results, OECD, Paris.

Torak, L. J. 1993. A modular finite-element model (MODFE) for areal and axisymmetric ground-water-flow problems, Part 1: Model description and user's manual. Techniques of Water-Res. Invests. of the U.S. Geol. Survey, Book 6, Ch. A3.

Van der Heijde, P. K. M., Bachmat, Y., Bredehoeft, J. D., Andrews, B., Holtz, D., and Sebastian, S. 1985. *Groundwater Management: The Use Of Numerical Models*, American Geophys. Union, Washington, D.C., Water Res. Monograph 5 (2nd Ed.).

Voss, C. I. 1984. SUTRA — Saturated Unsaturated Transport — A finite-element simulation model for saturated-unsaturated fluid-density-dependent ground-water flow with energy transport or chemically-reactive single-species solute transport. U.S. Geol. Survey Water-Res. Invest. Rep. 84-4369.

Wang, J. F. and Anderson, M. P. 1982 (reprinted, 1995). *Introduction to Groundwater Modeling*, Academic Press, San Diego.

Wexler, E. J. 1992. *Analytical Solutions for One-, Two-, and Three-Dimensional Solute Transport in Ground-Water Systems with Uniform Flow*, Techniques of Water-Res. Invests. of the U.S. Geol. Survey, Book 3, Ch. B7.

Woodward-Clyde Consultants. 1981. Well-field hydrology technical report for the ETSI coal slurry pipeline project. Bureau of Land Management, Washington, D.C.

Yager, R. M. 1997. Simulated three-dimensional ground-water flow in the Lockport Group, a fractured dolomite aquifer near Niagara Falls, New York. U.S. Geol. Survey Water-Supply Paper 2487.

Yeh, W. W.-G. 1986. Review of parameter identification procedures in groundwater hydrology: The inverse problem. *Water Resour. Res.* 22(1):95-108.

Zheng, C. 1990. *MT3D: A Modular Three-Dimensional Transport Model.* S.S. Papadopulos and Associates, Inc., Bethesda, MD.

Zheng, C. and Bennett, G. D. 1995. *Applied Contaminant Transport Modeling*, Van Nostrand–Reinhold, New York.

Zienkiewicz, O. C. 1971. *The Finite Element Method in Engineering Science.* McGraw-Hill, London.

Glossary

Analytical Model   A closed-form exact mathematical solution which is continuous in space and time.

Conceptual Model   A hypothesis for how a system or process operates.

Deterministic Model   A mathematical model based on conservation of mass, momentum, and energy.

Discretization   The process of representing a continuous system by a set of discrete blocks, cells, or elements.

Generic Model   The computer code used to solve one or more partial differential equations.

Mathematical Model   A set of equations, which include mathematical variables, constants, and coefficients, that represents relevant processes.

Model   A representation of a real system or process.

Numerical Model   An approximate solution of a differential equation obtained by replacing the continuous variables with a set of discrete variables defined at grid blocks, cells, or nodes.

Site-Specific Model   A numerical model with the parameters (such as hydraulic conductivity, dispersivity, etc.), boundary conditions, and grid dimensions of the generic model specified to represent a particular geographical area.

21

The Role of Geographical Information Systems in Groundwater Engineering

Bernard A. Engel
Purdue University

Kumar C. S. Navulur
Resource 21

21.1 Introduction to Geographic Information Systems (GIS)

Geographic Information Systems (GIS) have become important tools in efficiently solving many problems in which spatial data are important. Natural resources and environmental concerns, including groundwater, have benefited greatly from the use of GIS. This chapter provides a brief introduction to GIS and some of its applications in addressing groundwater issues.

21.1.1 Overview of GIS

GIS have evolved rapidly in the last decade, becoming powerful computer tools for varied applications ranging from sophisticated analysis and modeling of spatial data to simple inventory and management.

The rapid growth, diversity, and commercial orientation of GIS have resulted in multiple definitions of the term GIS. GIS can be described as a collection of computer software and hardware for storing, manipulating, analyzing, and displaying spatial or geographically referenced data. GIS can also be described as a combination of computer tools: cartographic and tools for displaying data and making maps, a database for storing spatial data, and an analysis tool for manipulating data. As a visualization tool, GIS allows graphical display of maps, tabular information, statistical summaries, and modeling solutions. As a database, GIS can store, maintain, and update spatial data and associated descriptive information. GIS supports most database functions such as browsing, tabular queries, updating data, and functions supported by commonly used computer software such as spreadsheets and statistical packages. The most significant difference between GIS and other information systems and databases is the spatial nature of the data in a GIS. The analysis functions in a GIS allow manipulation of multiple themes of spatial data to perform overlays, buffering, and arithmetic operations on the data.

The application of GIS grew dramatically in the 1980s and 1990s, with yearly expenditures on GIS experiencing double-digit growth. Estimated expenditures within the U.S. on GIS were approximately $6 billion in 1995. The ability of GIS to perform analysis on spatial data and corresponding attribute information and to integrate different types of data and at high speeds is unmatched by manual methods. There is an increasing trend on the part of organizations and public agencies at all levels to develop GIS for the management and analysis of spatial data.[1,2] The ability of a GIS to perform complex spatial analyses rapidly provides a quantitative as well as qualitative advantage in planning scenarios, decision models, change detection and analyses, and other problems requiring refinements to successive analyses. Digital databases like STATSGO (State Soils Geographic Database), land use data, satellite imagery, and Digital Elevation Model (DEM) data are now being distributed in GIS formats, and efforts are ongoing to integrate more data into GIS formats.

21.2 GIS Components

GIS, like other information systems, cannot exist on its own, but rather it exists in the context of an application. There must be an organization of people, facilities, and equipment responsible for implementing and maintaining a GIS. Three essential elements that constitute a GIS are computer software, hardware, and personnel.

21.2.1 Software

Existing GIS software are extremely diverse in functionality, database structure, and hardware requirements. A GIS user today is presented with a wide variety of commercial GIS software. Some of the popular GIS software is Arc/Info, ArcView, MapInfo, MGE (Modular GIS Environment), Genasys, ERDAS, GRASS, IDRISI, and Atlas. These software are sophisticated and have a variety of functionalities for GIS applications and analyses. Most of these software come with capabilities to handle data in various formats, thereby facilitating data transfer between GIS software. Some of the above software is hardware specific and others run on a variety of platforms. The marketing of GIS software today is increasingly oriented to an open system or platform-independent approach.

21.2.2 Hardware

The large memory and disk requirements of GIS software and data resulted in a trend toward workstations running the Unix operating system in the late 1980s and early 1990s. However, rapidly declining computer hardware costs and increasing capabilities of personal computers have made GIS software available on a variety of computer platforms including personal computers, high performance workstations, minicomputers, and mainframe computers. GIS applications often require use of special peripherals such as digitizers and scanners for developing digital data. Other hardware requirements for GIS applications

include plotters and printers for output of the final GIS products, including hard copy maps and summaries.

21.2.3 People

One of the most significant elements of a GIS is people. People are responsible for designing, implementing, and using GIS for an application. GIS projects usually require an interdisciplinary team including specialists from several disciplines such as geography, agriculture, photogrammetry, computer science, and natural resources to address multifaceted aspects of the application. Successful design and implementation of a GIS project requires coordination of personnel from these different disciplines. GIS training for key project personnel should be considered.

21.3 Data Representation

Data are the most important resource in a GIS. Digital data are usually very expensive to develop and requires large volumes of storage space. GIS data are commonly stored in one of two forms: (1) vector or (2) raster. Vector data are the set of points, lines, and polygons that are used to represent map feature locations (see Figure 21.1). The spatial relationships between various features must be explicitly coded for vector GIS data. A topological data structure is used to store the relationships between objects. Well locations, streams, and state boundaries are some of the examples of point, line, and polygon vector features, respectively. Vector data can be represented in various graphic formats such as DXF, DLG, and MOSS. GIS software is generally designed to accept and export data in one or more of these graphic exchange formats.

Vector Representation   **Raster Representation**
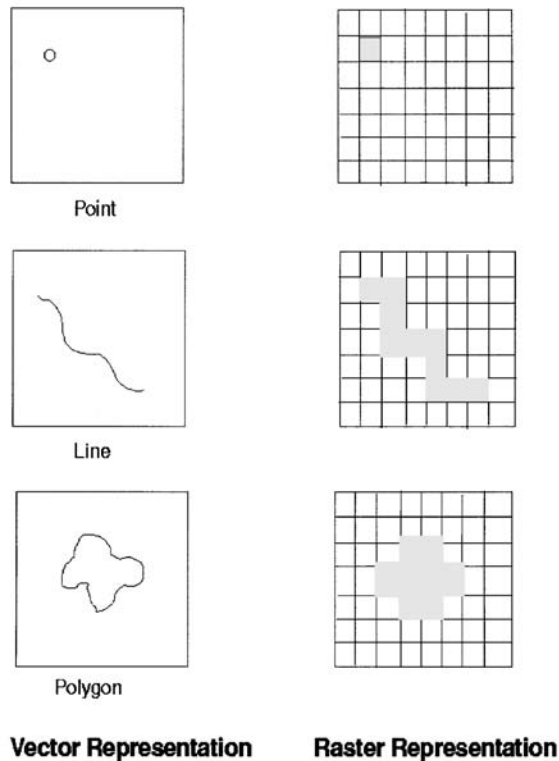
FIGURE 21.1 Vector and corresponding raster representation.

Raster data are continuous data and can be described as a matrix of columns and rows or as a regular grid of square or rectangular data. The location of each cell or pixel is defined by its row and column numbers. The value assigned to a grid or cell indicates the value of the attribute it represents. Figure 21.1 shows a comparison of vector and raster representation of GIS data sets. Satellite data are a typical example of raster data. Various graphic interchange formats for raster data are GIF, TIFF, BIL, BIP, JPEG, and ASCII. The choice between vector or raster GIS is usually driven by the type of application.

The most important aspects of data quality are accuracy, time, scale, and completeness of data sets. The error associated with each of the data sets will affect the accuracy of the final product and therefore defines the usefulness of the final products. Hence, GIS users are encouraged to document the data or create meta data for the data sets being developed. Meta data for a GIS data layer provide information that would usually include the scale of the data, its projection, how it was developed, and many other items of information useful in understanding the GIS data layer.

21.4 Analysis Capabilities of GIS

Analysis of spatial and attribute data in a GIS can be classified into five main types of procedures: (1) data transformation and restructuring, (2) data retrieval, classification, and measurement, (3) overlay, (4) neighborhood and statistical measures, and (5) connectivity. Each of these is discussed briefly in the remainder of this section.

Data transformation functions allow mathematical manipulation of raster or vector data to rotate, translate, scale, or mosaic different map layers. Such efforts are often necessary when acquiring GIS data sets. Transformation of data may be required to convert data being developed by digitizing or scanning into useful forms. Restructuring functions include data conversion from raster to vector, vector to raster, and converting data between different formats.

Retrieval functions allow the query and retrieval of spatial data and tabular information stored in the GIS database. They also facilitate spatial queries specific to a region on single or multiple map layers. Classification allows the user to rename attributes such as to indicate that areas within 0.5 miles of a well are sensitive areas and special precautions should be taken with potential contaminants in this area. Generalization allows the combining of classes such as all water table depths less than 15 feet can be considered shallow for a particular application. The data may have initially indicated water depths to the nearest foot. Measurement functions allow the user to determine distance from a feature, or length, area, volume, and perimeter of a feature or features.

Overlay functions constitute the third set of GIS analysis tools. Overlay allows the user to combine and overlay multiple thematic information (map layers) at different scales and in different formats. Overlay functions can also perform arithmetic operations such as addition, division, or multiplication on different map layers and logical operations between map layers such as AND, OR, and NOT. A typical logical operation might be finding areas with soils having high permeabilities and areas with shallow depths to water tables and defining areas having this combination as being susceptible to groundwater contamination.

The fourth set of analysis procedures are neighborhood functions that involve calculations based on the value of neighboring cells. For example, a search technique can be used to find the maximum contamination level in an area. The Thiessen polygon technique can be used to construct polygons around a set of points such as rain gauges or groundwater monitoring wells. Interpolation procedures such as inverse distance, kriging, and splines can be used for interpolating a value based on the values of neighboring locations. Such procedures are useful to create continuous surfaces from point data such as data observed in groundwater monitoring wells.

The fifth group of analysis procedures are connectivity functions such as contiguity measures that can be used to determine the size of a catchment. Proximity can be used to determine distance from a feature and generate buffer zones. Network functions can be used to predict water flow in a stream network,

and other connectivity functions can be used to define drainage basins, determine sinks in watersheds, and generate perspective views. Statistical measures allow calculation of spatial autocorrelation, correlation between maps, and confidence intervals.

21.5 Overview of Applications

GIS applications are commonly divided into the following categories:

- Facility and asset inventory
- Geographic data collection and production
- Map and chart publishing
- Resource allocation
- Route and flow optimization
- Route selection and navigation
- Site location and planning
- Surface and subsurface assessment
- Tracking and monitoring

GIS applications related to groundwater are found within several of these categories.

Typical examples of GIS applications in groundwater studies are site suitability analyses; mapping information such as water table depths, aquifer type and material, and aquifer recharge; managing site inventory data; computing statistics to estimate spatial correlation of a process occurring over a region such as nonpoint source pollution; estimating vulnerability of groundwater to pollution potential from nonpoint sources of pollution; integrating groundwater quality assessment models with spatial data to create spatial decision support systems; modeling groundwater movement: advection and dispersion modeling, developing flow field or vector fields representing groundwater seepage velocities, tracking particle movement, and particle retardation; modeling solute transport and leaching; and evaluating soil salinity and salt loading into groundwater.

21.6 Application of GIS to Groundwater Engineering

21.6.1 Mapping the Occurrence of Groundwater and Groundwater Features

Maps play an important role in portraying information about groundwater systems. GIS can be used to store these maps, allowing the user to create the desired map products. GIS further provide the ability to manipulate the mapped data.

Mapping the material of underlying aquifer systems requires information of geologic strata and other data including soil properties and depth to bedrock. Soil databases such as STATSGO contain information that can be used in conjunction with glacial geology to interpret the material of the aquifers. Figure 21.2 shows the aquifer media data layer developed from the glacial geology map of Indiana at a 1:100,000 scale.

Another application is monitoring the recharge flux to the water table. Recharge is a key component of the water balance and is critical for understanding groundwater systems. GIS can be utilized to evaluate and delineate patterns of recharge within a region by incorporating available information about observed temporal fluctuations of recharge at specified locations. Continuous remote acquisition of water table depths combined with GIS capabilities for distributing the flux spatially across the study region can result in near-real-time monitoring of spatial variability in recharge flux.[3] Recharge can also be estimated using information including current climate, soil, and vegetation/land use patterns and estimates of recharge for various soil-vegetation combinations.[4]
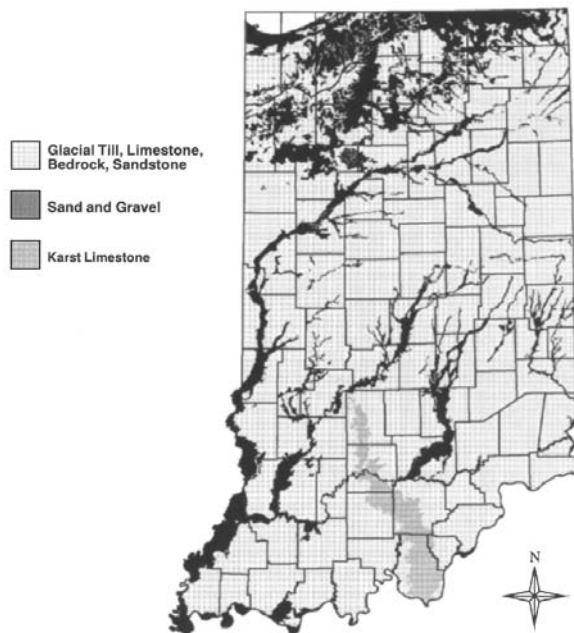
FIGURE 21.2 Aquifer material.

21.6.2 Managing Spatial Data Aspects of Groundwater Projects

Siting a waste disposal facility is an example for which GIS can help manage the spatial data considerations including the potential impacts on groundwater. The information on state and federal regulations for a site location can be implemented using GIS functions, overlays, buffers, and arithmetic computations on data layers. Regulations for a hazardous waste landfill indicate that they should not be located: (1) within 100 feet of any lake, reservoir or continuously flowing stream, (2) within 100 feet of any of the site's property boundaries, (3) within 600 feet of potable domestic water supplies, (4) within 600 feet of any dwelling, and (5) within 1200 feet of a public water supply. Such exclusionary areas are easily computed as buffers within GIS. Further, GIS maps of soils, geology, climate, hydrological, and cultural data sets can be integrated in a GIS to develop maps showing areas suitable for hazardous waste landfill siting that meet the above and other regulations. Figure 21.3 shows the areas suitable for hazardous waste landfill siting in Tippecanoe County, Indiana, that were identified using such an approach.[5]

21.6.3 Spatial Data Analyses Using Spatial Statistics

Spatial statistics are a rapidly emerging field being applied to many diverse applications.[6] Spatial autocorrelation studies are often used in geostatistics for determining the spatial dependencies of the factors being studied. In groundwater studies, spatial statistics can be computed to study the distributions of nonpoint source (NPS) contamination of groundwater on a regional scale, and spatial variation of hydrogeologic factors affecting the solute transport including hydraulic conductivity of the aquifer, slope, and groundwater properties.

Spatial autocorrelation refers to the spatial ordering of a single variable and to the relationship between pairs of observations of this variable. The ordering of n observed values of some variable X is usually described with the aid of a connectivity matrix, C. Non-zero c_{ij} entries in the $n \times n$ matrix indicate that the corresponding polygons are juxtaposed. For data measured on an interval/ratio scale, the Geary ratio and Moran coefficient statistics can be used[7] for assessing spatial correlation. These spatial statistics can be used to determine the spatial correlation of NPS pollution and to eliminate detections from point sources of pollution in a water quality database.

Unacceptable Sites
Poor Sites
Average Sites
Good Sites
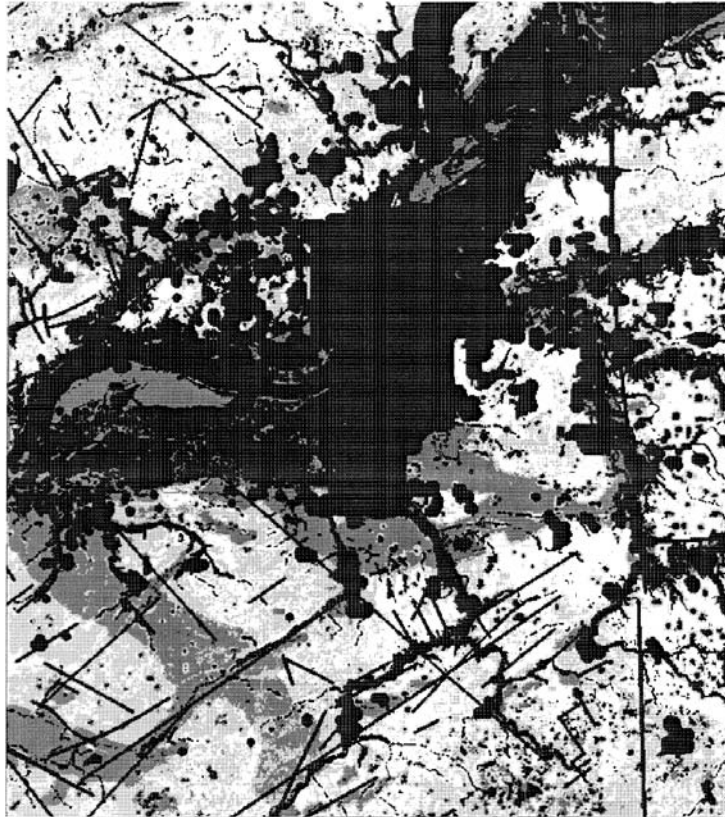
FIGURE 21.3 Suitability ratings of possible sites for hazardous waste landfills.

21.6.4 Surface Fitting and Interpolation

Surface fitting and interpolation techniques are commonly employed for creating continuous data (raster data) from a distributed set of data points over a geographical region (e.g., creating a depth to groundwater map from observations of water depth in wells). The underlying assumption in these techniques is that the data points are random, well-sampled, and representative of the spatial distributions of the attribute. Some of the commonly employed techniques for interpolation are the inverse distance weighting interpolation (IDW),[8,9] kriging using a semi-variogram, polynomial regression, and spline fitting.

Inverse distance weighting interpolation computes a cell value using a linear weighting combination of the set of sample points surrounding the cell. The weight is a function of inverse distance to the observed values. Usually a power of two is adopted for inverse distance, and values greater than two will increase the influence of nearby data. Also the interpolated value can be controlled by limiting the search radius or by limiting the number of sample points to be used for interpolation.

Kriging is based on the regional variable theory that assumes the spatial variation of an attribute is statistically homogeneous throughout the surface. The variation of the attribute is measured using a semi-variogram, a graphical plot of semi-variance against the distance between pairs of sample points. Using the surface estimators, ordinary kriging, and universal kriging, the value at a cell can be interpo-

lated. Polynomial regression interpolation depends on the order of polynomial used for fitting the surface. A first-order polynomial regression simply performs a least-squares fit of a plane to the set of data points. As the order of the polynomial increases, the interpolation becomes more complex.

Polynomial regression generally creates smooth surfaces, and the surface created seldom passes through the original data points. Spline interpolation is a two-dimensional minimum curve interpolation that creates a smooth surface, and unlike polynomial regression, the surface passes through the input data points. The commonly used minimum-curvature technique is also called thin plate interpolation. Spline interpolation ensures a smooth surface together with continuous first-derivative surfaces.

Functions for conducting surface fitting and interpolation are often included in many commercial GIS software.[8,9] These techniques are commonly used in geostatistics and can be applied in groundwater studies. The distribution of sample points will determine which method is appropriate. An example application of these techniques within GIS includes developing groundwater table depths for a region in unconfined aquifers using observed water table depths in drinking water wells. Figure 21.4 shows the map of water table depths in southern Indiana developed using the inverse distance weighting method. Another example is the study in northeastern Colorado, where inverse distance weighting interpolation was used for mapping spatial distributions of nitrate contamination in groundwater for the region.[10]
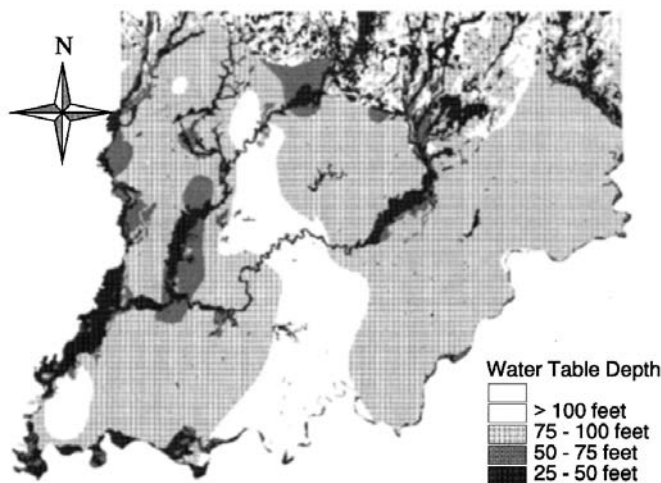
FIGURE 21.4 Interpolated water table depths based on water depths in wells.

21.6.5 Modeling Groundwater Vulnerability Using Spatial Data

GIS are increasingly being integrated with groundwater and surface water quality assessment models[11-17] to facilitate model use. The data required for these models can be stored and managed in a GIS to facilitate the model simulations.

GIS maps showing the areas vulnerable to groundwater contamination for a region would have many potential uses, such as implementation of groundwater management strategies to prevent degradation of groundwater quality and locating groundwater monitoring systems. These maps would also be helpful in evaluating existing and potential policies for groundwater protection. DRASTIC and SEEPAGE are simple techniques that can be used for developing regional scale groundwater vulnerability maps.

DRASTIC is a regional scale groundwater quality assessment model that estimates groundwater vulnerability of aquifer systems based on the hydrogeologic factors of a region.[18] It is an empirical model and was developed by the National Well Water Association in conjunction with the U.S. Environmental Protection Agency in the 1980s for evaluating pollution potential of groundwater systems on a regional scale. The factors considered in DRASTIC are:

- Depth to water table
- Recharge (aquifer)
- Aquifer media
- Soil media
- Topography
- Impact of vadose zone media
- Conductivity (hydraulic) of the aquifer

A region to be rated by DRASTIC is subdivided into smaller areas in which the factors considered are nearly homogeneous. A grid cell subdivision is often used and thus a raster GIS approach is well suited for conducting DRASTIC analyses. Each of the subdivisions of the factors considered in the model is weighted based on the factor values. The DRASTIC Index,[18,19] a measure of pollution potential, is computed by summing the weighted factors of each subdivision.

$$\text{DRASTIC Index} = D_r D_w + R_r R_w + A_r A_w + S_r S_w + T_r T_w + I_r I_w + C_r C_w$$

where
D_r = Ratings to the depth to water table
D_w = Weight assigned to the depth to water table
R_r = Ratings for ranges of aquifer recharge
R_w = Weight for the aquifer recharge
A_r = Ratings assigned to aquifer media
A_w = Weight assigned to aquifer media
S_r = Ratings for the soil media
S_w = Weight for soil media
T_r = Ratings for topography (slope)
T_w = Weight assigned to topography
I_r = Ratings assigned to vadose zone
I_w = Weight assigned to vadose zone
C_r = Ratings for rates of hydraulic conductivity
C_w = Weight given to hydraulic conductivity

The higher the DRASTIC index, the greater the relative pollution potential. Table 21.1 gives the typical weights for the DRASTIC factors. The DRASTIC Index can be converted into qualitative risk categories of low, moderate, high, and very high. See Table 21.2 for ranges used for classifying DRASTIC indices into various classes.

TABLE 21.1 DRASTIC Factor Weights

Feature	Weight
Depth to water	5
Net recharge	4
Aquifer media	3
Soil media	2
Topography	1
Vadose zone media	5
Hydraulic conductivity	3

Existing spatial databases like STATSGO (State Soils Geographic Database)[20] contain inputs required for these groundwater models. GIS can be used to develop and store data representing hydrogeologic settings of a region required for DRASTIC and SEEPAGE. GIS classification, overlay, and arithmetic operations can be used for integrating the data layers to develop maps showing areas vulnerable to

TABLE 21.2 DRASTIC Index Ranges for Qualitative Risk Categories

	DRASTIC Qualitative Category			
	Low	Moderate	High	Very High
DRASTIC Index	1–100	101–140	141–200	>200

DRASTIC Vulnerability Ratings

Low
Moderate
High
Very High

25   0   25   50 Miles

Developed at 1:250,000 Scale
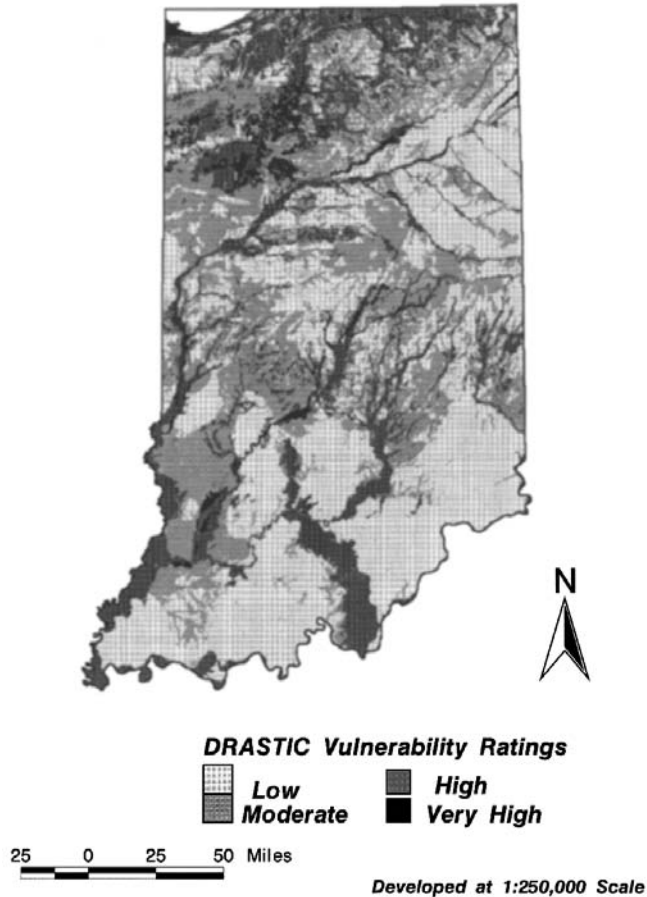
FIGURE 21.5 Groundwater contamination vulnerability as determined by DRASTIC analysis.

groundwater contamination from nonpoint source pollution. Figures 21.5 and 21.6 show the vulnerability of groundwater systems in Indiana as predicted by DRASTIC and SEEPAGE analyses. Using a GIS, a comparison of the predictions of DRASTIC and SEEPAGE analyses with a GIS-based water quality database of nitrate detections in drinking water wells from the U.S. Geological Survey (USGS) showed that 80% of nitrate detections > 2 ppm were within the DRASTIC highly and very highly vulnerable areas and 68% within SEEPAGE highly and very highly vulnerable areas.[21] These maps are useful as preliminary screening tools for policy and decision making in groundwater management strategies on a regional scale.

21.6.6 Bayesian Methods for Map Analysis

Statistical techniques used in conjunction with GIS show promise for various groundwater applications including identifying and assessing nitrate leached to groundwater on a regional scale.[10] The Bayesian approach uses prior and posterior probabilities of occurrence of an event for combining data sets.[22] The
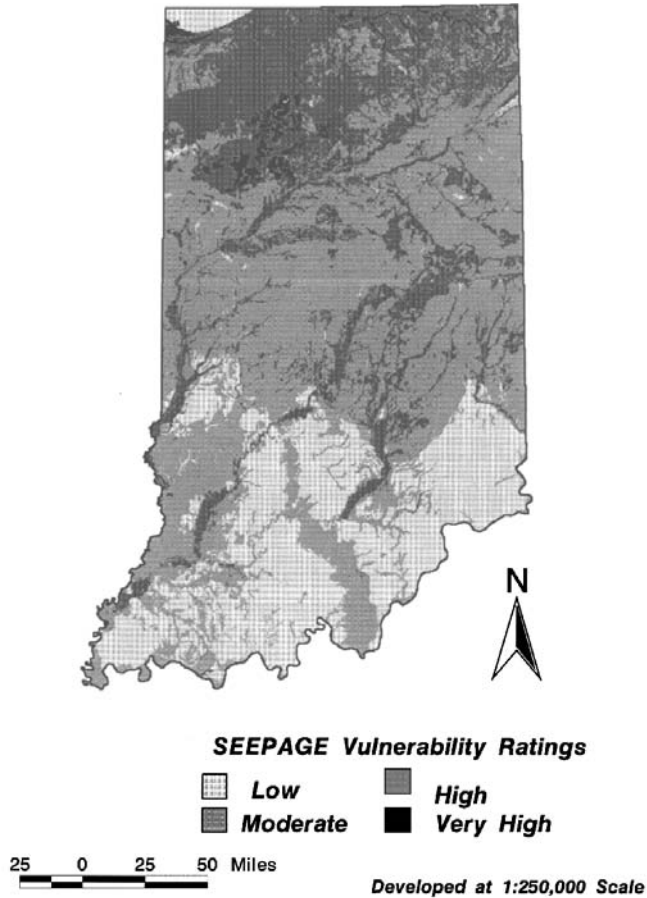
FIGURE 21.6 Groundwater contamination vulnerability as determined by SEEPAGE analysis.

posterior probability can be computed and updated based on the prior probability and addition of new evidence. The Bayesian model, in a log-linear form, known as the weights of evidence model, is a data-driven model that estimates the relative importance of evidence of data by statistical means. This method has been used for various purposes such as potential mineral mapping and determination of spatial distribution of seismic epicenters, and has been applied in diversified applications in medical diagnosis, geology, and other fields.[23] The Bayesian approach can be used to build a map showing areas vulnerable to groundwater contamination from nitrates and other pollutants using factors considered in DRASTIC or SEEPAGE and a water quality database of observed contamination levels in drinking water wells in a region. Figure 21.7 shows the Bayesian risk map for Indiana developed using the USGS water quality database and DRASTIC factors as evidence.[21] The data layers required for the analyses were integrated within the Arc/Info GIS.

21.6.7 Modeling Groundwater Movement

Groundwater flow modeling is usually completed using a two- or three-dimensional finite-difference or finite-element approach. In two dimensions, the finite-difference method employs a rectangular discretization in which input parameters are assigned to each pixel and hydraulic head is computed for each cell. A three-dimensional finite-difference can be visualized as a stack of grids with interaction between the grids as specified. Both common GIS data models, vector and raster, can be used to model groundwater movement using a finite-difference approach, but the grid/raster approach is more analogous to a
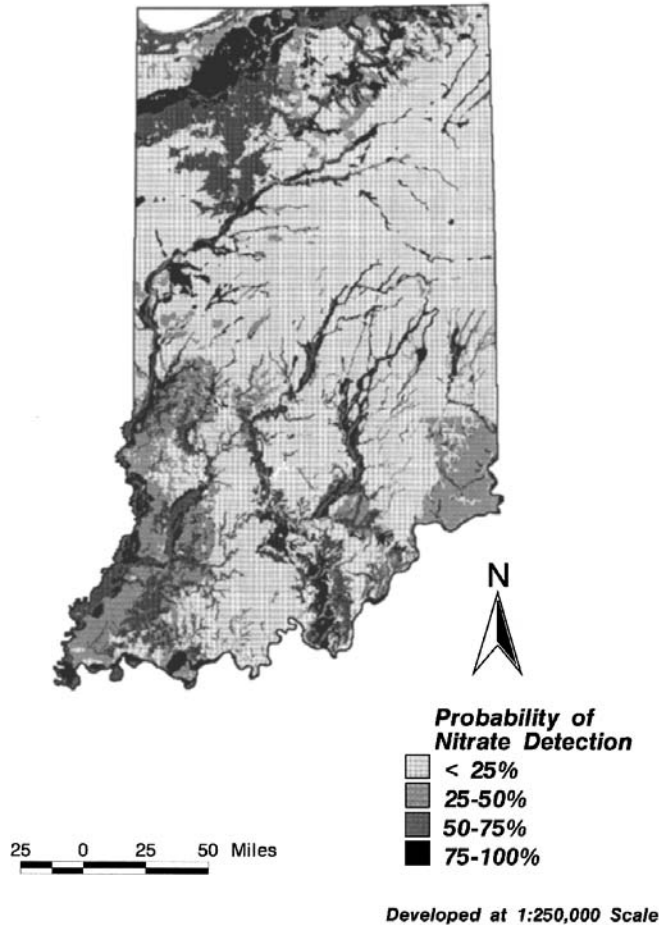
FIGURE 21.7 Bayesian risk map for Indiana.

finite-difference approach as each parameter/input into the model can be represented as a separate GIS data layer.[24]

In the finite-element approach, a triangular network is often used to represent two-dimensional flow. GIS are often equipped with functions for creating triangulated irregular networks (TIN), and this approach is similar to the finite-element approach. Hence, a vector data model with TIN capabilities may be more suitable for finite-element mesh generation for modeling groundwater movement.

GIS can facilitate the development, calibration, and verification of models as well as the display of model parameters and results in modeling groundwater movement. Spatial statistics and grid design capabilities of GIS can improve the modeling effort and aid in reliability assessment. A GIS-MODFLOW interface is an example of a groundwater finite-difference model linked with a GIS. Usually GIS and the model are autonomous and are linked by an interface or a set of programs. Typically, these programs simply convert the data from the GIS to a format that can be read by the model and the model output is converted for display in a GIS. GIS is primarily used for data management and spatial analysis capabilities in such interfaces.

Various functions are also generally available in a GIS that can be used for modeling groundwater advection and dispersion. The first step in groundwater flow modeling is to determine the flow velocity and direction at each point in the flow field. A flow field can be generated using Darcy's law for a regular grid. A volume balance can be computed for each cell. A zero volume balance residual might indicate a

balance between inflow and outflow in a cell. This represents the steady-state condition of groundwater flow.

21.6.8 Implementation of Regulations Involving Spatial Dimensions

Groundwater is clearly one of the most important natural resources. The government has mandated protection of these groundwater systems through the Safe Drinking Water Act,[25] which identifies and protects municipal wellfields by designating wellhead protection areas. A wellhead protection area (WHPA) is defined as the surface and subsurface areas surrounding a public well or wellfield through which contaminants are likely to move toward and reach the well or wellfield. Five general criteria in wellhead protection as identified by U.S. EPA are: (1) distance, (2) drawdown, (3) time of travel, (4) flow boundaries, and (5) assimilative capacity. A GIS can be used to integrate the above concepts and synthesize the inputs and outputs required for the analysis. The analytical capabilities of GIS software can be used for site characterization. Both the logical and visual overlay functions, as well as distance measurement capabilities in a GIS can be used to define conditions and locations meeting specified protection criteria.

GIS technology can be used for supporting remedial investigation of groundwater contamination. GIS can be used to integrate several software packages required for assessing environmental contamination during site characterization. Interfacing GIS with a groundwater movement model such as MODFLOW can facilitate investigation of three-dimensional plumes.[26]

GIS technology helps organize data about problems occurring over a geographic region, understanding their spatial association, and provides a powerful means of analyzing and synthesizing information on such problems. GIS can also provide decision makers information about the limitations of the accuracy of the data and the final products produced by GIS analyses.

For Further Information

Aronoff, S. 1989. *Geographic Information Systems: A Management Perspective*, WDL Publications, Ottawa, Canada.

Corwin, D.L., K. Loague, J.M. Bigham, D.M. Kral, and M.K. Viney. 1996. *Applications of GIS to the Modeling of Non-Point Source Pollutants in the Vadose Zone*, SSSA No. 48. Soil Science Society of America, Madison, WI.

Goodchild, M.F., B.O. Parks, and L.T. Steyart (editors). 1993. *Environmental Modeling with GIS*, Oxford University Press, New York.

Goodchild, M.F., L.T. Steyaert, B.O. Parks, C. Johnston, D. Maidment, M. Crane, and S. Glendinning (editors). 1996. *GIS and Environmental Modeling: Progress and Research Issues*, GIS World, Fort Collins, CO.

Heit, M. and A. Shortreid. 1991. *GIS Applications in Natural Resources*, GIS World, Fort Collins, CO.

Maguire, D., M. Goodchild, and D. Rhind (editors). 1991. *Geographical Information Systems*, John Wiley & Sons, New York.

References

1. Guptill, S. C., Cotter, D., Gibson, R., Liston, R., Tom, H., Trainor, T., and VanWyhe, H. P. A process for evaluating geographic information systems. Technical Report 88-105, U.S. Geological Survey, 1988.
2. Griffin, C. B. *Introduction to GIS for Public Agencies*. Technical Report, Conservation Technology Information Center, State University of New York, College of Environmental Science and Forestry, Syracuse, NY, 1995.
3. Keen, K. L. and Queen, L. P. An approach for continuously monitoring the recharge flux to the water table: Integration of field monitoring, laboratory characterization, and theory within GIS framework, in D.L. Corwin and R.J. Wagnet, Editors, *Applications of GIS to the Modeling of Non-Point Source Pollutants in the Vadose Zone.* 1996 Bouyaucos Conference, USDA-ARS, 1996.

4. Fayer, M. J., Gee, G. W., Rockhold, M. L., Freshley, M. D., and Walters, T. B. Estimating recharge rates for a groundwater model using GIS, in D.L. Corwin and R.J. Wagnet, Editors, *Applications of GIS to the Modeling of Non-Point Source Pollutants in the Vadose Zone.* 1996 Bouyancous Conference, USDA-ARS, 1996.

5. Levine, N. S. *The Development and Use of a GIS-Based Expert System for Hazardous Waste Landfill Siting.* Ph.D. Thesis, Purdue University, December, 1995.

6. Dou, C. and Woldt, D. E. Application of geostatistics for mapping nitrate contaminated groundwater. Number 94-2095, ASAE International Summer Meeting, 1994.

7. Griffith, D. A. *Spatial Auto-Correlation — A Primer.* Association of American Geographers, 1987.

8. ESRI. Redlands, CA. *ARC/Info User's Guide, Cell Based Modeling with GRID,* 1992.

9. Watson, D. F. and Philip, G. M. A refinement of inverse distance weighted interpolation. *Geo-Processing.* 2:315-327, 1985.

10. Follet, R. F., Shaffer, M. J., Brodahl, M. K., and Reichman, G. A. NLEAP simulation of residual soil nitrate for irrigated and non-irrigated corn. *Soil and Water Conservation.* 49(4):375-382, 1994.

11. Udoyara, S. T. and Jolly, R. Evaluating agricultural non-point source pollution using integrated geographic information systems and hydrologic/water quality model. *Environmental Quality.* 23:25-35, 1994.

12. Engel, B. A. Water quality modeling using Geographic Information System (GIS) data. Technical Report, John F. Kennedy Space Center, August, 1992.

13. Stallins, C., Huffman, R. L., Korram, S., and Guo, Z. Linking GLEAMS and GIS. Number 92-3613. ASAE International Winter Meeting, 1992.

14. Mamillapalli, S. and Engel, B. A. Use of GIS in running field scale models for watershed size areas. Number 94-3552. ASAE International Winter Meeting, 1994.

15. Wylie, B. K., Shaffer, M. J., Brodahl, M. K., Dubois, D., and Wagner, D. G. Predicting spatial distributions of nitrate leaching in northeastern Colorado. *Soil and Water Conservation.* 49(3):288-292. 1994.

16. Wylie, B. K., Shaffer, M. J., and Hall, M. D. Regional assessment of NLEAP NO3-N leaching indices. *Water Resources Bull.* 31(3):399-408, 1995.

17. Srinivasan, R. and Arnold, J. G. Integration of a basin-scale water quality model with GIS. *Water Resources Bull.* 30(3):453-462, 1994.

18. Aller, L., Bennett, T., Lehr, J. H., and Petty, R. J. DRASTIC: A standardized system for evaluating groundwater using hydrogeologic settings. Technical Report, USEPA EPA/600/2-85/0108, Robert S. Kerr Environmental Research Laboratory, Ada, OK, 1985.

19. Aller, L., Bennett, T., Lehr, J. H., Petty, R. J., and Hackett, G. DRASTIC: A standardized system for evaluating groundwater pollution using hydrogeologic settings. Technical Report, USEPA EPA-600/2-87-035, 1987.

20. SCS. State Soils Geographic Database (STATSGO) User's Guide. Technical Publication 1492, SCS, August, 1992.

21. Navulur, K. C. S. *Groundwater Vulnerability Evaluation to Non-Point Source Nitrate Pollution for Large Areas Using a GIS.* Ph.D. Thesis, Purdue University, Department of Agricultural and Biological Engineering, W. Lafayette, IN, December, 1996.

22. Bates, L. E., Barber, C., and Otto, C. J. Aquifer vulnerability mapping using GIS and Bayesian weights of evidence: Review of application, in *Hydro GIS.* Vienna, Austria, 1996.

23. Bonham-Carter, G. F. Geographic Information Systems for Geoscientists: Modeling with GIS. *Computer Methods in Geosciences.* 13:302-303, 1994.

24. Watkins, D. W., McKinney, D. C., Maidment, D. R., and Lin, M. Use of Geographic Information Systems in Ground-Water Flow Modeling. *Water Resources Planning and Management.* 122(2):88-96. 1996.

25. Gibas, D. R. and Barrett, D. P. A GIS-based aquifer characterization methodology for wellhead protection, in D.L. Corwin and R.J. Wagnet, Editors, *Applications of GIS to the Modeling of Non-Point Source Pollutants in the Vadose Zone.* 1996 Bouyaucos Conference, USDA-ARS, 1996.

26. Dale, S. and Barbato, L. GIS technology and a client-server environment for remedial investigation, in D.L. Corwin and R.J. Wagnet, Editors, *Applications of GIS to the Modeling of Non-Point Source Pollutants in the Vadose Zone.* 1996 Bouyaucos Conference, USDA-ARS, 1996.

Glossary

Buffer   A zone of a specified distance around coverage features. Buffers are useful for proximity analysis (e.g., finding all wetlands within 150 meters of streams).

Cell   Basic element of spatial information in a grid or raster data set.

Coordinate System   A system used to measure horizontal and vertical distances on a planimetric map. Its units and characteristics are defined by a map projection. A common coordinate system is used to spatially register geographic data for the same area.

Coverage   A set of thematically associated data considered as a unit. A coverage usually represents a single theme or layer such as roads, streams, land use, or aquifer depth.

Digitize   To encode map features as x, y coordinates in a digital form. A digitizer (a table device with a cursor and keys for tracing map features) is used to create GIS coverages.

Geographic Data   The locations and descriptions of geographic features. The composite of spatial data and descriptive data.

Geographic Database   A collection of spatial data and related descriptive data organized for efficient storage and retrieval.

Georeference   To establish the relationship between coordinates on a planar map and known real-world coordinates.

GIS   Geographic information system. An organized collection of computer hardware, software, geographic data, and personnel to efficiently capture, store, update, manipulate, analyze, and display geographically referenced information.

Intersect   Topological integration of spatial data sets. Common analysis function within GIS software.

Layer   A set of thematic data stored for use within GIS. Layers organize data by subject matter (e.g., aquifer depths, slope, roads).

Line   A set of ordered coordinates that represent the shape of linear features within a GIS coverage or layer. Data such as streams and roads are commonly represented as lines.

Map Projection   Conversion of locations on the earth's surface from spherical to planar coordinates. A projection is required to define a map's coordinate system.

Point   An x,y coordinate that represents a geographic feature. Data representation technique used within GIS to represent small features such as well locations.

Polygon   A multisided figure used to represent an area on a map. A polygon is defined by a series of arcs. Data such as the locations of soil series are commonly represented as polygons within a GIS coverage or layer.

Raster   A cellular or grid data structure composed of rows and columns. Groups of cells represent features such as land use or elevation. The value of each cell represents the value of the feature. Image and remotely sensed data are commonly stored in this format.

Resolution   Resolution is the accuracy at which a map scale can depict the location and shape of map features. As map scale decreases, resolution diminishes and feature boundaries must be smoothed, simplified, or not shown at all.

Scanning   Data input in raster format with a device called a scanner. Used to convert paper-based maps into GIS coverages or layers.

Spatial Analysis   The process of modeling, examining, and interpreting model results. Spatial analysis is the process of extracting or creating new information about a set of geographic features. Spatial analysis is useful for evaluating suitability and capability, for estimating and predicting, and for interpreting and understanding.

Spatial Data   Information about the location, shape, and relationships among geographic features usually stored as coordinates and topology.

Surface   Representation of geographic information as a set of continuous data in which the map features are not spatially discrete. Surfaces are used to represent spatial features such as depth to groundwater and are generally created by fitting a surface to point data.

Topology   The spatial relationships between connecting or adjacent coverage features. Topology is useful in GIS because many spatial analysis functions do not require coordinates, only topological information.

22

U.S. Laws and Regulations

John A. Veil,
Deborah Elcock,
Nancy L. Ranek, and
David R. Green
Argonne National Laboratory

22.1 Introduction

This chapter describes the four major federal environmental laws that place controls on the management of groundwater and the relevant regulations that have been developed based on the statutory authority of those laws. These are:

- The Clean Water Act (CWA), based on the Federal Water Pollution Control Act (FWPCA) Amendments of 1972, Public Law 92-500, 42 U.S.C. §1251 *et seq.**
- The Safe Drinking Water Act (SDWA), Public Law 93-523, 42 U.S.C. §300f *et seq.*
- The Resource Conservation and Recovery Act of 1976 (RCRA), Public Law 94-580, 42 U.S.C. 6901 *et seq.*
- The Comprehensive Environmental Response, Compensation, and Liability Act of 1980 (CERCLA or Superfund), Public Law No. 96-510, 42 U.S.C. §9601 *et seq.*, as amended.

To meet the needs of the readers of this book, the laws and regulations affecting groundwater quality and protection will be summarized as concisely as possible while still conveying the important issues. We strongly recommend consultation with the laws and regulations themselves before making any judgments or interpretations. Readers can obtain the laws and regulations from federal and state agencies and can access them through numerous online services or through the Internet.

Another important component of groundwater regulation is in the area of groundwater quantity. Historically, it has been the role of the states to determine who could withdraw groundwater and how much could be withdrawn. State water quantity laws date back many years and can be very complicated. Water withdrawal rights in arid western states are closely guarded. Even in some states with bountiful water resources, persons wanting to withdraw groundwater for drinking, industrial, agricultural, or other uses must obtain permits from state agencies. This chapter does not attempt to summarize the web of water quantity requirements. We do urge any reader who must deal with water quantity issues to consult the laws and regulations of the state in which he or she is interested for the details.

22.1.1 Relationship Between Laws and Regulations

Many persons who do not regularly work with legal requirements do not understand the relationship between laws and regulations. Laws are passed by legislative bodies such as Congress or state legislatures. Normally, laws outline general requirements without going into great detail on how to accomplish the requirements. Regulatory agencies, such as the U.S. Environmental Protection Agency (EPA) or state and local environmental, health, or natural resources agencies, are responsible for interpreting the general mandates of the laws and implementing them through regulations. The regulations are issued by the regulatory agencies following a lengthy administrative procedure involving opportunity for public comment.

22.1.2 Relationship Between Federal, State, and Local Laws and Regulations

Groundwater protection efforts are managed by all three levels of government. The federal laws and regulations establish basic frameworks and minimum national standards. Congress designed many of the regulatory programs so that they can be delegated to states, if the states are interested and can demonstrate that they have an adequate legal basis under state authority to administer the programs. Other responsibilities, such as funding programs, conducting large research programs, and preparing baseline regulatory criteria, are left to the federal government. When states assume federal programs, their programs must be at least as strict as the federal programs and can be stricter. In addition to assuming federal programs, states can adopt laws and regulations that create and implement other

---

* For readers not familiar with legal nomenclature, this is the reference for section 1251 and the following sections in volume 42 of the United States Code, the body of U.S. laws.

groundwater protection programs that are not considered under the federal laws. For example, many states operate groundwater discharge permit programs which cover more discharges than the comparable federal program, the Underground Injection Control (UIC) program.

Local governments also play a role in groundwater protection. Local health departments frequently represent the first line of defense against unsound septic system installation and operation. Local and state governments are the driving force behind wellhead protection programs that are geared to prevent pollution from entering a drinking water aquifer. Under RCRA, local governments must protect groundwater by complying with Federal and State regulations applicable to municipal solid waste landfills (MSWLFs).

22.1.3 Organization of Chapter

The remainder of this chapter discusses the four main federal environmental laws that play a role in groundwater protection — CWA, SDWA, RCRA, and CERCLA. For each law, information is presented on historical perspectives, goals and objectives, and general provisions. Under each law, a summary is provided of the regulations that protect groundwater quality, along with a discussion of the issues and outlook for changes to the law and the regulations. A final section discusses the overlap between RCRA and CERCLA environmental cleanup programs.

22.2 CWA Legislative Background

22.2.1 Historical Perspective

The CWA was passed in 1972 as a comprehensive package of amendments to the federal Water Pollution Control Act. Although the name "Clean Water Act" was not applied to this law until 1977 when another set of amendments was passed, the 1972 amendments serve as the federal basis for water quality protection and water pollution control. The CWA was passed in response to the country's growing awareness of water quality degradation in its rivers, lakes, and oceans. One commonly cited impetus for congressional action was the Cuyahoga River in Cleveland, which actually caught fire due to floating oils and other flammable contaminants. Given the nation's greater awareness of surface water and its problems, nearly all of the CWA focuses on surface water issues. Major amendments to the law in 1977 and 1987 did little to change this focus. Congress has been working to reauthorize and amend the CWA each year since 1991, but no additional changes have been passed yet. By many accounts, the CWA has been the most successful federal environmental law. Water quality has improved in hundreds of water bodies across the nation.

22.2.2 Goals and Objectives

The stated objective of the CWA is "to restore and maintain the chemical, physical, and biological integrity of the Nation's waters." The term "Nation's waters" is never defined, but the term "navigable waters" is used throughout the CWA. Navigable waters are defined as "waters of the United States," which includes surface water but not groundwater. None of the goals listed in §101 of the CWA directly involves groundwater. However, the next section of the CWA, §102(a) states that EPA "shall, after careful investigation … prepare or develop comprehensive programs for preventing, reducing, or eliminating the pollution of the navigable waters and groundwaters and improving the sanitary condition of surface and underground waters." However, no CWA regulatory programs have been enacted to implement this section.

There are several other references to groundwater in the CWA. §104 directs EPA to maintain a groundwater monitoring network. §208 directs states to identify areas that have substantial water quality problems and develop area-wide waste treatment management plans for these areas. These plans must include processes to identify contaminated underground mine drainage and salt water intrusion due to

excessive groundwater extraction, and a process to control disposal of pollutants on the land or in subsurface excavations to protect groundwater quality.

§304 requires EPA to develop water quality criteria after considering the presence of pollutants in groundwater and to publish information on the factors necessary to restore and maintain the chemical, physical, and biological integrity of all navigable waters and groundwater. §304 also directs EPA to issue processes, procedures, and methods to control pollution resulting from, among other sources: underground mining activities; the disposal of pollutants in wells or in subsurface excavations; salt water intrusion resulting from reductions of fresh water flow from any cause, including extraction of groundwater; and changes in the movement, flow, or circulation of any navigable waters or groundwaters. EPA has fulfilled some of these obligations through other programs, particularly those under the SDWA, but has not developed a CWA program to address all of the obligations. The only actual program for protecting groundwater is found in §319; this is discussed below.

Although groundwater is not mentioned in §106, grant funding through this section has been used by states to develop their state groundwater protection and wellhead regulations.

22.2.3 General Provisions

The primary regulatory emphasis of the CWA is control of point source discharges (mainly, discharges from industries and sewage treatment plants such as publicly owned treatment works — POTWs) through the National Pollutant Discharge Elimination System (NPDES) program (§402). Congress also established a comparable pretreatment program for industries that discharge wastewater to POTWs rather than directly to navigable waters.

Nonpoint sources of pollution are not subject to a formal federal regulatory program such as the NPDES program. Under §319, states are required to submit nonpoint source management programs to EPA for approval. The state programs must outline the measures that will be used to reduce nonpoint source pollution and the regulatory or nonregulatory programs that will achieve the necessary measures. The programs would include enforcement, technical and financial assistance, education, training, technology transfer, and demonstration projects.

§319 is one of the few sections of the CWA that refers to groundwater by name and attempts to protect groundwater. State management programs are supposed to take into account the impact of their recommended practices on groundwater quality (§319(b)(2)(A)). When making federal nonpoint source grants, EPA must give priority to projects that will carry out activities to protect groundwater from nonpoint sources (§319(h)(5)(D)). Further, EPA may make grants to states to assist in carrying out groundwater protection activities that will assist in implementation of a comprehensive nonpoint source program (§319(i)). Unfortunately, nonpoint source control programs have not been widely successful. This may be attributable to the lack of a strong federal program, insufficient levels of federal financial support, and the difficulty of finding a responsible and financially viable party like an industry or a POTW to carry out the control program. Historically, only limited water pollution controls have been placed on the agricultural community, which is perhaps the largest contributor to nonpoint source pollution. From a groundwater protection standpoint, §319 held much promise, but little actual improvement has been realized.

Although the federal NPDES program is limited to discharges to surface water, many state water pollution control laws are broader in that they mandate regulatory programs to protect both the surface water and groundwater of their state. For example, some states require state discharge permits for discharges of wastewater not only to surface water but also to groundwater or the land surface. Septic systems for disposal of sewage or household wastes are not regulated at the federal level through the CWA, but are controlled by state and local regulations. The details of these state programs vary widely, so readers are advised to consult with the appropriate state regulatory agencies.

22.2.4 Issues and Outlook

Congress has attempted to reauthorize and amend the CWA for the past five years but has been unable to pass a bill through both the House and Senate. The only changes to the CWA that are likely to place any significant requirements on groundwater following reauthorization are expanded nonpoint source controls and watershed management. Nonpoint sources contribute to pollution in surface waters but also are responsible for groundwater pollution. As described above, the CWA already envisions that state nonpoint source control programs will protect groundwater. Any legislative change is likely to strengthen the requirements to protect both surface waters and groundwater from nonpoint source impacts.

The concept of watershed management involves consideration of all activities occurring within a watershed rather than focusing on a single regulatory program. A well-operated watershed management program would consider impacts of various activities on both groundwater quantity and quality. EPA has begun to encourage watershed management programs, but these programs would have greater success and clout with full statutory authority to back them up.

The outlook for significant amendments to the CWA is not clear. As of early 1998, Congress had placed a low priority on amending the CWA.

22.3 CWA Regulations

All federal regulations are codified in the Code of Federal Regulations (CFR). Readers can find the CFR in most libraries and can access it through numerous online services or through the Internet. The CWA regulations do not place regulatory requirements directly on groundwater. The storm water regulations (40 CFR 122.46*) could have an indirect impact on groundwater. They require that storm water discharges associated with industrial activity, from construction sites, and from large and medium municipal separate storm sewer systems must obtain NPDES permits. In most instances, storm water permits do not contain numerical limits by which to measure compliance. Instead, they require that permittees install or follow best management practices (BMPs). There are three main categories of BMPs required or recommended by general permits: (1) those that keep clean storm water from coming into contact with contaminant sources (e.g., berms, curbing, roofing); (2) those that treat the contaminated runoff to lower pollutant levels (e.g., oil/water separators, straw bale dikes, sediment ponds); and (3) those that reduce the total volume or rate of runoff to surface waters (e.g., detention or retention ponds, storm water drainage wells).

The third category of BMPs represents a very effective mechanism for controlling surface water contamination but can contribute to groundwater contamination. Dissolved pollutants contained in the storm water can infiltrate through the bottom of detention or retention ponds and enter the groundwater. Disposal of storm water into drainage wells introduces pollutants directly into the groundwater.

Nonpoint source storm water discharges contain the same contaminants as do point source storm water discharges. EPA has not adopted nonpoint source regulations, but existing nonpoint source control programs rely almost entirely on BMPs. Consequently, the same potential for groundwater contamination exists.

One other regulatory area under the CWA that could indirectly affect groundwater is the sewage sludge regulations (40 CFR 503). Two of the approved methods for sewage sludge management are land application and surface disposal. The land application and surface disposal regulations do not contain direct groundwater protection requirements, but the application rates and specified management practices are designed to protect surface water (and indirectly groundwater) from excessive land application of sludge contaminants.

---

* 40 CFR 122.46 refers to Part 122.46 of Title 40 of the Code of Federal Regulations

22.4 CWA Summary

The CWA has little to do with groundwater protection. Only a few sections of the CWA actually mention groundwater, and the only existing part of the CWA that has some serious potential to place groundwater protection requirements on landowners, businesses, farmers, or other dischargers is §319, the nonpoint source control program.

22.5 SDWA Legislative Background

22.5.1 Historical Perspective

The SDWA was passed in 1974 to establish a national framework for ensuring a clean drinking water supply. The SDWA establishes two main programs: (1) the public water system program, which specifies how clean drinking water must be when it is delivered to customers; and (2) programs to protect underground sources of drinking water (USDWs) from potentially contaminating activities. Minor amendments were made to the SDWA in 1976, 1977, and 1980, but the first significant changes were adopted in 1986. The 1986 amendments greatly expanded the scope of contaminants for which drinking water standards would be developed and increased the level of monitoring for nonregulated contaminants. States were required to establish wellhead protection areas.

In 1996, Congress passed comprehensive amendments to the SDWA. Many of the amendments affected the public water supply system requirements through changes in how drinking water standards are established and implemented. Consideration was given for small public systems with limited economic resources. Much greater emphasis was placed on protecting source water. The new source water requirements are likely to improve groundwater quality.

22.5.2 Goals and Objectives

The general goal of the public water system program is ensuring that drinking water at the tap is clean and safe to drink. This is accomplished through a combination of treatment requirements (e.g., filtration, disinfection) and numerical quality standards for individual contaminants. The SDWA focuses solely on public water systems, which are defined as those that have at least 15 service connections or that serve at least 25 persons. The SDWA does not require testing, controls, or standards for those drinking water systems that supply fewer than the specified number of connections or persons (e.g., private wells). The general goal of the remaining SDWA programs is to protect USDWs from contamination. This is accomplished through restricting activities that can take place in sensitive areas and carefully controlling discharges of contaminants to groundwater or to the subsurface soils. The SDWA does not contain any program specifically requiring protection of surface water supplies. Under the SDWA, not all groundwater must be protected; protection is required only for USDWs.

22.5.3 General Provisions

All major SDWA programs have some effect on groundwater. The heart of the public water system program is the national primary drinking water regulations. EPA has published maximum contaminant levels (MCLs) and maximum contaminant level goals (MCLGs) for more than 75 contaminants (see Table 22.1). The MCLG represents the desired concentration goal, in terms of health protection, for drinking water supplied to customers. For many contaminants, the MCLG cannot be economically or technically achieved. The MCL is the highest concentration of a contaminant that is allowed for water delivered to customers. The MCL is often less stringent than the MCLG. In some cases, EPA may establish treatment requirements to control a particular contaminant in lieu of establishing a numerical MCL. The SDWA provides a mechanism for public drinking water systems to obtain variances and waivers from meeting MCLs or treatment requirements under certain circumstances. The MCLs described above

are enforceable; EPA also has adopted secondary MCLs for 16 drinking water contaminants or constituents. Secondary MCLs are not enforceable but are intended to protect the aesthetic qualities of drinking water.

The public water system requirements under the SDWA do not directly control groundwater quality since they are designed to protect drinking water consumers and not the groundwater itself. A water supplier can withdraw groundwater that does not meet all MCLs and treat it prior to supplying it to customers. However, MCLs are important for several regulatory reasons. The Underground Injection Control (UIC) program under the SDWA frequently requires that MCLs be met at the point of waste injection, at the point where the discharge plume intersects the USDW, or the point at which the USDW leaves the property boundaries. Under CERCLA, the MCLs are often taken to represent appropriate cleanup standards. This is discussed in a later section of this chapter.

The SDWA contains several important USDW protection programs. The first of these is the UIC program, which regulates injection of fluids into wells (§§1421-1424). The SDWA envisions that states will have the opportunity to administer the UIC program. Many states have been granted UIC program authority for some or all classes of injection wells. EPA or state UIC programs, which authorize injections through permit or rule, must ensure that injection wells do not endanger drinking water sources.

The second SDWA program dealing with USDW protection is the sole source aquifer program, under which EPA can designate certain aquifers as the sole or principal drinking water source for an area (§1424(e) and §1427). When an aquifer is designated as the sole source aquifer, EPA may not commit federal financial assistance (through a grant, contract, loan guarantee, or otherwise) for any project which EPA determines may contaminate such aquifer through a recharge zone so as to create a significant hazard to public health. As of September 1996, EPA had designated 66 aquifers as "sole source" under this provision.

State and local governments and regional planning entities may apply to EPA to have certain critical aquifer protection areas under their jurisdiction selected for funding of demonstration programs. In selecting critical aquifer protection areas EPA must consider: (1) the vulnerability of the aquifer to contamination; (2) the population using the aquifer as their drinking water source; (3) the economic, social, and environmental benefits resulting from protecting the aquifer; and (4) the economic, social, and environmental costs resulting from aquifer degradation.

The application for a demonstration program must include a comprehensive management plan for the proposed protection area. The plan would include, among other items: (1) an identification of existing and potential point and nonpoint sources of groundwater degradation; (2) an assessment of the relationship between activities on the land surface and groundwater quality; (3) specific actions and management practices to be implemented in the critical protection area to prevent adverse impacts on groundwater quality; (4) a determination of the quality of the existing groundwater recharged through the special protection area and the natural recharge capabilities of the special protection area watershed; (5) requirements designed to maintain existing underground drinking water quality or improve underground drinking water quality if prevailing conditions fail to meet drinking water standards; and (6) actions in the special protection area which would avoid adverse impacts on water quality, recharge capabilities, or both.

The third SDWA program dealing with USDW protection is the wellhead protection program, under which states are directed to develop programs to protect the areas around water supply wellheads from contaminants (§1428). Wellhead protection programs must delineate the size of the protection area for each wellhead and the activities that will be undertaken to protect the water supply within wellhead protection areas from contaminants. The federal requirements for the wellhead protection program are concerned with setting up the framework for the program and providing funding. The actual regulatory requirements and controls are administered at the state or local level.

The final SDWA programs dealing with USDW protection were new in 1996. §1429 authorizes EPA to make grants to states for developing and implementing comprehensive programs for groundwater protection. Two other source water protection programs are geared to protect both surface and groundwater drinking water supplies. States are required to develop and implement source water assessment

programs (§1453). State source water assessment programs must delineate the boundaries of assessment areas, identify contaminants that may present a public health threat, and identify the origins of those contaminants.

States may establish programs under which public water systems or local governments may submit source water quality protection partnership petitions to request assistance in developing voluntary, incentive-based partnerships to reduce the presence of contaminants in drinking water and to provide for long-term source water protection (§1454).

22.5.4 Issues and Outlook

The SDWA was reauthorized and amended in mid-1996. The information presented here reflects those amendments.

In June 1996, EPA published a "National Drinking Water Program Redirection Strategy." The proposal was based on four major objectives: sound science and adequate data; risk-based priorities for setting high quality standards; strong, flexible partnerships with states and local governments in implementation; and community-based, effective source water protection. The proposal was an attempt to focus on the activities that would provide the greatest health protection for the least cost. Under the specific details of EPA's proposal, the agency planned to scale back its efforts in the UIC and wellhead protection programs and delegate more of these activities to states.

22.6 SDWA Regulations

EPA has published extensive regulations for the public drinking water program in 40 CFR 141-143. As discussed above, most of the drinking water requirements are not directly relevant to groundwater protection. Since this chapter is concerned with groundwater issues, not drinking water issues, only the MCLs and MCLGs are discussed here. Many parts of the drinking water regulations were adopted at different times; therefore, not all MCLs or MCLGs are found in the same section of the regulations. MCLs are found in 40 CFR 141.11, .12, .13, .15, .16, .61, .62, and .63. MCLGs are found in 40 CFR 141.50, .51, and .52. All promulgated MCLs and MCLGs are listed in Table 22.1. EPA is presently working on several proposed or final regulations that will add or revise MCLs and MCLGs.

EPA has published detailed regulations for much of the UIC program in 40 CFR 144-148. Any underground injection, except as authorized by permit or rule issued under the UIC program, is prohibited by 144.11. Further, 144.12 specifies that: "no owner or operator shall construct, operate, maintain, convert, plug, abandon, or conduct any other injection activity in a manner that allows the movement of fluid containing any contaminant into underground sources of drinking water, if the presence of that contaminant may cause a violation of any primary drinking water regulation under 40 CFR Part 142 or may otherwise adversely affect the health of persons."

Part 146.4 allows for an aquifer to be considered an exempted aquifer if: (1) it does not currently serve and cannot in the future serve as a USDW by virtue of its depth, location, existing contamination, or its use for mineral or hydrocarbon extraction; and (2) the total dissolved solids content of the groundwater is more than 3000 and less than 10,000 mg/l and it is not reasonably expected to supply a public water system.

Part 146.5 establishes five classes of injection wells. Class I wells are used for disposal of hazardous or nonhazardous wastes below the lowermost USDW. Parts 146.11 – .15 and .61 – .73 contain detailed requirements for siting, constructing, operating, monitoring, and closing Class I wells to ensure that contaminants are unable to migrate into a USDW. Parts 144.60 – .70 define the financial responsibilities for Class I well owners and operators. There are over 500 Class I wells in 22 states, particularly in the Gulf Coast region.

Class II wells are used for injection of oil and gas industry wastes, injection of fluids for enhanced oil recovery, or underground storage of liquid hydrocarbons. Parts 146.21 – .25 contain detailed

TABLE 22.1 EPA's National Primary Drinking Water Standards

Contaminant	MCLG (mg/l)	MCL (mg/l)
Inorganics		
Antimony	0.006	0.006
Asbestos (>10 um)	7 MFL[a]	7 MFL
Barium	2	2
Beryllium	0.004	0.004
Cadmium	0.005	0.005
Chromium (total)	0.1	0.1
Copper	1.3	TT[b]
Cyanide	0.2	0.2
Fluoride	4	4
Lead	0	TT
Mercury (inorganic)	0.002	0.002
Nitrate	10	10
Nitrite	1	1
Selenium	0.05	0.05
Thallium	0.0005	0.002
Coliform and Surface Water Treatment		
Giardia lambia	0 detected	TT
Legionella	0 detected	TT
Standard plate count	N/A	TT
Total coliform	0 detected	[c]
Turbidity	N/A	TT
Viruses	0 detected	TT
Organics		
Acrylamide	0	TT
Adipate (di-2-ethylhexyl)	0.4	0.4
Alachlor	0	0.002
Aldicarb	0.001	0.003
Aldicarb sulfone	0.001	0.002
Aldicarb sulfoxide	0.001	0.004
Atrazine	0.003	0.003
Benzene	0	0.005
Benzo-a-pyrene	0	0.0002
Carbofuran	0.04	0.04
Carbon tetrachloride	0	0.005
Chlordane	0	0.002
Chlorobenzene	0.1	0.1
2,4-D	0.07	0.07
Dalapon	0.2	0.2
Dibromochloropropane	0	0.0002
0-dicholorobenzene	0.6	0.6
p-dichlorobenzene	0.075	0.075
1,2-dichloroethane	0	0.005
1,1-dichloroethylene	0.007	0.007
cis-1,2-dichloroethylene	0.07	0.07
trans-1,2-dichloroethylene	0.1	0.1
Dichloromethane	0	0.005
1,2-dichloropropane	0	0.005
Dinoseb	0.007	0.007
Diquat	0.02	0.02
Dioxin	0	0.00000003
Endothall	0.1	0.1
Endrin	0.002	0.002
Epichlorohydrin	0	TT
Ethylbenzene	0.7	0.7
Ethylene dibromide	0	0.00005
Glyphosate	0.7	0.7

TABLE 22.1 (continued) EPA's National Primary Drinking Water Standards

Contaminant	MCLG (mg/l)	MCL (mg/l)
Heptachlor	0	0.0004
Heptachlor epoxide	0	0.0002
Hexachlorobenzene	0	0.001
Hexachlorocyclopentadiene	0.05	0.05
Lindane	0.0002	0.0002
Methoxychlor	0.04	0.04
Oxamyl	0.2	0.2
PCBs	0	0.0005
Pentachlorophenol	0	0.001
Phthalate (di-2-ethylhexyl)	0	0.006
Picloram	0.5	0.5
Simazine	0.004	0.004
Styrene	0.1	0.1
Tetrachloroethylene	0	0.005
Toluene	1	1
Toxaphene	0	0.003
2,4,5-TP	0.05	0.05
1,2,4-trichlorobenzene	0.07	0.07
1,1,1-trichloroethane	0.2	0.2
1,1,2-trichloroethylene	0.003	0.003
Trichloroethylene	0	0.005
Total trihalomethanes	0	0.1
Vinyl chloride	0	0.002
Xylenes (total)	10	10
Radionuclides		
Radium 226 + radium 228	0	5 pCi/l
Gross alpha particle activity	0	15 pCi/l
Beta particle + photon radioactivity	0	4 mrem/yr

[a] MFL = million fibers per liter

[b] TT = treatment techniques in lieu of a numerical standard

[c] no more than 5% of samples can be total coliform-positive

requirements for construction, monitoring, operations, and closure. There are approximately 170,000 Class II wells in 31 states.

Class III wells are those used for mineral extraction. There are two primary mining processes that use Class III wells — solution mining and *in situ* leaching. Parts 146.31 – .35 contain detailed requirements for construction, monitoring, operations, and closure. There are approximately 21,000 Class III wells in 16 states, primarily in the south-central and southwestern states. Under federal UIC regulations, all Class I, II, or III wells must be permitted by the EPA or a delegated state (144.31).

Class IV wells are those in which hazardous or radioactive wastes are injected into or above a USDW (144.13). Class IV wells are considered to be a threat to USDWs and are prohibited, except for wells used as part of a CERCLA or RCRA cleanup, or injection made to an exempted aquifer. If a Class IV well is identified, it must be closed and may possibly be subject to remediation.

The final class of injection wells is not a specific type of well but a catchall for all injection wells not falling into one of the other classes. Class V wells include drywells, sumps, drain fields, drainage wells, and some septic systems, among others. Class V wells have been widely used in many parts of the country, but they have not typically been subject to extensive permitting requirements. Most Class V wells are drainage wells or wells involving sewage disposal (i.e., septic tanks and drainfields, cesspools). EPA estimates that there are approximately 1 million Class V wells in the United States, although it is probable that many more Class V wells have not yet been identified or reported to regulators.

Currently, Class V wells are authorized by rule (approval for injection is given through blanket regulatory approval), and EPA has promulgated no regulations governing Class V well operation, although

EPA can require a permit for individual wells under certain circumstances (144.25). Class V well operators must submit an inventory of their wells to EPA or a delegated state agency (144.26). On August 28, 1995 (60 FR 44652)*, EPA proposed new regulations for Class V wells. Because of the large number and diverse nature of Class V wells, EPA did not propose permit requirements. Instead, most Class V wells would continue to be authorized by rule and those wells that are most likely to endanger USDWs would be closed. EPA plans to publish proposed Class V regulations in 1998.

Parts 144.51 — .55 contain various requirements and conditions that are applicable to all UIC permits.

EPA established procedures whereby a state can be delegated authority to administer the UIC program for one or more classes of wells. Part 145 outlines the specific requirements for state delegation. Part 147 lists for each state whether the state or EPA administers the UIC program. Readers should consult Part 147 to determine which agency is responsible for UIC activities in a particular state.

Part 148 identifies hazardous wastes that are restricted from disposal into Class I hazardous waste injection wells and defines those circumstances under which a waste that is otherwise prohibited from injection may be injected. These regulations actually implement §3020 of RCRA, but are discussed here because EPA's UIC regulatory program covers both SDWA and RCRA UIC requirements.

For the sole source aquifer program, the federal regulations are minimal since the language in the SDWA is already rather specific. Part 149.3 defines a critical aquifer protection area as: (1) an area designated as a sole or principal source aquifer prior to June 19, 1986, and for which an area-wide groundwater quality protection plan was approved; or (2) a major recharge area of a sole or principal source aquifer, designated before June 19, 1988, for which contamination is likely to occur without a control program.

Parts 149.100 -.110 describe federal requirements for protecting the Edwards Underground Reservoir, which is the sole or principal drinking water source for the San Antonio area and which, if contaminated, would create a significant hazard to public health. This is the only aquifer for which EPA has published specific requirements. EPA's Office of Ground Water and Drinking Water periodically publishes a list of the sole source aquifers in the United States. The types of requirements placed on these other sole-source aquifers through state programs are similar to those required for the Edwards Underground Reservoir.

EPA has not adopted regulations for wellhead protection programs, although EPA does require states to submit wellhead protection plans as mandated by §1428 of the SDWA. As is the case for the sole source aquifer program, the SDWA language for the wellhead protection program is rather specific. Most states have developed wellhead protection programs. The actual implementation of most wellhead protection programs is done at the state and local levels. Over 4000 communities nationwide have some form of wellhead protection in place.

This section has presented only the federal regulations associated with the SDWA. Since states, for the most part, take the lead role in groundwater protection activities, readers are encouraged to check with state regulatory agencies to learn specific groundwater protection requirements. Although not required by the SDWA, EPA has encouraged each state to develop a Comprehensive State Ground Water Protection Program (CSGWPP) focusing on the following hierarchy of priorities: (1) prevention of contamination whenever possible; (2) prevention of contamination based on the relative vulnerability of the resource, and where necessary, the groundwater's use, value, and vulnerability; and (3) remediation based on the relative use and value of groundwater. A CSGWPP will incorporate goal setting, priority setting, assignment of responsibilities to different stakeholders, implementation, data collection, and public participa-

_____

* This is a reference to a proposed regulation published in the *Federal Register*, a daily publication of the Federal government's regulatory actions. This proposed regulation was published for comment, along with background and explanatory information, in Volume 60 of the *Federal Register*, beginning at page 44652. The *Federal Register* can be found in most libraries and readers can access it through numerous online services or through the Internet (http://www.access.gpo.gov/su_docs/). The *Federal Register* differs from the Code of Federal Regulations. Each new proposed or final regulation is published in the *Federal Register* whereas the Code of Federal Regulations contains the current final versions of all Federal regulations, including additions, deletions, or changes that have been published in the *Federal Register*.

tion. States with CSGWPPs are eligible for greater flexibility in grant funding and program administration. As of early 1998, 19 states had submitted or were working on CSGWPPs, and 9 of these have been approved.

22.7 SDWA Summary

The SDWA places both indirect (MCLs for delivered drinking water) and direct (UIC, sole source aquifer, wellhead protection, groundwater protection, and source water protection) controls on groundwater quality. The MCLs serve as overriding targets for groundwater quality even though they are intended as drinking water standards to be met at the tap. The UIC program regulates injection of fluids and wastes into underground formations. The sole source aquifer, wellhead protection, groundwater protection, and source water protection programs are designed to prevent drinking water aquifers from becoming contaminated. States have the largest role in managing these programs and receive extensive support from local governments.

22.8 RCRA Legislative Background

22.8.1 Historical Perspective

RCRA, as enacted in 1976, gave EPA broad authority to develop a hazardous and solid waste management regulatory program to promote recycling and ensure protective waste management practices. However, little guidance was included on how to do it, and EPA focused on the regulation of nonhazardous solid wastes. As a result, by 1984 the agency had managed to issue permits (EPA's approach to ensuring protective waste management practices) to only five of the thousands of operating hazardous waste land disposal facilities. Meanwhile, Congress had become even more aware of the effects that land disposal of hazardous wastes was having on the environment (especially groundwater). Congress was intent on preventing contaminated sites that could require future cleanups under CERCLA, but lacked confidence in EPA's ability to develop an effective program. Hence, the Hazardous and Solid Waste Amendments of 1984 (HSWA) were passed, containing a level of detail unprecedented in environmental law and fundamentally altering the RCRA program. Provisions were included which would automatically take effect unless EPA issued regulations to significantly reduce, and comprehensively control, land disposal of hazardous wastes within specified time frames. To date, EPA is still striving to implement many of HSWA's provisions.

22.8.2 Goals and Objectives

§1003(b) of RCRA declares the national policy of the United States to be: "Wherever feasible, the generation of hazardous waste is to be reduced or eliminated as expeditiously as possible. Waste that is nevertheless generated should be treated, stored, or disposed of so as to minimize the present and future threat to human health and the environment."

22.8.3 General Provisions

In general, RCRA regulates solid waste, which includes both nonhazardous solid waste and hazardous waste produced by industrial processes. Subtitle D of the statute deals with nonhazardous solid waste by directing EPA to promulgate guidelines for use by states, which are directed to develop and implement solid waste management plans and to establish permitting programs for certain solid waste management facilities. Subtitle C, on the other hand, authorizes EPA to develop and implement all or portions of an extensive federal program for regulating management of hazardous wastes. Upon request, qualified states can be authorized by EPA to implement the federal hazardous waste program. State regulations must be at least as stringent as EPA's.

When EPA finalizes new federal hazardous waste regulations, EPA implements them in all states until the states become authorized, unless the new federal regulations are based on RCRA authority not created by HSWA, or are less stringent than existing ones. States also sometimes use their own laws as authority to regulate hazardous wastes, independent of RCRA. Readers should check with state waste management agencies to determine which state and federal hazardous waste regulations are applicable for that state.

Also of interest in RCRA is Subtitle I, which regulates underground storage tanks (USTs) containing petroleum and hazardous substances as defined by the CERCLA. Subtitle I does not cover USTs containing hazardous wastes that are regulated under Subtitle C. Like Subtitle C, Subtitle I allows states to be authorized by EPA to implement the federal UST program, provided that their regulations are no less stringent than the corresponding EPA requirements.

22.8.4 Issues and Outlook

RCRA is presently due for reauthorization. Although a comprehensive reauthorization package did not pass during 1996, Congress considered modifications to RCRA addressing several issues, including: (1) over-regulation of low-risk remediation wastes; (2) lack of above-ground storage tank regulations; (3) permitting requirements that slow corrective action; (4) skyrocketing corrective action costs; and (5) lack of "acceptable risk" definition to clarify corrective action cleanup levels.

22.9 RCRA Regulations

Because of the large number of RCRA regulations that deal with groundwater protection, this section of the chapter relies primarily on summary tables.

22.9.1 Hazardous Waste

Regulations implementing RCRA Subtitle C, "Hazardous Waste Management," are located in 40 CFR 260 – 279. Of these, Part 264, "Standards for Owners and Operators of Hazardous Waste Treatment, Storage and Disposal Facilities," Part 265, "Interim Status Standards for Owners and Operators of Hazardous Waste Treatment, Storage and Disposal Facilities," and Part 280, "Underground Storage Tanks," contain specific groundwater monitoring provisions. Part 260 contains definitions and general provisions that apply to all facilities and wastes governed by RCRA Subtitle C. Part 261 defines the terms *solid waste* and *hazardous waste*. For a waste to be regulated under RCRA Subtitle C, it must first fall within the definition of solid waste. Waste which does not meet this definition cannot be RCRA hazardous waste. Unfortunately, the definitions of solid waste and hazardous waste are very complicated, making it sometimes difficult to determine whether a particular waste is regulated.

Facility Standards. Once a waste meets EPA's hazardous waste definition, generators, transporters, and all facilities that treat, store, or dispose of it must comply with applicable RCRA regulatory requirements. Such requirements are intended to encourage minimization of hazardous waste generation, and ensure that hazardous waste management facilities and practices prevent migration of hazardous constituents into the environment, including groundwater. They are also intended to ensure that if hazardous constituents do migrate into the environment, corrective action will occur.

Of particular concern with respect to groundwater protection are hazardous waste facilities and practices involving placement of hazardous waste on or into the land. Therefore, RCRA regulations governing such facilities and practices contain specific groundwater protection provisions, which are summarized briefly in Tables 22.2 and 22.3. Table 22.2 covers the regulations applicable to owners and operators of hazardous waste treatment, storage, and disposal facilities that hold RCRA permits (Part 264). Table 22.3 covers the regulations applicable to interim status facilities (Part 265). To qualify for interim status, the owner or operator must: (1) have been treating, storing, or disposing of the hazardous waste, or commenced facility construction on or before October 31, 1976 (the date RCRA was enacted), or on or before the date on which the waste being managed became subject to RCRA permitting

TABLE 22.2 Groundwater Protection Requirements for Hazardous Waste Treatment, Storage, and Disposal Facilities with RCRA Permits

Provision	Description
40 CFR 264.90	Requires solid waste management units to comply with corrective action requirements in 40 CFR 264.90 - 264.101, but allows surface impoundments, waste piles, land treatment units, or landfills that received hazardous waste after July 26, 1982 (referred to as "regulated units") to comply with 40 CFR 264.91 – 264.100 in lieu of 264.101.
40 CFR 264.91	(1) Requires owner/operator to implement compliance monitoring under 264.99 whenever hazardous constituents under 264.93 from a regulated unit are detected at a compliance point under 264.95; (2) requires owner/operator to institute corrective action under 264.100 whenever the groundwater protection standard under 264.92 is exceeded; (3) requires owner/operator to institute corrective action under 264.100 whenever hazardous constituents under 264.93 from a regulated unit exceed concentration limits under 264.94 in groundwater between the compliance point under 264.95 and the downgradient facility property boundary; and (4) in all other cases, requires owner/operator to institute detection monitoring under 264.98.
40 CFR 264.92	Requires owner/operator to comply with permit conditions designed to ensure that hazardous constituents detected in the groundwater from a regulated unit do not exceed concentration limits in the uppermost aquifer underlying the waste management area beyond the point of compliance during the compliance period.
40 CFR 264.93	Indicates that facility permit will specify the hazardous constituents to which the groundwater protection standard of 264.92 applies.
40 CFR 264.95 and 264.96	Provides that facility permit will specify the point of compliance at which the groundwater protection standard applies and at which monitoring must be conducted, as well as the compliance period.
40 CFR 264.97	Requires groundwater monitoring programs to satisfy Parts 264.98-264.100.
40 CFR 264.98 - 264.100	Establishes owner/operator responsibilities to develop detection monitoring, compliance monitoring, and corrective action programs, respectively, for regulated units.
40 CFR 264.101	Requires applicants for RCRA permits to establish corrective action programs.
40 CFR 264, Subpart G, Closure and Post Closure (264.110 - 264.120)	Requires closure of hazardous waste management units in a manner that, among other things, protects groundwater. Requires closure and post-closure plans that include groundwater monitoring corrective action provisions, if applicable.
40 CFR 264.193, 264.221, 264.251, 264.272, and 264.301	Establishes unit-specific design and operating requirements for tanks, surface impoundments, waste piles, land treatment units, and landfills, respectively, to prevent migration of hazardous wastes or hazardous constituents into groundwater.
40 CFR 264.552	Allows corrective action management units to be designated and establishes requirements for groundwater monitoring at such units.

TABLE 22.3 Groundwater Protection Requirements for Hazardous Waste Treatment, Storage, and Disposal Facilities with Interim Status

Provision	Description
40 CFR 265.90	Requires owners/operators of interim status surface impoundments, landfills, or land treatment facilities which are used to manage hazardous waste to implement, and operate during the active lives of the facilities as well as during post-closure, groundwater monitoring programs which meet the requirements of 265.91, and comply with 265.92 - 265.94.
40 CFR 265.91 and 265.92	Establishes requirements for groundwater monitoring systems and sampling and analysis plans at interim status units, respectively.
40 CFR 265.93	Requires more comprehensive, alternative groundwater monitoring when an owner/operator of an interim status unit assumes (or knows) that monitoring of indicator parameters in accordance with 265.91 and 265.92 would show statistically significant increases (or decreases in the case of pH) when evaluated under 265.93(b).
40 CFR 265.94	Establishes record keeping and reporting requirements for groundwater monitoring programs at interim status units.
40 CFR 265, subpart G, Closure and Post-Closure (265.110 - 265.120)	Establishes requirements for closure and post-closure care of interim status units in a manner protective of human health and the environment, including post-closure groundwater monitoring. Mandates corrective action if migration of hazardous constituents is detected.

requirements; (2) comply with the §3010 notification requirements; and (3) apply for a permit under Part 270.

Generally, the Part 265 regulations specify a minimal groundwater monitoring program to detect hazardous releases to the uppermost aquifer at interim status surface impoundments, landfills, and land treatment facilities. In contrast, Part 264 imposes minimum technology designs on tanks, surface impoundments, waste piles, land treatment facilities, and landfills that are seeking RCRA permits, and requires all solid waste management units (SWMUs) at facilities seeking RCRA permits to comply with corrective action requirements designed to identify and remediate hazardous releases.

Land Disposal Restrictions and Corrective Action Programs. HSWA created two regulatory programs particularly designed to prevent migration of hazardous constituents into groundwater as a result of hazardous waste management activities: the Land Disposal Restrictions (LDR) program and the Corrective Action program. Under the LDR program, all hazardous wastes cannot be land disposed after specified dates unless they have been: (1) treated prior to disposal in accordance with standards established by EPA; (2) treated prior to disposal in accordance with a case-by-case treatability variance granted by EPA or an authorized state ; (3) disposed of in a facility that will prevent migration of hazardous constituents for as long as the waste remains hazardous; (4) approved for a national capacity variance granted by EPA; or (5) approved for a case-by-case capacity variance granted by EPA. While this program does not directly address groundwater monitoring or remediation, it has significantly affected hazardous waste management practices that could result in groundwater contamination.

Prior to HSWA, corrective action requirements under RCRA were applicable only to onsite hazardous waste releases to groundwater from surface impoundments, waste piles, land treatment units, and landfills that received wastes after July 26, 1982 (referred to as regulated units). HSWA expanded corrective action requirements to cover both onsite and offsite releases to all environmental media from solid waste management units at all facilities seeking RCRA permits, including interim status facilities. While EPA incorporated the HSWA language regarding corrective action very quickly into its rules (existing 40 CFR 264.90 -.101), not until 1990 did the agency propose comprehensive implementing regulations (55 FR 30798; July 27, 1990). As proposed, these regulations would require collection of enough information to support determinations of whether corrective action would be needed, and if so, to support selection and implementation of appropriate corrective action. Frequently, groundwater monitoring would be required to satisfy the proposed information collection requirements, and selected corrective actions would probably include groundwater remediation, with accompanying monitoring to evaluate success.

EPA received numerous comments on the proposed regulations and has not yet finalized them, except for a small section dealing with corrective action management units (CAMUs) and temporary units (TUs) (58 FR 8658; February 16, 1993), which are units limited to handling remediation wastes. Until additional corrective action regulations are finalized, the agency is using its 1990 proposal as guidance.

22.9.2 Nonhazardous Waste

Guidelines for state and regional solid waste plans are located in 40 CFR 255–258. Parts 255 and 256 cover the process by which states are expected to develop and implement solid waste management plans. Part 257 sets criteria for determining which solid waste disposal facilities and practices pose a reasonable probability of adverse effects on health or the environment. Such facilities and practices must be appropriately addressed in state waste management plans.

Additionally, Part 257 sets minimum national standards governing nonmunicipal, nonhazardous land disposal units that receive hazardous waste from certain small quantity generators that are conditionally exempt from most RCRA Subtitle C requirements. These minimum national standards impose requirements on such units regarding location, groundwater monitoring, corrective action, and recordkeeping. Table 22.4 summarizes the provisions in Part 257 pertinent to groundwater protection.

Finally, Part 258 establishes minimum national standards for location, operation, design, closure, post-closure care, and financial assurance of MSWLF units. It also delineates minimum national standards for corrective action and groundwater monitoring at MSWLFs, and therefore, is of particular interest

TABLE 22.4 Groundwater Protection Guidelines for Nonmunicipal Nonhazardous Waste Disposal Units that Receive Hazardous Waste from Conditionally Exempt Small Quantity Generators (CESQGs)

Provision	Description
40 CFR 257.8	Establishes demonstration requirements for new units, existing units, and lateral expansions located within a 100-year floodplain.
40 CFR 257.9	Establishes demonstration requirements for new units and lateral expansions located within wetlands.
40 CFR 257.13	Prohibits acceptance after January 1, 1998, of CESQG hazardous waste by existing units located in a 100-year floodplain if the unit has not made the demonstrations required by 40 CFR 257.8.
40 CFR 257.21	Requires units receiving CESQG hazardous waste to install groundwater monitoring systems unless hazardous constituents are shown to have no potential for migration to the uppermost aquifer during the active life of the unit plus 30 years.
40 CFR 257.22–257.25	Establishes requirements for groundwater monitoring systems design and sampling and analysis. Requires detection monitoring, assessment monitoring if a statistically significant increase in hazardous constituents occurs, and public notice if a release of hazardous constituents is confirmed.
40 CFR 257.26–257.27	Establishes requirements for assessment of corrective measures, selection of a remedy, and implementation of corrective action when certain hazardous constituents have been detected above specified levels.
40 CFR 257.30	Establishes recordkeeping requirements for demonstrations, certifications, monitoring records, testing records, analytical data, and findings related to groundwater monitoring.

TABLE 22.5 Groundwater Protection Guidelines for Municipal Solid Waste Landfill Units

Provision	Description
40 CFR 258, subpart B, Location Restrictions (258.10 - 258.19)	Establishes criteria for locating MSWLFs with respect to airports, wetlands, floodplains, seismic zones, fault areas, and unstable areas.
40 CFR 258, subpart C, Operating Criteria (258.20 - 258.29)	Establishes criteria for operating MSWLFs, including among other things, procedures for excluding the receipt of hazardous waste, cover material requirements, requirements for run-on/run-off control systems, limitations on surface water discharges, and restrictions on placing liquids in a landfill.
40 CFR 258, subpart D, Design Criteria (258.40)	Requires design of MSWLFs to assure that concentrations of specified constituents do not exceed certain limits in the uppermost aquifer, and requires composite liners and leachate collection systems.
40 CFR 258, subpart E, Corrective action and Groundwater Monitoring (258.50 - 258.58)	Requires groundwater monitoring systems for contaminant detection at MSWLFs. Establishes requirements for groundwater monitoring system design and sampling and analysis. Requires assessment monitoring program, as well as corrective measures evaluation, selection, and implementation if contaminants are detected.
40 CFR 258, subpart F, Closure and Post-closure Care (258.60 - 258.74)	Requires all MSWLF units to prepare a closure plan, install a final cover system designed to minimize infiltration and erosion, and conduct post-closure care for 30 years, including monitoring the groundwater and maintaining the groundwater monitoring system, if applicable.

with respect to groundwater protection issues. Table 22.5 summarizes provisions of Part 258 pertinent to groundwater protection.

22.9.3 Underground Storage Tanks

In 1984, HSWA added Subtitle I, "Regulation of Underground Storage Tanks," to RCRA as a result of Congress' realization that, of several million USTs containing petroleum or hazardous substances in the United States, tens of thousands were leaking, and more would leak in the future. Since leaking USTs can cause fires or explosions and can contaminate drinking water and groundwater, Subtitle I was designed to address these issues. EPA's regulations implementing Subtitle I are located in Parts 280, 281,

and 282. Parts 281 and 282 deal primarily with the process by which EPA approves state UST programs and documents such approvals. Part 280, however, sets technical standards and corrective action requirements for owners and operators of USTs, and is of most interest with respect to groundwater protection planning.

Groundwater Protection at USTs. USTs are identified and classified by Part 280 according to five criteria: (1) physical characteristics; (2) exclusions; (3) deferrals; (4) contents (i.e., hazardous substance or petroleum); and (5) age (i.e., new or existing). Together, these criteria define how each UST is regulated. Table 22.6 briefly summarizes the UST groundwater protection provisions of Part 280.

TABLE 22.6 Groundwater Protection Standards for Underground Storage Tanks

Provision	Description
40 CFR 280.20	Establishes performance standards for new USTs in order to prevent releases due to structural failure, corrosion, or spills and overfills.
40 CFR 280.21	Requires upgrading of existing USTs.
40 CFR 280.30 - 280.33	Requires operating practices to prevent releases caused by spills, overfills, corrosion, incompatibility of content and tank materials, or repairs.
40 CFR 280.40 - 280.43	Requires new and existing UST systems to be equipped with a method, or combination of methods, of release detection, and establishes standards for such methods. Standards for groundwater monitoring are included.
40 CFR 280.50 - 280.52	Requires reporting and investigation of suspected releases.
40 CFR 280.53	Requires reporting and cleanup of spills and overfills.
40 CFR 280.60 - 280.67	Requires response and corrective action for confirmed releases, including reporting, initial abatement, initial site characterization, free product removal, site investigations for soil and groundwater contamination, preparation of corrective action plan, and public notice.
40 CFR 280.70 - 280.72	Establishes requirements for temporary and permanent closure, including measurements for the detection of releases.

22.10 RCRA Summary

RCRA, as amended by HSWA, attempts to protect groundwater quality through various programs. States are required to regulate land disposal of nonhazardous solid waste (Subtitle D). EPA is required to regulate underground storage tanks containing petroleum and hazardous substances (Subtitle I), as well as the generation, transportation, treatment, storage, disposal, and release remediation of hazardous waste (Subtitle C).

22.11 CERCLA Legislative Background

22.11.1 Historical Perspective

In the late 1970s, many U.S. citizens were becoming distressed about the dangers of sites contaminated with hazardous wastes. Perhaps the most famous case was Love Canal. Residents of this small New York community were evacuated after authorities learned that hazardous waste buried over a 25-year period had contaminated the groundwater and surrounding soils. Responding to concerns over Love Canal and similar sites, Congress passed CERCLA, popularly known as the Superfund law, in 1980. (The term *Superfund* refers to the trust fund that Congress established to pay for CERCLA cleanup and enforcement activities.) Because it covers all environmental media (groundwater, surface water, air, and soil), the scope of the Superfund law is broader than most environmental laws. In 1986, Congress reauthorized CERCLA for five years and strengthened many of its provisions. In 1990, Congress reauthorized CERCLA through 1994 but made no other significant changes. On December 31, 1995, the taxing authority for the trust fund expired. Congress has been trying to reauthorize the controversial law since 1994, but while several bills have been introduced, none has passed.

22.11.2 Goals and Objectives

The primary goal of Superfund is to protect human health and the environment. It provides federal funding and enforcement authority to clean up hazardous waste sites that were created in the past. CERCLA also provides for response to hazardous spills and requires businesses to report releases of hazardous substances.

In 1989, EPA evaluated progress of the Superfund program to that point, and based on its review, developed eight goals for the future. The eight goals were as follows: control acute threats immediately; emphasize enforcement; address worst sites first; monitor and maintain sites over the long term; develop and use new technologies; improve efficiency; encourage public participation; and foster cooperation with other federal and state agencies. Many representatives of industry, environmental organizations, and government would argue that these goals are not being met under the current system, and as a result, major changes to the law are needed.

22.11.3 General Provisions

Because the statute is so broad, few CERCLA provisions pertain exclusively to groundwater. Pertinent portions of those sections of the law that relate to groundwater are summarized below.

§104 — Response Authorities. §104 authorizes EPA to respond when an actual or threatened release of a hazardous substance may endanger public health or welfare. CERCLA provides for two types of response actions. Removal actions are short-term responses that stabilize or clean up a hazardous site that poses a threat to public health or the environment. Removal actions include capping of contaminated soils to reduce migration of hazardous substances into groundwater and providing temporary alternate sources of drinking water to local residents. The second type of response action is a remedial action. Remedial actions are taken instead of or in addition to removal actions to minimize the migration of hazardous substances so that they do not cause substantial danger to present or future public health or the environment. Remedial actions include the study, design, and construction of longer-term actions aimed at permanent remedy. Typical actions include constructing underground walls to control the movement of groundwater and pumping and treating of contaminated groundwater. Both removal and remedial actions must be conducted in accordance with the National Contingency Plan (NCP).

§105 — National Contingency Plan. CERCLA requires EPA to revise and republish the NCP for the removal of oil and hazardous substances, originally prepared and published pursuant to §311 of the CWA. CERCLA requires that the plan reflect the responsibilities created by CERCLA, including the establishment of procedures for responding to releases of hazardous substances. The NCP is published as an EPA regulation in 40 CFR 300. It specifies procedures for identifying, removing, and remedying releases of hazardous substances. §105 also provides for the Hazard Ranking System (HRS). The HRS is a scoring system used to evaluate potential relative risks to human health and the environment from releases or threatened releases of hazardous substances. EPA uses the HRS to determine if a site should be on the National Priorities List (NPL). The NPL is EPA's list of the most serious hazardous waste sites identified for possible long-term remedial response using money from the Trust Fund. The NCP and the HRS are described further under CERCLA Regulations.

§121 — Cleanup Standards. §121 provides principles governing the cleanup of hazardous substances. §121(a) states that remedial actions are to be selected in accordance with the NCP. §121(b) states that remedies in which treatment permanently and significantly reduces the volume, toxicity, or mobility of the hazardous substances are preferred over remedies that do not provide such treatment. CERCLA requires that remedial actions are to be protective of human health and the environment, cost-effective, and utilize permanent solutions and alternative treatment technologies to the maximum extent practicable. In determining the degree of cleanup, CERCLA provides no standards of its own. However, §121(d)(2) requires site cleanups to attain standards from other federal and state environmental programs that are applicable or relevant and appropriate requirements (ARARs). §121(d)(2)(A)(ii) states that such

remedial actions shall require a level or standard of control which at least attains MCLGs established under the SDWA and water quality criteria established under the CWA, where such goals or criteria are relevant and appropriate under the circumstances of the release or threatened release. §121(d)(2)(B) states that in determining whether or not any water quality criteria under the CWA is relevant and appropriate under the circumstances for the release or threatened release, the designated or potential use of the groundwater must be considered. §121(d)(4) provides limited waivers for achieving ARARs.

§121(e) provides that CERCLA cleanups require no federal, state, or local permits for the portion of any removal or remedial action conducted entirely onsite, where such remedial action is selected and executed in compliance with the rest of §121. §121(f) provides for "substantial and meaningful" state involvement in initiation, development, and selection of remedial actions to be undertaken in that state.

22.11.4 Natural Resource Damages

CERCLA §107(a) provides that owners, operators, transporters, or generators potentially responsible for contamination problems at a Superfund site are liable for natural resource damages (NRD) resulting from a release of a hazardous substance. NRDs are monetary compensations for injury to, destruction of, or loss of natural resources. They are distinct from response costs, which are the costs of actions taken under the NCP to remove threats to human health and the environment caused by hazardous substance releases. §301(c) of CERCLA requires the Department of the Interior to promulgate regulations for assessing NRDs resulting from hazardous substance releases. The regulations must include two types of assessment procedures. "Type A" procedures are standard procedures for simplified assessments. "Type B" procedures are "alternative protocols for conducting assessments in individual cases."

Unlike many environmental regulatory programs, CERCLA does not provide for state delegation. Cleanup of sites financed by the Superfund trust fund is administered by EPA and its ten federal regions. Before the federal government initiates a fund-financed remedial action, the state must assure that it will provide for future maintenance of response actions and pay for 10% of all costs of remedial action, or 50% if the state had operated the facility undergoing cleanup. Many states have comparable programs that provide for the cleanup of contaminated sites that are not on the federal NPL. In all cases, concerns of the state and local community are considered in cleanup decisions.

22.11.5 Issues and Outlook

When enacted in 1980, Congress expected that only a few hundred sites would require cleanup and that the program would require relatively modest funding. However, today there are more than 1300 sites on the NPL, and after 15 years and $25 billion spent, fewer than 300 sites have been cleaned up. Superfund has many critics. A key criticism pertains to the liability system, which requires parties to pay for cleanups for which they may not be responsible; this can result in excessive litigation and other transaction costs that divert money from actual cleanup. Another relates to the remedy selection process. Many argue that the prescriptive, cumbersome, and confusing nature of the process has delayed cleanups. Significant controversy surrounds the requirements for ARARs, permanent treatment versus containment, and the use of risk assessment. Congress has been struggling to modify and reauthorize CERCLA since 1991. Legislation introduced in the 104th Congress would have eliminated ARARs and the preferences for permanence and treatment in remedy selection. It would have required increased use of risk and consideration of future water and land uses in decision making. Remedy selection decisions for groundwater would have to consider type and timing of use; technical feasibility and reasonableness of cost where contamination threatens uncontaminated usable groundwater; and a range of possible remedies, including pump and treat, point-of-use treatment, containment, and natural attenuation. However, controversy over these provisions and others, including proposed changes to the liability scheme and to the natural resource damage provisions, has prevented these bills from passing in either the House or the Senate.

22.12 CERCLA Regulations

Superfund regulations are contained in the NCP (40 CFR 300) and in the Natural Resource Damage Assessment (NRDA) regulations (43 CFR 11).

The NCP contains procedures for responding to discharges of hazardous substances. Subpart E describes methods and criteria for assessing sites and determining the appropriate response to hazardous substance releases. Site assessment determines whether a site is eligible for cleanup funding and includes the following activities:

- Site Discovery — identifies hazardous substance releases.
- Preliminary Assessment — evaluates existing data to determine need for further action.
- Site Investigation — assesses onsite conditions to determine need for HRS scoring.
- HRS scoring — Applies a mathematical approach to assess the relative risk of a site and determine its eligibility for the NPL. The HRS considers four exposure pathways: groundwater migration, surface water migration, soil exposure, and air migration. The groundwater migration pathway is based on three factor categories: likelihood of release, waste characteristics, and targets. The HRS evaluates these pathways and then combines them to produce an overall score for the site. Appendix A to Part 300 provides details on HRS scoring calculations; §3.0 of that appendix pertains to groundwater pathway calculations.
- NPL listing — determines sites eligible for Superfund-financed remedial action. If the score resulting from the HRS process is at least 28.5, the site is placed on the NPL.

Once a site is placed on the NPL, the remedial process begins. The NCP provides general expectations to help identify and implement appropriate remedial actions. Some of these pertain to groundwater. For example, 40 CFR 300-430(a)(iii)(D) expects EPA to use institutional controls such as water use restrictions to supplement engineering controls in limiting exposure to hazardous substances. However, institutional controls cannot substitute for active response measures (e.g., restoration of groundwater to its beneficial uses) as the sole remedy, unless such active measures are determined not to be practicable. EPA expects to return usable groundwater to its beneficial uses whenever practicable within a time frame that is reasonable; when restoration of groundwater to beneficial uses is not practicable, EPA expects to prevent further migration of the plume, prevent exposure to the contaminated groundwater, and evaluate further risk reduction.

The NCP prescribes the following specific activities for the remedial action process:

- Remedial Investigation (RI) — Determines the nature and extent of the problem presented by a release. The RI includes field investigations to assess the characteristics of groundwater and other environmental media. A site-specific risk assessment is conducted to characterize the current and potential threats to human health and environment that may be posed by contaminants migrating to groundwater, surface water, and other pathways. The results of the baseline risk assessment help establish acceptable exposure levels for developing remedial alternatives in the feasibility study.
- Feasibility Study — Develops and evaluates options for remedial action. Remediation goals are set to establish acceptable exposure levels that are protective of human health and the environment. For groundwater, MCLGs established under the SDWA that are set at levels above zero are to be attained by remedial actions for ground or surface waters that are current or potential sources of drinking water, where the MCLGs are relevant and appropriate. If an MCLG is determined not to be relevant and appropriate, the corresponding MCL is to be attained where relevant and appropriate. Where the MCLG for a contaminant has been set at zero, the MCL is to be attained by remedial actions for ground or surface waters that are current or potential sources of drinking water, where the MCL is relevant and appropriate.

- Remedy Selection — Identifies the likely remedial alternative and provides for public comment. The Record of Decision (ROD) is the official report that documents the background information on the site and describes the chosen remedy and the justification for that remedy.
- Remedial Design — Prepares technical plans for implementing the chosen remedy.
- Remedial Action — Implements the remedial design.
- Operation and Maintenance — Ensures that the cleanup methods are working properly after a response action occurs. A groundwater restoration activity will be considered administratively complete when: (1) actions restore ground or surface water quality to a level that assures protection of human health and the environment; (2) actions restore groundwater or surface water to such a point that reductions in contaminant concentrations are no longer significant; or (3) ten years have elapsed, whichever is earliest. Ground or surface water measures initiated for the primary purpose of providing a drinking water supply and not for the purpose of restoring groundwater do not constitute treatment.

22.12.1 Natural Resource Damage Assessments (NRDA)

NRDA regulations (43 CFR 11) define the process whereby a natural resource trustee may pursue compensation on behalf of the public for injury to natural resources resulting from releases of hazardous substances. As noted above, Type B procedures are used for conducting assessments in individual cases. Type B procedures prescribe specific procedures for determining whether an injury has occurred to groundwater (and other) resources and for determining if the exposure pathway of the hazardous substance was a groundwater (or other) pathway. It also prescribes testing and sampling methods for use in determining injury and procedures for use in quantifying baseline conditions and effects of the release on groundwater (and other) resources for determining the appropriate amount of compensation.

22.13 CERCLA Summary

In summary, CERCLA provides for the cleanup of groundwater and other media at old or abandoned sites contaminated by releases of hazardous substances. The law contains no cleanup standards of its own, relying instead on requirements from other federal and state environmental programs such as the CWA and the SDWA. Many Superfund critics blame the law and its regulations for high cleanup costs and slow cleanup progress. Congress has been trying to reauthorize and improve the law since 1991, but controversy over key issues has frustrated reform efforts.

22.14 Overlap of RCRA and CERCLA Groundwater Remediation Programs

The two principal programs governing groundwater remediation are the RCRA corrective action program and the removal and remedial response programs of CERCLA. These two programs, while attempting to reach the same goal, are not mutually exclusive and do not necessarily require the same actions to achieve compliance. A key difference between CERCLA and RCRA is that, in general, CERCLA addresses sites that have been abandoned, or for which no financially viable owners can be identified; RCRA addresses operating sites and facilities. Also, CERCLA covers hazardous substances, while RCRA addresses a subset of hazardous substances — hazardous wastes. Many facilities, but in particular those owned or operated by agencies of the federal government, find that they are required to comply with the requirements of both programs. Such a situation is further complicated by the way the two programs are administered, with RCRA corrective action being a state-led, EPA-led, or shared state-EPA responsibility, depending on the specifics of a given state's RCRA authorization, and CERCLA being primarily an EPA-lead program.

The potential for overlapping application of these authorities has been cause for confusion and concern not only in the regulated community, but among state and federal regulators as well. In an attempt to resolve these concerns, EPA has repeatedly referred to the concept of parity between the two programs, suggesting that since both programs address cleanup of potential and actual releases, both programs should arrive at similar remedial solutions and that any procedural differences between the two programs should not substantively affect the outcome of remedial activities.

With respect to groundwater restoration, there are two excellent examples of the EPA's efforts to achieve true parity between CERCLA and RCRA corrective action. First, both programs seek to restore groundwater to beneficial use. The goal of both programs with respect to the restoration of groundwater is, for all intents and purposes, identical. Therefore, the only significant differences in conducting groundwater restoration activities are procedural.

The other example of how EPA has attempted to achieve parity between the two programs is in the issuance of a determination of technical impracticability for groundwater restoration. In EPA's 1993 *Guidance on Evaluating the Technical Impracticability of Ground-Water Restoration*, the agency stated that "experience over the past decade has shown that achieving the required final cleanup standards may not be practicable at some sites due to the limitations of remediation technology." The guidance embodied in this EPA document outlines the agency's approach to evaluating the practicability of achieving the required degree of groundwater restoration and for establishing alternative, protective remedial strategies where restoration is determined to be technically impracticable. The parity of this concept is best expressed by the statement that the guidance in the document applies to both the RCRA corrective action and Superfund programs.

Notwithstanding the intent of EPA to achieve parity and to fully integrate environmental restoration activities under these two programs, there is little definitive guidance on other mechanisms to integrate remedial activities under both programs. Facilities that have both RCRA corrective action and CERCLA remedial activities have several opportunities for integration of the two programs, as described in the following paragraphs.

22.14.1 Follow One Regulatory Program Whenever Possible

There often is a large amount of duplication of effort in meeting the administrative requirements of both authorities, with a corresponding waste of resources. This problem can, however, be solved by convincing regulators to use a single authority to conduct the response, or alternatively, to allow the facility to conduct an integrated response which is designed to satisfy both sets of requirements.

Incorporating RCRA requirements into CERCLA actions (and vice versa) requires linking logically connected actions performed sequentially over time or concurrently at different parts of the facility into a single integrated action, and requires consideration of the remedial alternative for the individual phase or unit *and* the ultimate remedial goals for the entire facility. Such an approach has an additional benefit in that it works to foster a view of the facility as a whole, rather than as a series of separate problems to be addressed.

22.14.2 Undertake Voluntary Response Actions

EPA encourages voluntary response actions. Voluntary cleanups have a number of advantages, including timeliness, flexibility, and efficient use of facility and regulatory agency resources. Unfortunately, in some cases procedural barriers have delayed voluntary cleanups, but EPA and the state regulatory agencies are looking for means to address these problems and to further encourage and facilitate voluntary actions.

22.14.3 Use the Data Quality Objectives Process to Focus and Integrate Investigative Efforts

Data quality objectives (DQOs) are used to develop an effective sampling plan which avoids the collection of data that are inconsequential to decision making. DQOs are developed as part of designing the

investigative and sampling protocol to be followed during the investigation. In summary, the DQO process will help: establish the objectives of the data collection effort; specify how the data will be used to support risk management decisions; define the scope of the data collection effort; specify acceptable levels of decision errors that will be used as the basis for establishing the quantity and quality of data needed; and specify the quantity and quality of data to be collected.

With respect to integrating groundwater investigations and remedial activities under RCRA and CER-CLA, the DQO process is an invaluable tool for determining how to meet the specific sampling and analysis requirements imposed by the regulatory agencies. The process is useful in the analysis of requirements for investigative activities such as sampling and analysis to minimize duplication of effort (e.g., requirements to use two different analytical procedures for the same substance), avoiding repeated analysis for contaminants that are not present, and for establishing the frequency of sampling.

Another aspect of using the DQO process is working with the regulatory agencies to establish guidelines for determining that a release poses no significant threat to human health or the environment and that additional action is unnecessary, as early in the process as possible. With respect to groundwater restoration, this would involve: (1) developing and continually narrowing (as appropriate) a target list of constituents for sampling (as opposed to always sampling for a very large suite of contaminants); (2) early characterization of the overall quality of the groundwater (i.e., analyzing for factors such as hardness, total dissolved solids, and other parameters generally related to water quality); (3) making an assessment of known and potential beneficial uses of the groundwater; (4) assessment of whether the release might be a candidate for an early determination of technical impracticability; (5) early assessment of potential exposure of humans or environmental receptors to contaminated groundwater; and (6) early assessment of the potential health or environmental affects of exposure to the contaminated groundwater.

22.14.4 Develop Integrated Planning Documents and Standardized Report Formats

Another suggestion for a facility conducting both RCRA corrective action and CERCLA responses is to work with its regulatory agency to develop integrated planning documents for conducting investigations and remedial activities. There is nothing to prevent a facility subject to both RCRA and CERCLA from developing planning documents that fulfill both sets of requirements; the documents required to conduct corrective action differ only slightly from those required for CERCLA actions, but are referred to by different names.

22.14.5 Integrate the Evaluation of Alternatives to Address Contamination

The specific criteria for selecting a corrective measure under RCRA and the criteria for CERCLA remedial actions are not identical, but do share common elements. Facilities required to comply with both programs should consider combining the evaluation process and criteria into a single evaluation.

For Further Information

SDWA

The following documents provide useful information on establishing and implementing wellhead protection programs:

Witten, J., S. Horsley, S. Jeer, and E. Flanagan, 1995, *A Guide to Wellhead Protection*, prepared for the American Planning Association, Chicago, IL and EPA, Washington, DC, report number 457/458, August.

EPA, *Wellhead Protection: A Guide for Small Communities*, EPA/625/R-93/002, Feb. 1993.

EPA, *Delineation of Wellhead Protection Areas in Fractured Rocks*, EPA 570/9-91-009, June 1991.

RCRA and CERCLA

The following documents have been issued by EPA or the U.S. Department of Energy (DOE) to provide guidance on a variety of RCRA and CERCLA groundwater topics. Where available, the NTIS* order number is provided in brackets at the end of the citation. While this list is not intended to be comprehensive, it may assist readers in locating pertinent federal guidance.

Groundwater Monitoring at Hazardous Waste Treatment, Storage, and Disposal Facilities (TSDFs)

EPA, *RCRA Groundwater Monitoring: Draft Technical Guidance*, EPA/530-R-93-001, Nov. 15, 1992. [PB93-139 350] Provides technical guidance for implementing the groundwater monitoring regulations for regulated units contained in 40 CFR 264 Subpart F and the permitting standards of 40 CFR 270. Also provides guidance to owner/operators of TSDFs that are required to comply with other subparts of 40 CFR 264.

EPA, *Handbook of RCRA Groundwater Monitoring Constituents: Chemical and Physical Properties (Appendix IX to 40 CFR Part 264*, EPA/530-R-92-022, Sept. 15, 1992. [PB92-233 287] Contains the physical and chemical properties of constituents listed in 40 CFR 264 Appendix IX. The handbook organizes groundwater monitoring constituents by Appendix IX name and includes constituents that are being considered for addition to or deletion from Appendix IX.

EPA, *RCRA Permit Policy Compendium, Volume 7 (9460.1980 — 9482.1990)*, EPA/530-SW-91-062G, Aug. 15, 1991 (updated version complete through 1996 available on Internet at http://www.epa.gov/epaoswer/hazwaste/permit/compend.htm). [PB92-111 772] Compiles related EPA Office of Solid Waste and Emergency Response (OSWER) directives dealing with RCRA permit policy. Includes, RCRA/Superfund Hotline Monthly summaries, letters, memos, and responses to transporter standards, contingency plans, TSDF technical requirements, groundwater standards, management of containers, and more.

EPA, *Status of Contaminated Groundwater and Limitations on Disposal and Reuse*, OSWER Directive #9481.00-11, July 13, 1990.

EPA, *Statistical Analysis of Groundwater Monitoring Data at RCRA Facilities; Interim Final Guidance*, EPA/530-SW-89-026, April 15, 1989. [PB89-151 047] Assists regional and state personnel in evaluating groundwater monitoring data from RCRA facilities. Guides statistical analysis of groundwater monitoring pertaining to spatial relationships between monitoring wells and potential contaminant sources.

EPA, *Ground-Water Monitoring at Clean Closing Surface Impoundment and Waste Pile Units*, OSWER Directive #9476.00-14, Mar. 31, 1988.

EPA, *RCRA Regulatory Status of Contaminated Ground Water*, OSWER Directive #9481.00-06, Nov. 13, 1986.

EPA, *RCRA Ground Water Monitoring Technical Enforcement Guidance Document*, EPA/530-SW-86-055, Sept. 9, 1986. [PB87-107 751/AS] Describes the essential components of a groundwater monitoring system to meet goals of RCRA.

EPA, *Criteria for Identifying Areas of Vulnerable Hydrogeology under RCRA (Complete Set of 5 Volumes)*, EPA/530-SW-86-022, -022A, -022B, -022C, and -022D, July 15, 1986. [PB86-224 946] Provides RCRA permit writers with a standardized technical method for evaluating hydrogeologic data submitted in permit applications for hazardous waste land treatment, storage, and disposal facilities. The methodology determines if facilities are located in areas of vulnerable hydrogeology.

EPA, *Groundwater Monitoring Above the Uppermost Aquifer*, OSWER Directive #9481.02-85, April 8, 1985.

_____

* The U.S. Department of Commerce operates the National Technical Information Service, from which copies of many U.S. government documents can be purchased by calling (703) 487-4650 between the hours of 8:30-5:00 eastern time, Monday through Friday.

EPA, *Responses and Mechanisms to Prevent Groundwater Monitoring Deficiencies in Part B Applications*, OSWER Directive #9504.02-84, Nov. 29, 1984.

EPA, *Enforcing Ground Water Monitoring Requirements in RCRA Part B Permit Applications*, OSWER Directive #9504.01-84, Aug. 16, 1984.

EPA, *Clarification of the Definition of Aquifer in 40 CFR 260.10*, OSWER Directive #9481.06-84, June 4, 1984.

EPA, *Guidance on Implementation of Subpart F Requirements for Stating Significant Increase in Indicating Parameter Values*, OSWER Directive #9481.06-83, Nov. 30, 1983.

EPA, *Groundwater Monitoring Guidance for Owners and Operators of Interim Status Facilities*, EPA/SW-963, Mar. 15, 1983. [PB83-209 445] Assists owners and operators of hazardous waste management facilities required to meet groundwater monitoring standards. Discusses the purpose of groundwater monitoring requirements and presents the rationale behind the policy.

EPA, *Waiving Groundwater Monitoring Requirements in an Authorized State*, OSWER Directive #9481.03-82, Oct. 14, 1982.

EPA, *Groundwater Monitoring (40 CFR 265, Subpart F); Standards Applicable to Owners and Operators of Hazardous Waste Treatment, Storage, and Disposal Facilities under RCRA, Subtitle C, Section 3004*, May 20, 1980. [PB81-189 797] Provides EPA support for setting groundwater monitoring standards at hazardous waste management facilities. Specifies monitoring requirements.

Groundwater Monitoring at Municipal Solid Waste Landfills

EPA, *Solid Waste Disposal Facility Criteria: Technical Manual*, EPA/530-R-93-017, Nov. 15, 1993. [PB94-100 450] Addresses general applicability of the Part 258 criteria, location restrictions, operating requirements, design standards, groundwater monitoring and corrective action, and closure and post-closure care for landfills. Written for municipal solid waste landfill owners and operators.

EPA, *Safer Disposal for Solid Waste: The federal Regulations for Landfills*, EPA/530-SW-91-092, Mar. 15, 1993. Summarizes the federal regulations covering landfill location, operation, design, groundwater monitoring and corrective action, closure and post-closure care, and financial assurance. Gives owners/operators and local officials dates for compliance and additional sources of information.

EPA, *Criteria for Municipal Solid Waste Landfills: Groundwater Monitoring and Corrective Action (Subpart E)*, EPA/530-SW-88-043, July 15, 1988. [PB88-242 490] Provides detailed information on development and implementation of groundwater monitoring and corrective action requirements. Options for groundwater monitoring and corrective action programs are discussed, and EPA rationale is explained.

Groundwater Remediation

EPA, Guidance for Evaluating the Technical Impracticability of Ground-Water Restoration, OSWER Directive No. 9234.2-25, October, 1993.

EPA, *Considerations in Groundwater Remediation at Superfund Sites and RCRA Facilities: Update,* OERR, May 27, 1992. Clarifies and expands OSWER's general policy concerning remediation of contaminated groundwater, especially with regard to nonaqueous phase liquid (NAPL) contaminants.

EPA, *Guide to Pump and Treat Groundwater Remediation Technology*, OERR, Nov. 1990. Summarizes how to use available hydrogeological and chemical data to determine when, where, and how pump and treat technology can be used successfully to contain and remediate contaminant plumes.

EPA, *Suggested ROD Language for Various Groundwater Remediation Options*, OERR, Oct. 10, 1990.

EPA, *Considerations in Groundwater Remediation at Superfund Sites,* OERR, Oct. 18, 1989. Transmits findings from a study of several sites where groundwater extraction was conducted to contain or reduce levels of contaminants in groundwater.

EPA, *Guide on Remedial Actions for Contaminated Groundwater*, OERR, April 1989. Summarizes key issues in development, evaluation, and selection of groundwater remedial actions at Superfund sites.

Groundwater Investigations at Remediation Sites

DOE, A Comparison of the RCRA Corrective Action and CERCLA Remedial Action Processes, DOE/EH-0365, February 1994.

DOE, *Remedial Investigation/Feasibility Study (RI/FS) Process, Elements, and Techniques Guidance*, December 1993.

EPA, *Guidance for Planning for Data Collection in Support of Environmental Decision-Making Using the Data Quality Objective Process*, QA/G4, Interim Final, October 1993.

DOE, *RCRA Corrective Action Program Guide*, Interim Guidance, May 1993.

EPA, *Quality Assurance/Quality Control Guidance for Removal Activities, Sampling QA/QC Plan and Data Validation Procedures*, Interim Final, EPA/540/G-90/004, April 1990. [PB90-274481].

EPA, *RCRA Facility Investigation (RFI) Guidance; Interim Final; Volume II: Soil, Groundwater and Subsurface Gas Releases*, EPA/530-SW-89-031, May 15, 1989. [PB89-200 299] Provides guidance to regulatory agency personnel on overseeing owners or operators of hazardous waste management facilities conducting the second phase of the RCRA Corrective Action Program, a RCRA Facility Investigation (RFI).

Glossary

The following definitions are provided for the readers' better understanding and are not necessarily the same as the applicable legal definitions.

ARARs　Applicable or relevant and appropriate requirements include any standard, requirement, criterion, or limitation under any federal environmental law, (e.g., CWA, SDWA, TSCA [Toxic Substances and Control Act], RCRA) and any promulgated standard, requirement, criterion, or limitation under a state environmental or facility siting law that is more stringent than any federal standard requirement, criterion, or limitation.

CAMUs　Corrective action management units are land areas located within an RCRA-regulated facility, that have been designated by EPA or the authorized state for the purpose of managing remediation wastes generated from corrective action activities at that facility.

CESQG　A conditionally exempt small quantity generator is a hazardous waste generator that produces less than 100 kg of hazardous waste in a calendar month.

Corrective Action　A corrective action is any action taken to remedy a hazardous release.

CSGWPP　The Comprehensive State Ground Water Protection Program is an EPA program that attempts to focus state resources on prevention of contamination and remediation based on the relative use and value of groundwater.

DQOs　Data quality objectives are qualitative or quantitative statements which specify the quality of data required to support remedial response or other environmental management decisions.

Hazardous Substance　A hazardous substance includes, but is not limited to, any substance designated under the CWA, the Clean Air Act (CAA), TSCA, or any "hazardous waste" under RCRA.

Hazardous Waste　A hazardous waste is defined as any solid waste that is not excluded from RCRA Subtitle C regulation, and that exhibits a hazardous characteristic, is listed as hazardous waste under RCRA regulations, or is mixed with or derived from hazardous waste.

HRS　The Hazard Ranking System is a scoring system used to evaluate potential relative risks to human health and the environment from releases or threatened releases of hazardous substances.

Injection Well　Practically speaking, an injection well is a hole in the ground that is deeper than it is wide that receives wastes or other fluids. Under the UIC program, drain fields, leach fields, drywells, and sumps are considered to be injection wells even though traditionally they have not been thought of as wells.

LDRs Land disposal restrictions treatment standards are EPA regulations adopted under RCRA specifying concentration levels or methods of treatment that substantially diminish the toxicity or likelihood of migration of hazardous constituents from hazardous wastes placed into or onto the land.

MCLs Maximum contaminant levels are enforceable standards for drinking water quality. MCLs may be less stringent than MCLGs since they consider not only human health protection but also the availability of economically achievable treatment technology.

MCLGs Maximum contaminant level goals are nonenforceable targets set to protect human health.

MSWLF A municipal solid waste landfill is a discrete, publicly or privately owned area of land that receives household waste (e.g., garbage, trash and sanitary waste from homes, hotels, campgrounds, etc.), and may receive commercial and industrial solid waste, nonhazardous sludge, or hazardous waste from generators exempt from RCRA Subtitle C because they produce such small quantities.

NCP Officially known as the National Oil and Hazardous Substances Pollution Contingency Plan (40 CFR 300), the NCP outlines the responsibilities and authorities for responding to releases into the environment of hazardous substances and other pollutants and contaminants under the statutory authority of CERCLA and §311 of the CWA.

NPDES The National Pollutant Discharge Elimination System is a federal program that issues waste-water discharge permits to all point source discharges to surface water.

NPL The National Priorities List is EPA's list of the most serious hazardous waste sites identified for possible long-term remedial response using money from the Superfund Trust Fund.

NRD Natural Resource Damages are monetary compensations for injury to, destruction of, or loss of natural resources.

NRDA Natural Resource Damage Assessments are processes whereby a trustee assesses the damages to natural resources resulting from a release of a hazardous substance under the Superfund law.

Natural Resources Natural resources are the land, fish, wildlife, biota, air, water, groundwater, drinking water supplies, and other such resources controlled by federal, state, local, or foreign governments, or Indian tribes.

Nonpoint sources Nonpoint sources of pollution are those that do not enter surface or groundwater through discrete pipes or conveyances. They include agricultural runoff, atmospheric deposition, groundwater inflow to stream beds, acid mine drainage, and some types of storm water runoff.

POTW A publicly owned treatment works is a wastewater treatment plant owned by a state or local government.

RCRA Base Program The RCRA base program is defined as all RCRA Subtitle C requirements imposed prior to enactment of the Hazardous and Solid Waste Amendments of 1984.

RI A Remedial Investigation is a process used under the National Contingency Plan to determine the nature and extent of the problem(s) presented by a release of a hazardous substance.

ROD A record of decision is a public document that explains which cleanup alternative(s) will be used at a National Priorities List site.

Release A release is any spilling, leaking, pumping, pouring, emitting, emptying, discharging, injecting, escaping, leaching, dumping, or disposing into the environment. Excluded from the definition of release are workplace exposures covered by OSHA, vehicular engine exhausts, radioactive contamination covered by other statutes, and the normal application of fertilizer.

Remediation Wastes Remediation wastes are wastes managed for the purpose of implementing corrective action.

Sole Source Aquifer A sole source aquifer is a sole or principal drinking water source for an area.

Solid Waste A solid waste is any material that has been discarded by being abandoned, recycled, or considered inherently waste-like, and that is not excluded from RCRA Subtitle C requirements.

SWMU A solid waste management unit is any discernable unit at which solid wastes have been placed at any time, irrespective of whether the unit was intended for the management of solid or hazardous wastes.

TUs Temporary units are tanks or container storage units, located within the boundaries of an RCRA-regulated facility, that have been designated by EPA or the authorized state for use in treating and storing remediation wastes generated from corrective action activities at that facility, and that will operate for no longer than one year.

UIC The Underground Injection Control program is a federal program that regulates injection of fluids underground through injection wells.

USDW An underground source of drinking water is an aquifer that presently serves or could reasonably be expected to serve as a drinking water supply.

USTs Under RCRA, USTs are underground storage tanks containing petroleum and hazardous substances.

Wellhead Protection Programs Wellhead protection programs are designed to impose controls on activities located in surface and subsurface areas surrounding a well that supplies a public water system, through which contaminants are reasonably likely to move toward and reach the well.

23
Landfills

Pedro C. Repetto
Woodward-Clyde International

23.1 Introduction

This chapter presents a general overview of the primary aspects associated with landfill site selection and design, with special emphasis on hydrogeologic aspects and goundwrater monitoring. The main landfill elements and the sequence in which they must be considered in design are discussed in detail. Regulatory aspects related to siting, design, and groundwater monitoring are also reviewed throughout the paper.

Development of a landfill project requires proper attention to the diverse and, sometimes, conflicting array of regulatory, physical, and technical considerations relevant to siting and design. An essential step in the development of a landfill project is site selection. A thorough understanding of the physical setting of the site and the regulatory requirements for the design and permitting are of paramount importance for successful development of a project.

Landfill designers must keep in mind that environmental regulations are intended to protect human health and the environment, and that they do not necessarily provide all the requirements of a sound engineering design. Therefore, as part of the design process, appropriate hydrogeologic engineering analyses must be performed to ensure adequate performance of all landfill elements.

23.2 Regulatory Background

The basic federal regulation for sanitary (municipal) landfills in the United States is the Environmental Protection Agency (EPA) Criteria for Municipal Solid Waste Landfills (MSWLF) (40 CFR 258), generally referred to as Subtitle D (of the Resource Conservation and Recovery Act). Other federal regulations that establish additional criteria for MSWLFs have been promulgated under the Clean Water Act and the Clean Air Act.

The regulations for the design of hazardous waste landfills have been promulgated under Subtitle C of the Resource Conservation and Recovery Act. The regulations for the design of municipal and hazardous waste landfills are similar, except for the requirements for liner and final cover systems. Therefore, only the regulations for MSWLFs are discussed for brevity.

23.2.1 Resource Conservation and Recovery Act

The purpose of the Subtitle D regulations is to establish minimum national criteria for MSWLFs. Subtitle D regulations are organized into the following subparts: Location Restrictions, Operating Criteria, Design Criteria, Groundwater Monitoring and Corrective Action, Closure and Post-Closure Care, and Financial Assurance Criteria. The Location Restrictions specified under Subtitle D are discussed below. Other Subtitle D requirements are discussed elsewhere in this chapter.

Subtitle D specifies the following location restrictions: Airport Safety, Floodplains, Wetlands, Fault Areas, Seismic Impact Zones, and Unstable Areas. A brief description of each of these restrictions is presented below.

Airport Safety. Existing units, new units, and lateral expansions located within 10,000 feet of any airport runway end used by turbojet aircraft, or within 5,000 feet of any airport runway end used only by piston-type aircraft must demonstrate that the facility does not pose a bird hazard to aircraft.

Floodplains. Existing units, new units, and lateral expansions located in 100-year floodplains must demonstrate that the facility does not restrict the 100-year flood flow, reduce the temporary water storage capacity of the floodplain, or result in the washout of solid waste so as to pose a hazard to human health or the environment.

Wetlands. This restriction requires that new units and lateral expansions shall not be located in wetlands, unless the following demonstrations are made:

- Where applicable under Section 404 of the Clean Water Act, a demonstration must be made that no practicable alternative to the proposed landfill is available.
- Construction and operation of the MSWLF unit will not cause or contribute to:
  - Violations of any applicable state water quality standard
  - Violate any applicable toxic effluent standard or prohibition under Section 307 of the Clean Water Act
  - Significant degradation of wetlands
  - Jeopardize the continued existence of threatened or endangered species or result in the destruction or adverse modification of a critical habitat protected under the Endangered Species Act of 1973
  - Violate any requirement under the Marine Protection, Research, and Sanctuaries Act of 1972 for protection of marine sanctuaries.
- To the extent required under Section 404 of the Clean Water Act or applicable state wetlands regulations, a demonstration that steps have been taken to attempt to achieve no net loss of wetlands (as defined by area and function). This is to be done by first avoiding impacts to wetlands to the maximum extent practicable, minimizing unavoidable impacts to the maximum extent practicable, and finally offsetting remaining unavoidable wetland impacts through all appropriate and practicable compensatory mitigation actions.

Fault Areas. New units and lateral expansions may not be located within 200 feet of faults that have experienced displacement in Holocene time, unless it is demonstrated that a smaller setback distance will be protective of human health and the environment.

Seismic Impact Zones. New units and lateral expansions may not be located in seismic impact zones, unless it is demonstrated that the facility is designed to resist the maximum horizontal acceleration in lithified material for the site.

Unstable Areas. Existing units, new units, and lateral expansions may not be located in unstable areas unless it is demonstrated that engineering measures have been incorporated into the design to ensure that the integrity of the structural components of the unit will not be disrupted.

23.2.2 Clean Water Act

The regulations applicable to MSWLFs promulgated under the Clean Water Act are the wetlands regulations, discussed under Subtitle D location restrictions, and the National Pollutant Discharge Elimination System (NPDES) program. The NPDES program establishes discharge criteria and standards for the imposition of technology-based treatment requirements in permits. For MSWLF units, the NPDES program is applicable to discharges from storm water management systems and leachate treatment plants, as well as storm water discharges during construction.

23.2.3 Clean Air Act

Under the authority of the Clean Air Act, the EPA has promulgated the "Standards of Performance for New Stationary Sources and Guidelines for Control of Existing Sources: Municipal Solid Waste Landfills." These standards and guidelines regulate MSWLF emissions, commonly referred to as "landfill gas," which includes methane and non-methane organic compounds (NMOCs).

This regulation requires MSWLFs with a design capacity of greater than 2.5 million Megagrams (Mg) (2.75 million tons) or 2.5 million cubic meters (3.3 million cubic yards), and NMOCs emissions exceeding 50 Mg per year (55 tons per year), to design and install gas collection systems and then to combust the captured landfill gases. The control device is a flare which must be capable of reducing NMOCs emissions by 98% by weight. Other control systems that meet the 98% destruction criterion are acceptable, including gas utilization applications. When determining the design capacity and the NMOC emission rate, the entire landfill must be considered, even if portions are closed. If installation of collection and control systems is required, these systems are to be installed in the entire landfill.

The NMOC emission rate is calculated as a function of the mass and age of the refuse, the gas generation constant (k), and the concentration of NMOCs (C_{NMOC}). The regulation allows the calculation of gas emission based on conservative default values of k and C_{NMOC} provided by the EPA, or using site-specific values of C_{NMOC} and k determined from gas samples and pumping tests, respectively.

23.2.4 State and Local Regulations

State and local regulations vary significantly and may affect both the lateral and vertical limits of landfill development. These frequently include minimum setback distances from:

- Surface water bodies
- Sources of drinking water
- Public roadways
- Property lines
- Structures such as homes, schools, hospitals, and nursing homes
- Cultural resources such as historic parks, monuments, structures, and museums
- Recreational parks
- Designated wild and scenic rivers

- Coal seams or limestone outcrops
- Gas or oil wells.

Other regulatory aspects that may affect the design or permitting are:

- Zoning
- Buffers
- Deed restrictions
- Special protection areas
- Traffic/access road
- Height restrictions
- Groundwater quality protection
- Maximum/minimum final cover slopes.

23.3 Hydrogeologic and Geotechnical Considerations

To verify that the contemplated landfill development will not adversely affect groundwater resources, regulations require that a number of aspects be adequately addressed. These include:

- Depth to groundwater
- Delineation of regional confined and unconfined aquifers
- Direction and rate of groundwater flow, including possible spatial and temporal variations
- Aquifer characteristics including thickness and saturated (and, as applicable, unsaturated) hydraulic conductivity
- Regional uses of the groundwater
- Background groundwater quality.

A complete characterization of the hydrogeologic regime must include an understanding of the previous, current, and potential uses of the groundwater that occurs in the site vicinity. For example, in semi-arid agricultural areas, groundwater pumping to serve irrigation needs may seasonally lower the water table. Significant regional pumping and floodwater storage impoundments may also modify localized groundwater flow directions. Furthermore, in some areas of the country, a reduction or cessation in groundwater pumping for public water supplies has resulted in a significant rise in the regional groundwater table. In addition, background levels of chemical constituents in the groundwater may be due to the past or current existence of local industrial or agricultural facilities.

A comprehensive understanding of the site aquifer characteristics and soil permeability is also needed. For example, both saturated and unsaturated hydraulic conductivity may be required for predictive contaminant transport modeling. In addition, in the absence of other controlling criteria, the presence of a thick uniform stratum of near-surface low permeability soils should be considered when siting a landfill cell, since the presence of such a material may permit the deletion of a clay liner component.

Many state regulations provide specific requirements related to hydrogeologic and geotechnical field investigations, such as the minimum number and depth of test borings, test pits, piezometers, and monitoring wells. The number of test holes is usually based on the size of the proposed facility. Depths are based on the intended depth of the facility bottom, depth to the groundwater table, and presence of additional lower aquifer units. Some regulations may provide guidance on supplemental investigative techniques such as geophysics and other remote sensing methods. Some state regulations specify *in situ* hydrogeologic testing of aquifers such as individual well or multiwell pumping tests to determine aquifer characteristics. These tests may be essential in evaluating the presence of an aquitard. Frequencies of groundwater level measurements may be specified in order to construct appropriate seasonal or regional groundwater potentiometric surface maps.

Some state regulations specify a minimum vertical separation distance between the bottom of the MSWLF unit and the groundwater table. Some states allow alteration of the uppermost unconfined aquifer or perched aquifers to be able to meet the separation requirements, while others do not. An important design consideration is to obtain adequate groundwater level data to determine satisfactorily the potentiometric surface to be used to determine compliance with the separation criteria. This may require several months or a year or more of monitoring to determine seasonal high groundwater levels. Using groundwater elevations that do not represent the true seasonal high can lead to costly redesign later if groundwater elevations rise.

To properly determine potentiometric surfaces or seasonal high groundwater contours, it is important to select an adequate number and spread of piezometers, taking into account the site's topographic features and subsurface soil profile. Hydrogeologic controlling features, such as wetlands and streams, must also be considered when selecting the location of piezometers.

Often, specific geotechnical soil laboratory tests and frequencies are specified in hydrogeologic boreholes to characterize subsurface soils or to evaluate suitability of potential borrow soil sources for construction, operation, and closure of the landfill. This can avoid the duplication of drilling efforts for hydrogeological and geotechnical investigations.

Additional geotechnical factors which must also be considered in the formulation of investigation programs include:

- Presence of near-surface bedrock
- Interrelated topographic/geotechnical considerations (e.g., substantial cuts or fills)
- Stability of natural and cut slopes
- The presence of soft, compressible, collapsible, or otherwise unsuitable foundation soils
- Availability of cover materials and uses for any excavated soils.

23.4 Conceptual Design

The first step in the development of a new landfill project, after site selection, is preparation of a Site Master Plan. The Master Plan considers the facilities required at the site, site-specific constraints (setbacks, buffer zones, wetlands, etc.), and optimal use of the land. The main objectives of preparing the Site Master Plan are: (1) to optimize the use of the land to be used for waste disposal; (2) to ensure that the relative locations of the facilities are adequate for landfill operation; and (3) to limit future design changes to accommodate facilities that were not initially planned.

The type of facilities required and their sizes depend mainly on the total airspace desired, the expected waste stream, and specific regulatory requirements. Typically, a Site Master Plan includes at least the following facilities:

- Waste disposal area
- Buffer zones
- Access, entrance, and internal roads
- Sedimentation/detention basin(s)
- Scale(s) and scale house
- Office building
- Maintenance building
- Stockpile area(s)
- Leachate holding tank(s) or pond(s), if no leachate treatment is to be performed onsite, or leachate pretreatment/treatment plant if leachate treatment will be performed onsite
- Fence and gate.

Other facilities that may be included in a Site Master Plan are recycling and composting facilities, leachate truck loading facility, landfill gas flare, and onsite borrow areas.

23.4.1 Landfill Configuration

The main landfill elements, or systems, are:

- Landfill footprint
- Subbase grading
- Final cover grading
- Cell layout, including cell access ramps
- Perimeter and intercell berms
- Liner system section
- Final cover section
- Storm water management system
- Leachate collection and removal system
- Gas management system
- Environmental monitoring system.

The interdependence between the various landfill elements and the sequence in which they are considered in the design plays a very important role in landfill design. The most common landfill types, from a geometrical configuration point of view, are (Figure 23.1):

- *Area fill.* Landfilling progresses with little or no excavation. Normally used in relatively flat areas with shallow water table.
- *Above-ground and below-ground fill.* The landfill consists of a two-dimensional arrangement of large cells that are constructed one at a time. Once two adjacent cells are filled, the area between them is also filled. Normally used in relatively flat areas without a shallow water table.
- *Valley fill.* The area to be filled is located between natural slopes. It may include some below ground excavation.
- *Trench fill.* This method is similar to the above-ground and below-ground configuration, except that the cells are narrow and parallel. It is generally used only for small waste streams.

The design of a landfill must take into account the interdependence of the various landfill elements, which can lead to several design iterations. In order to minimize design iterations, the design is generally subdivided into two stages: conceptual design and detailed design.

The conceptual design comprises the basic elements that define the cost of the project and its economic feasibility. These basic elements are those related to optimal use of the land, the airspace, and the soils balance. The airspace should include an allowance for daily, intermediate, and final covers. Soils balance is defined as the comparison between the volume of soils to be excavated and the volume of soils required for the entire project. Soils are excavated at landfill projects for cell development and for sedimentation/detention basins. Excavated soils can be used for structural fill (subgrade fill, perimeter and intercell berms); daily, intermediate, and final covers; soil liners; sedimentation/detention basin berms; and roads. The soils balance applies not only to the total volume, but to soils of specific properties, since soils used for different purposes have different specifications.

The basic elements that must be included in the conceptual design are the landfill footprint, the subbase grading plan, the final grading plan, sedimentation/detention basins, borrow areas, stockpile areas, cell layout, and entrance and internal roads. Detailed design of the other elements, such as the leachate collection and removal system, the storm water management system, and the gas management system is not needed until the conceptual design has been completed.
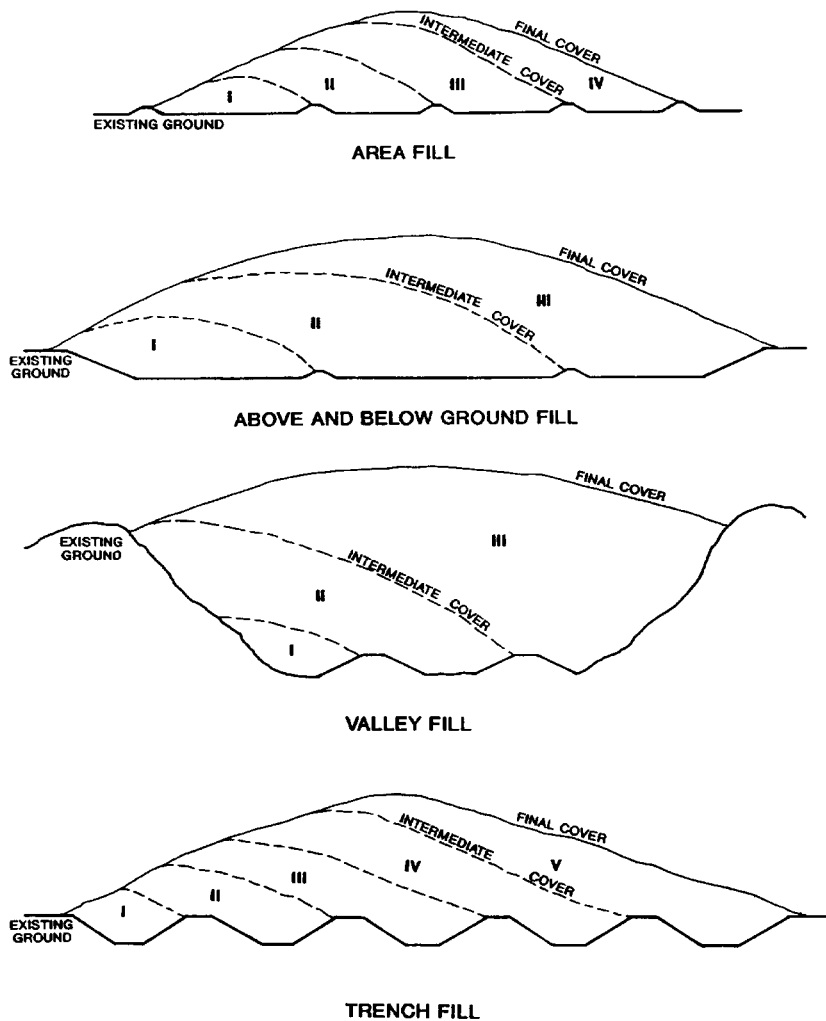
FIGURE 23.1 Landfill configurations.

Although the conceptual design must account for all the site constraints and regulatory requirements, it should be preliminary and simplified. The landfill elements included in a conceptual design are discussed below.

23.4.2 Landfill Footprint

The landfill footprint is generally selected by maximizing its area, considering the rest of the required facilities and site constraints. Prime importance must be given to critical issues that could impair permitting (e.g., wetlands, groundwater issues). At the conceptual design level, selection of the footprint does not need to be very accurate; for example, an allowance should be included for the width of perimeter berms, but the berms do not need to be designed.

The landfill footprint selected in the conceptual design stage will require some adjustments when other landfill elements are designed as part of the final design. However, if the conceptual design has been properly analyzed, the adjustments should be minor and should not significantly modify the soils balance or the location of other facilities at the site.

23.4.3 Subbase Excavation and Grading

Once the footprint has been preliminarily selected, the next step is to prepare a subbase grading plan, which identifies areas and depths of excavation, as well as areas and heights of fill. Selection of the landfill footprint and the subbase grading plan is an iterative process that depends significantly on the soils balance. At the conceptual design level, the subbase grading should be sufficient to obtain a preliminary estimate of excavation and fill volumes and, together with the final cover grading, to estimate the airspace. It should consider vertical separation from groundwater, side slopes, and overall leachate collection system slope. However, at this level it does not need to include detailed subbase grading (leachate collection swales) or design of the intercell berms.

It is frequently intended to design the subbase excavation as deep as possible, so that airspace is maximized and soils are made available for landfill development. However, the footprint shape and extent, cell layout, excavation depth, and subbase design are strongly influenced by a number of factors, such as:

- *Depth to the water table or uppermost aquifer.* Some regulations require a minimum separation from groundwater. Even if regulations do not require separation from groundwater, the cost of dewatering for construction would have a significant impact.

- *Depth to bedrock.* Some regulations require a minimum depth to bedrock. Also, the cost of rock excavation negatively impacts the economic feasibility of a project. Additionally, excavated rock generally has little use in landfill projects.

- *Stability of the foundation.* Weak foundation soils, as well as areas previously mined or susceptible to sinkhole development, may cause instability problems.

- *Site topography and stability of natural and cut slopes.* In flat areas, only fill embankments and cut slopes for cell excavation need to be considered, whereas excavations close to natural slopes may affect their stability and need to be considered in selecting the landfill footprint.

- *Stability of cut slopes.* The stability of cut slopes is a function of their height, inclination, groundwater conditions, and the strength and unit weight of the *in situ* soils. The inclination and height of the cut slopes affects the landfill airspace and the volume of soil to be excavated.

- *Soils balance.* The soils balance consists of the comparison of soils available and needed. An adequate soils balance allows a sufficient volume of soils to be available from onsite excavations and, at the same time, avoids excessive excavation that would have to be stockpiled onsite or disposed offsite.

- *Permeability of natural soils.* Sites having naturally low permeability surface soils and extensive non-water bearing formations below are generally preferable. In some cases, an alternative liner system that incorporates a natural clay layer may be acceptable, which would eliminate the need for low permeability borrow material for an engineered liner.

- *Required airspace and available area.* A minimum airspace may be required to make the landfill project feasible.

- *Waste stream.* The waste rate is directly related to the life of each cell. Normally cells are dimensioned so that they do not remain empty for an extended period of time.

- *Filling sequence.* One of the design goals is to avoid excessive stockpiling of excavated soils and to minimize leachate production by placing the final cover as soon as possible.

- *Grading of the landfill floor.* A minimum slope is required for the leachate collection system.

23.4.4 Sedimentation/Detention Basins

The location and size of sedimentation/detention basins is one of the basic elements to be considered at the beginning of the design process, since they frequently occupy significant area and influence the landfill footprint. Sedimentation/detention basins also influence the soils balance, since they require excavation

and/or the construction of berms. A detailed design of sedimentation basins is not required at the conceptual design level, since only their location, areal extent, and volume are important at this level.

23.4.5 Borrow Areas

The need for and the location of borrow areas must be evaluated as part of the conceptual design, since they have a significant influence on the landfill footprint and the soils balance. Borrow areas may be onsite or offsite. This decision depends on soil types available onsite, land availability, desired airspace, cost of an imported borrow, etc. It also depends on whether clay or a soil-bentonite admixture is used for a low permeability soil liner/cover. Generally, soils from cell excavation are used as borrow material; however, the need for stockpiling for future use must be considered. The selection of onsite borrow areas may affect the landfill footprint and the sequence of filling.

23.4.6 Final Cover Grading

At the conceptual design level, the final cover grading should be sufficient to obtain a preliminary estimate of the airspace. It should consider maximum and minimum regulatory slopes, any height restrictions, slope stability, and allowance for appropriate runoff control. However, at this level it does not need to include design of the benches/channels for runoff control; the influence of the channels/benches in the airspace can be considered by assuming an average slope.

The final grading is frequently designed as steep and high as possible, so that the maximum airspace is attained. However, both steepness and height need to be compatible with other requirements, such as regulatory restrictions regarding side slopes and/or height, overall stability of slopes and sliding of the final cover, and storm water management and erosion control.

23.4.7 Cell Layout and Filling Sequence

Landfills are generally subdivided into cells. The objective of dividing the landfill into cells is to provide separate operational units by constructing one cell at a time, so that the initial capital investment is limited and the period of time over which the liner would be exposed is minimized. Also, separate cells may be used to dispose of different types of waste.

Cell sizes are selected based mainly on the waste stream and the desired life of each cell. In general, it is unusual for cells to be less than two acres or more than eight acres. The desired life of each cell is generally decided as a function of the desired schedule of construction operations. When analyzing the life of individual cells, it should be remembered that isolated cells form a single mound, while cells adjacent to a previously filled cell also add a wedge on the interim slope of the previous cell. Therefore, the ratio volume/area of the cells increases with further cell development.

Other important factors to take into account in designing the cell layout are leachate collection and removal, and the filling sequence. Preferably, the cells are designed to slope toward the landfill perimeter, so that leachate can be removed without crossing other cells. The cell layout plan should show all the leachate sumps and riser pipes and bottom grading plan.

It is generally advantageous to design the cell layout and the filling sequence so that excessive stockpiling of excavated soils is avoided and part of the final slopes are formed as early as possible. Once a part of a final slope has been formed, final cover can be placed on that area. This reduces the cost of leachate treatment, since placement of final cover reduces infiltration into the waste and allows managing the runoff as storm water rather than leachate.

23.5 Detailed Design

The detailed design includes refining the design of the basic landfill elements. The following sections discuss only the additional landfill elements not included in the conceptual design.

The main geoenvironmental containment systems used in modern landfills are the bottom liner, the final cover, and the storm water management system. Other types of containment systems associated with landfills are cutoff walls and blanket drains. A detailed discussion of the layers that comprise liner and cover systems has been presented in several recent papers (Mitchell et al., 1990a,b; Koerner and Daniel, 1992; Mitchell and Mitchell, 1992; Repetto, 1995).

23.5.1 Liner System

Liner systems are containment elements constructed under the waste to control leachate infiltration into the subsoil or groundwater. The leachate may be part of the waste itself or may originate from precipitation that has infiltrated into the waste.

Landfill liner systems consist of multiple layers which fulfill specific functions. Landfill liner systems may consist, from top to bottom, of the following functional layers:

- *Protective layer*. This is a layer of soil, or other appropriate material, that separates the refuse from the rest of the liner to prevent damage from large objects.
- *Leachate collection layer*. This is a high-permeability layer, whose function is to collect leachate from the refuse and convey it to sumps, where it is removed. Frequently the functions of the protective layer and the leachate collection layer are integrated in one single layer of coarse granular soil.
- *Primary liner*. This is a low permeability layer (or layers of two different low-permeability materials in direct contact with each other). Its function is to control the movement of leachate into the subsoil.
- *Secondary leachate collection layer* (or leakage detection layer). This is a high-permeability (or transmissivity, if geosynthetic) layer designed to detect and collect any leachate seeping through the primary liner. This layer is used only in conjunction with a secondary liner.
- *Secondary liner*. This is a second (or backup) low permeability layer (or layers of two different low-permeability materials in direct contact with each other). Not all liner systems include a secondary liner.
- *Drainage layer*. In cases where the liner system is close to or below the water table, a high-permeability (or high-transmissivity, if geosynthetic) blanket drainage layer is generally placed under the liner system to control migration of moisture from the foundation to the liner system.
- *Subbase*. This layer is generally of intermediate permeability. Its function is to separate the liner system from the natural subgrade or structural fill.

These layers are normally separated by geotextiles to prevent migration of particles between layers, or to provide cushioning or protection of geomembranes.

Liner systems comprising only the primary liner are called *single liner systems*, and those having a primary and a secondary liner are called *double liner systems*. Further, each of the liners (primary or secondary) may consist of one layer only (low-permeability soil, geomembrane, or geosynthetic clay liner) or adjacent layers of two of these materials, in which case it is called a composite liner. Typical sections of double composite liner systems are presented in Figure 23.2. The leachate collection layer(s) may be constructed of a permeable soil, a geonet, or a geocomposite. The geocomposite used for leachate collection layers consists of a geonet welded between two nonwoven geotextile layers.

According to Subtitle D, the liner system must comply with either a design standard or a performance standard. The design standard under Subtitle D is a single composite liner comprised of a minimum 30-mil geomembrane (60-mil if HDPE) placed in direct contact with at least a two-foot layer of compacted soil with a hydraulic conductivity of no more than 1×10^{-7} cm/sec.

Any alternative liner system design must demonstrate compliance with the performance standard. The performance standard sets limits on the chemical concentrations in downgradient groundwater and is discussed later in this chapter.
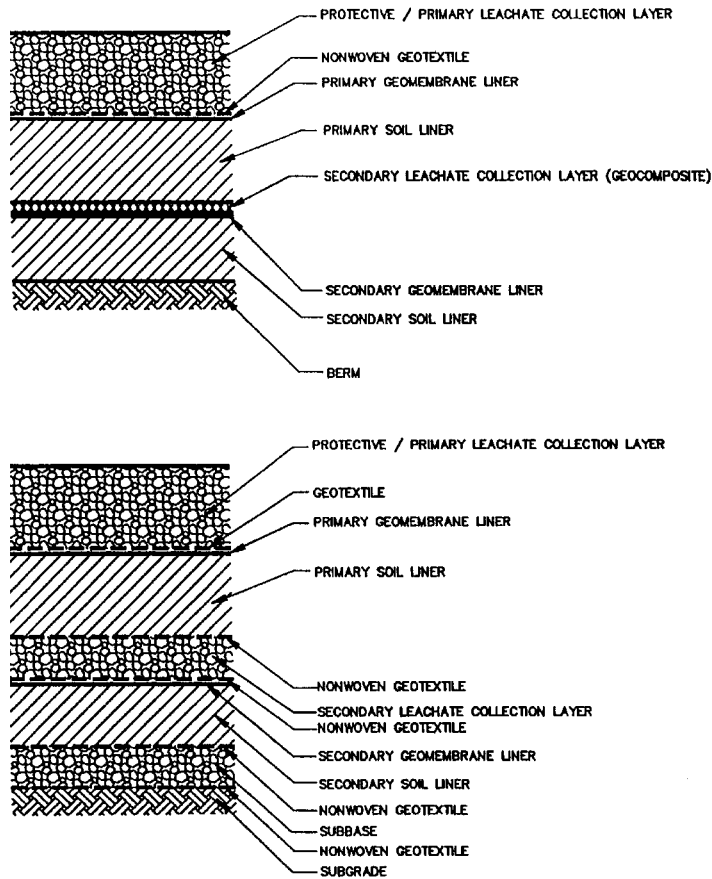
FIGURE 23.2 Double composite liner systems.

23.5.2 Final Cover System

A low-permeability cover is a containment system constructed on top of the waste, primarily to control the infiltration of precipitation. Cover systems control infiltration by providing a low-permeability barrier and by promoting runoff with adequate grading of the final surface. Other functions of cover systems are to prevent contact of runoff with waste, to prevent spreading or washout of wastes, to reduce disease vectors, and to control the emission of gases.

Cover systems (or caps) play a primary role in groundwater remediation at old landfill sites without a bottom liner system. In these cases, if the bottom of the waste is above the capillary fringe of the water table, the cover system may be an effective means of minimizing infiltration through the waste that would eventually migrate to the groundwater.

Cover systems also consist of multiple layers with specific functions related to the management of storm water. Of the precipitation that falls on the cover, part becomes surface water runoff, a portion infiltrates through the cover into the underlying materials, part is retained in the cover soil, and the remainder is lost through evapotranspiration. A low-permeability cover system may include, from top to bottom, the following functional layers (see Figure 23.3):

- *Erosion (vegetative) layer.* This is a layer of soil capable of supporting vegetation (typically grass) and with good resistance to erosion due to surface water runoff.
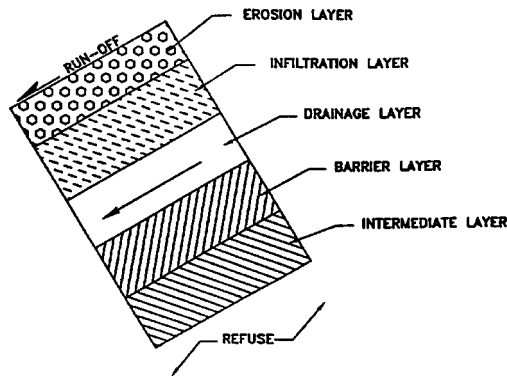
FIGURE 23.3 Final cover system functional layers.

- *Infiltration layer.* The functions of this layer, also frequently called cover material, are to separate the precipitation that does not evaporate into runoff and infiltration, and to protect the underlying barrier layer from frost penetration.
- *Drainage layer.* This is a permeable layer whose function is to convey the water that infiltrates into the cover system. A cover without a drainage layer is susceptible to be damaged by water seeping out of the cover.
- *Barrier layer.* This is a low-permeability layer (or layers of two different low-permeability materials in direct contact with each other), whose function is to control infiltration into the waste.

The erosion and infiltration layers are normally constructed of soils, since they must support vegetation and provide frost protection. The drainage layer can be constructed of a permeable soil, a geonet, or a geocomposite. The barrier layer can be constructed of compacted low-permeability soil, a geomembrane, a geosynthetic clay liner (GCL), or adjacent layers of two of these materials, in which case it is called a *composite barrier layer.* Some of these layers can also be separated by geotextiles.

The runoff can produce erosion of the cover surface and must be controlled by means of a storm water management system. The infiltrated water primarily flows through the drainage layer, which must have adequate hydraulic capacity to convey the expected flow rate without head buildup.

Subtitle D specifies that the final cover system must comprise, as a minimum, a 6-inch erosion (vegetative) layer capable of sustaining native plant growth, underlain by an 18-inch infiltration layer. The infiltration layer must have a permeability less than or equal to the permeability of any bottom liner system or natural soils, or 1×10^{-5} cm/sec, whichever is less. If the liner system includes a geomembrane, then the cover system must also include a geomembrane with a minimum thickness of 20 mils (60 mils if HDPE) to comply with the permeability requirement; in this case, the permeability requirement for the infiltration layer is not to exceed 1×10^{-5} cm/sec.

23.5.3 Berms

The design of landfills includes perimeter and intercell berms. The primary functions of the perimeter berms are to prevent the spreading of refuse outside of the disposal area during filling operations and to protect the landfill from storm water runon from surrounding areas. The intercell berms provide separation between active cells and the rest of the landfill area. This allows construction of new cells to be performed simultaneously to the filling of active cells, and to prevent runoff from inactive cells and undeveloped areas from coming into contact with waste.

The perimeter and intercell berms are subject to driving forces produced by refuse slopes and, therefore, require structural strength. Consistent with this requirement, landfill berms are constructed of structural fill consisting generally of a coarse-graded granular soil.

In addition to the landfill berms, sedimentation/detention basins may also include perimeter berms. These berms require low permeability in addition to strength.

23.5.4 Storm Water Management Systems

Storm water management systems typically consists of two parts. One part controls runoff on the landfill itself to prevent erosion of the final cover; the other part is external to the landfill, and its objective is to prevent runon to the landfill from the surrounding areas. Subtitle D specifies that storm water management systems be designed to collect and control at least the water volume resulting from the 24-hour, 25-year storm. Some state regulations also require verification of the outlet structures of sedimentation/detention basins for the 100-year storm.

The runoff control system on the landfill consists generally of a system of benches/channels located on the final cover. It is difficult to control runoff without benches, although it is possible for flat or short slopes. The vertical distance between the benches/channels is generally selected such that erosion of the final cover, calculated by the U.S. Department of Agriculture Universal Soil Loss Equation, does not exceed 2 tons/acre/year as recommended by the EPA (USEPA, 1982). An additional criterion for selecting the vertical distance between benches/channels is to prevent the surface water flow regime changing from sheet flow to shallow concentrated flow. By preventing this change in the flow regime, the storm water has less energy with which it can mobilize soil. Final cover channels are designed with minimal longitudinal slope, more or less paralleling the final cover contours. Depending on the flow velocity, the channels may be lined with grass or rip-rap.

The final cover channels generally discharge into gabion-lined downchutes that convey the flow downslope. A gabion lining is generally used for the downchutes because of their ability to withstand uneven settlements and the relatively high water velocities resulting from steep slopes. Sometimes pipes are also used to convey the flow downslope. The downchutes, in turn, discharge to other channels that convey the flow to sedimentation/detention basins. These other channels typically collect runoff from the final cover and runon from the surrounding areas, and discharge to sedimentation basins which are often converted to detention basins after closure of the facility.

Subtitle D and some state regulations typically require the design of runon and runoff control systems for the 25-year, 24-hour storm event. Methods outlined by the U.S. Department of Agriculture, Soil Conservation Service in the publication entitled, *TR-55: Urban Hydrology for Small Watersheds* (USDA, 1986), are generally used to calculate flow rates for the design of storm water control structures.

23.5.5 Leachate Collection and Removal Systems

The leachate collection system (LCS) is a drainage system located on the bottom of the landfill, immediately above the primary liner. The purpose of the LCS is to collect leachate from the waste and to convey it to sumps, where it is removed from the cell. The LCS comprises a high-permeability collection layer, which often contains perforated pipes in swales surrounded by gravel. The collection layer generally consists of gravel, and the perforated pipes are made of HDPE or PVC; crushing resistance of the collection pipes is a critical design factor. Subtitle D requires the LCS be designed to maintain less than one foot of leachate head over the liner. The LCS transfers the leachate to sumps located at the lowest point of each cell, preferably adjacent to the toe of the inside slope of the perimeter berm.

The leachate removal system may consist of either a gravity flow pipe passing through the perimeter berm, a riser pipe which parallels the inside slope of the perimeter berm, or a vertical manhole which daylights out of the final cover (see Figure 23.4). The latter two systems require submersible pumps to remove the leachate. Removal by means of a gravity flow pipe passing through the perimeter berm avoids the need for pumping, but requires a liner penetration that may affect liner performance.

The leachate removed from the cells is generally conveyed by a header pipe to holding tanks or ponds. The header pipe may be gravity or force, depending on the site topography. The holding tanks or ponds are generally designed to hold 30 days of leachate production. Some regulations require double contain-

LINER PENETRATION

VERTICAL MANHOLE

SIDESLOPE RISER PIPE
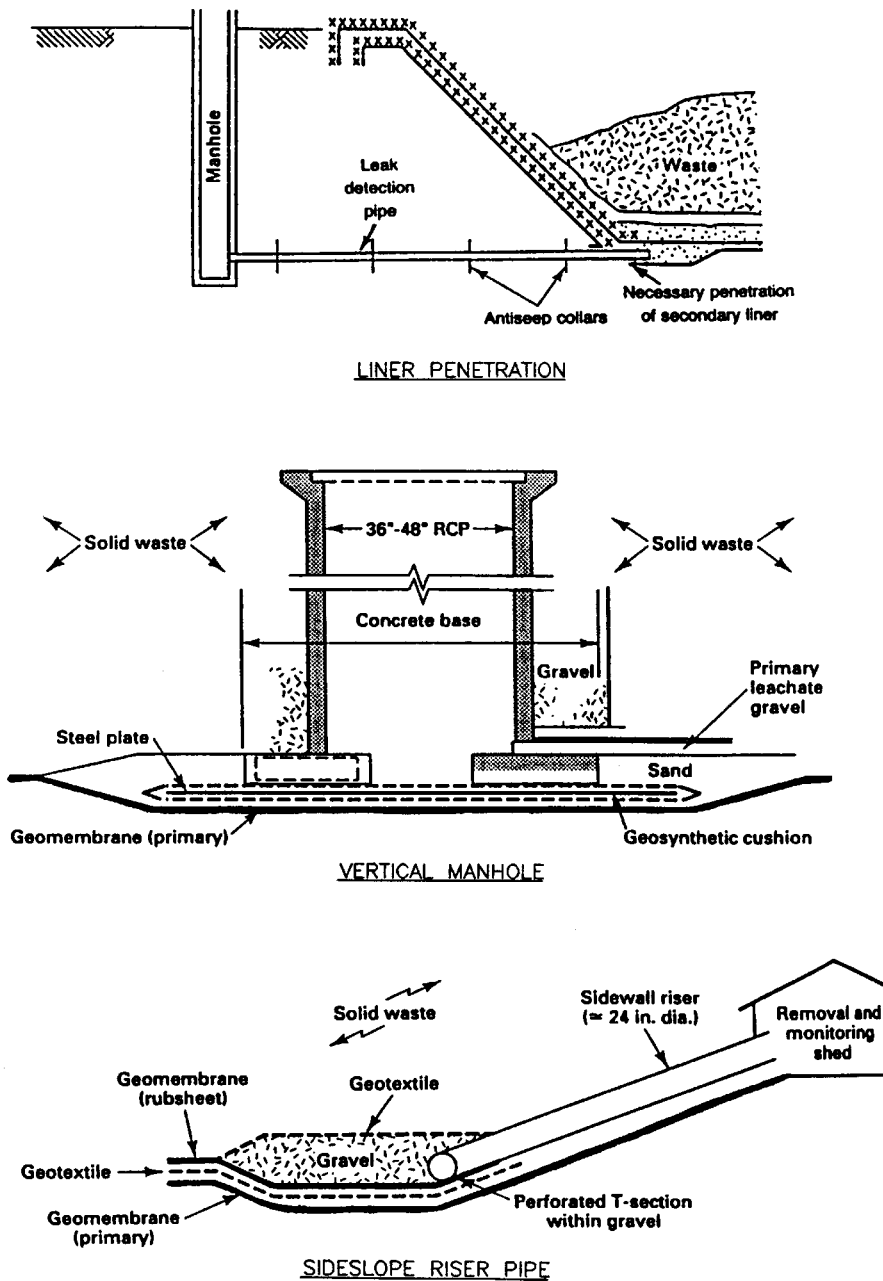
FIGURE 23.4 Leachate removal systems. (Courtesy Prof. R.M. Koerner.)

ment for leachate header pipes and holding tanks. Finally, the leachate may be discharged to an onsite treatment plant, or may be piped or transported by truck to an offsite treatment facility.

The size and spacing of the leachate collection pipes depend upon the amount of leachate generated. The amount of leachate generated can be calculated using two methods: the water budget (or balance) method and the Hydrologic Evaluation of Landfill Performance (HELP) model. The water budget method, based on the interrelationship of precipitation, evapotranspiration, runoff, soil moisture storage, and infiltration as discussed by Viessman et al. (1989), is a manual computation that can be performed

using a spreadsheet-type setup. This method does not, however, allow consideration of a geomembrane component in the cover.

Conversely, the HELP model, developed by the U.S. Army Corps of Engineers (USEPA, 1994), allows the inclusion of geomembranes within the cover or liner. This program requires input parameters similar to the water budget method, yet the program includes significant default data (e.g., climatological for 102 cities, soil and vegetation types, etc.). The results of the program include runoff, evapotranspiration, vertical drainage through the liner, percolation through the layers, peak daily leachate generation rates, average annual and monthly leachate generation rates, and lateral drainage from layers above the liner.

23.5.6 Gas Management Systems

Gas management systems can be passive or active. The type of system depends on state regulations, the total annual emissions of NMOCs, which in turn is a function of the size of the landfill, and site-specific issues related to odor control. As discussed under regulatory considerations related to the Clean Air Act, the federal regulation establishes that an active system must be installed if the design capacity of the MSWLF exceeds 2.75 million tons or 3.3 million cubic yards, and the total emissions of NMOCs exceed 55 tons per year.

A passive system is designed to vent the gas by providing a path for gas flow without the use of additional mechanical components. A passive system generally consists of perforated pipes embedded some depth into the waste to allow ventilation of the gas. Another system may include a high-permeability layer located under the cover system, into which vent pipes terminate.

The primary components for an active system are gas extraction wells or trenches, piping network, condensate collection traps, blowers or compressors with related power and control equipment, and disposal. From the piping network, the gas is transferred to the disposal point where it can be destroyed by means of open gas flares, or it can be utilized in energy generation, heating, or purified for other uses.

23.6 Liner System Performance Evaluation

In order to demonstrate the design performance required to justify using an alternative liner system design, it is necessary to conduct an assessment of leachate quality and quantity, leachate leakage rates through the liner system, and subsurface transport to the relevant point of compliance.

As mentioned above, the performance standard sets limits on the chemical concentrations in down-gradient groundwater for chemicals assumed to originate from the landfill. It requires that the maximum concentration levels (MCLs) of the 24 chemical parameters listed in Table 23.1 (reproduced from Subtitle D) not be exceeded in the uppermost aquifer at the relevant point of compliance. The relevant point of compliance may be located at the waste management unit boundary or, if approved by the director of an EPA-approved state, up to a maximum of 150 meters downgradient from the waste management unit boundary on land owned by the owner of the landfill. A site-specific alternative design in compliance with the performance standard may be approved by the director of an EPA-approved state. The following factors should be considered by the state director in determining the relevant point of compliance:

- The hydrogeologic characteristics of the facility and surrounding land
- The volume and physical and chemical characteristics of the leachate
- The quantity, quality, and direction of groundwater flow
- The proximity and withdrawal rate of the groundwater users
- The availability of alternative drinking water supplies
- The existing quality of the groundwater, including sources of contamination and their cumulative impacts on the groundwater, and whether the groundwater is currently used or reasonably expected to be used for drinking water
- Public health, safety, and welfare effects
- Practicable capability of the owner or operator.

TABLE 23.1 Maximum Concentration Levels

Chemical	MCL (mg/l)
Arsenic	0.05
Barium	1.0
Benzene	0.005
Cadmium	0.01
Carbon tetrachloride	0.005
Chromium (hexavalent)	0.05
2,4-Dichlorophenoxy acetic acid	0.1
1,4-Dichlorobenzene	0.075
1,2-Dichloroethane	0.005
1,1-Dichloroethylene	0.007
Endrin	0.0002
Fluoride	4
Lindane	0.004
Lead	0.05
Mercury	0.002
Methoxychlor	0.1
Nitrate	10
Selenium	0.01
Silver	0.05
Toxaphene	0.005
1,1,1-Trichloroethane	0.2
Trichloroethylene	0.005
2,4,5-Trichlorophenoxy acetic acid	0.01
Vinyl chloride	0.002

From Subtitle D, 40 CFR 258

The demonstration required for an alternative design should consider an assessment of leachate quality and quantity, leachate leakage to the subsurface, and subsurface transport to the relevant point of compliance. These factors are governed by the hydrogeology and climatic factors of the site. When leachate constituents can be expected to exceed MCLs, then a profile of leachate quality and production rates (volumes) to be used in evaluating its fate and transport from the point of release to the relevant point of compliance should be developed.

The U.S. EPA has developed a modeling package to perform the analyses required for this demonstration, called the Multimedia Exposure Assessment Model (MULTIMED). It is intended for sites where certain simplifying assumptions can be made to qualify its selection, since it does not account for:

- Site-specific spatial variability
- The shape of the landfill
- Site-specific boundary conditions
- Multiple aquifers and pumping wells, or
- Flow in fractures and chemical reactions between contaminants.

Table 23.2, reproduced from the Subtitle D Technical Manual (USEPA, 1993b) indicates the issues to be considered before applying MULTIMED.

Chapter D of the Subtitle D Technical Manual presents a comprehensive list of other models, some of which allow more accurate modeling when the site conditions are not as simple as those that can be represented in MULTIMED.

TABLE 23.2 Issues To Be Considered Before Applying Multimed

Objectives of the Study
- Is a "screening level" approach appropriate?
- Is modeling a "worst-case scenario" acceptable?

Significant Processes Affecting Contaminant Transport
- Does MULTIMED simulate all the significant processes occurring at the site?
- Is the contaminant soluble in water and of the same density as water?

Accuracy and Availability of the Data
- Have sufficient data been collected to obtain reliable results?
- What is the level of uncertainty associated with the data?
- Would a Monte Carlo simulation be useful? If so, are the cumulative probability distributions for the parameters with uncertain values unknown?

Complexity of the Hydrogeologic System
- Are the hydrogeologic properties of the system uniform?
- Is the flow in the aquifer uniform and steady?
- Is the site geometry regular?
- Does the source boundary condition require a transient or steady-state solution?

23.7 Groundwater Monitoring

Landfill projects must include groundwater and gas monitoring systems. If there are important streams nearby, surface water monitoring may also be required. Only groundwater monitoring is discussed in this chapter.

23.7.1 Subtitle D Requirements

As part of the overall environmental monitoring program for a landfill, a groundwater monitoring system and program must be developed to:

- Establish background groundwater quality levels
- Detect any releases of contamination from the landfill into the groundwater system
- Assist in formulating corrective action measures should a release occur.

The program is based, in part, on the results of hydrogeologic investigations conducted prior to facility operations, and on any applicable regulations established by local, state, and federal agencies.

The groundwater monitoring system specified by Subtitle D consists of a sufficient number of wells installed at appropriate locations and depths to yield groundwater samples from the uppermost aquifer at the site. A number of wells must provide samples representative of groundwater reaching the site (background) and others of the quality of groundwater passing the relevant point of compliance (downgradient). The purpose of comparing upgradient and downgradient groundwater quality is to determine if the landfill is impacting the aquifer.

In order to obtain the data necessary to design the groundwater monitoring system, a hydrogeologic investigation needs to be conducted to determine the hydrogeological regime at the site (aquifers, potentiometric surface, flow direction, gradient, etc.) and to obtain background groundwater quality. It should be noted that up to one year of background groundwater data may be required in some states prior to beginning landfill operations.

23.7.2 Site Hydrogeology

To determine the hydrogeological regime at the site and to facilitate the design of a groundwater monitoring system, hydrogeologic investigations must be conducted prior to the design of a landfill. Depending on site- or project-specific conditions, these investigations may include:

- Reviewing available literature and data concerning regional and site hydrogeology, including any information on wells at or near the site.
- Conducting field reconnaissance mapping to determine the extent and nature of bedrock geology beneath the site.
- Installing test wells and piezometers at the site and conducting aquifer tests (slug tests, pumping tests, and packer tests).
- Evaluating the hydrogeologic properties of the aquifer(s) beneath the site, based on data collected and on computer modeling, including delineating aquifer, aquiclude, and aquitard zones and determining the hydraulic conductivity, transmissivity, storativity, gradients, and flow directions of the various zones.
- Establishing background groundwater quality levels prior to the operation of the facility.

Data collected during these investigations are used to formulate a conceptual hydrogeologic and chemical model for the site which serves as the foundation for the design of a groundwater monitoring system.

23.7.3 Monitoring Wells

Upon completion of the hydrogeologic assessment, a permanent groundwater quality monitoring network and sampling plan for the site should be prepared. Pertinent aspects regarding monitoring wells include:

- The number and locations of monitoring wells are selected based on the site hydrogeologic conditions and on the requirements of regulatory agencies.
- Most agencies require a minimum number of wells be installed both upgradient (minimum of 1) and downgradient (minimum of 3) of the facility. In many cases, regulatory agencies require that more than the above minimum number of wells be installed.
- The number of wells can also be affected by the nature of the hydrogeologic system(s) beneath the site. Sites underlain by multiple, interconnected aquifers usually require that each aquifer be separately monitored.
- Following drilling and prior to well installation, additional aquifer tests may be conducted to refine the site hydrogeologic characterization and to ensure that the well will produce enough water to allow for sampling.
- Following construction, the wells must be properly developed to ensure that good communication is established between the aquifer and the well.

23.7.4 Monitoring Plan

In addition to the design of the groundwater monitoring system (well locations, depths, etc.), a monitoring plan should be prepared, including procedures for:

- Sample collection procedures
- Sampling frequency
- Sample preservation and shipment
- Analytical procedures
- Chain-of-custody control
- Testing parameters
- Quality assurance and quality control
- Statistical methods for data evaluation.

Typical sampling parameters for analysis include:

- Volatile organic compounds
- Semivolatile compounds
- Organic compounds
- Metals
- Other inorganic compounds.

Groundwater sampling procedures are formulated based on agency requirements and/or on established industry standards. A Field Sampling Plan is created which presents the detailed methodology to be employed during sampling. The ultimate accuracy of any data generation depends on a well-conceived and carefully implemented sampling and measurement procedure. A general description of the sampling process includes:

- *Purging.* Each well to be sampled must be purged by pumping or bailing. The wells must be purged until three well volumes of water have been removed and/or until the pH and conductivity of the water stabilize. This task is undertaken to ensure that stagnant water is removed from the well casing and that a representative aliquot of aquifer water is obtained.
- *Sampling.* Following purging, water samples will be removed from the well with a decontaminated (or disposable) sampling bailer and placed in labeled sampling jars. Typically, throughout the sampling process, pH, specific conductivity, and temperature are measured and recorded on field parameter information forms which also indicate the volume of water purged from the well, and the water level in the well before and after sampling.
- *Preservatives.* Preservatives must be added to samples that require them.
- *Cleaning.* All sampling equipment must be thoroughly cleaned with deionized water between sampling stations to ensure that cross contamination will not occur. Field personnel must also wear unpowdered latex gloves when handling the samples.
- *QA/QC Samples.* Specific Quality Assurance (QA) and Quality Control (QC) samples must also be collected. These QA/QC samples can include:
  – Equipment rinsate blanks
  – Trip blanks
  – Laboratory control spikes, matrix spikes, and duplicate samples
- *Transportation.* The samples must be placed in a cooler/transport shuttle, ice added, the chain-of-custody information completed, and the cooler sealed and shipped via overnight courier to the agency-approved and certified analytical lab. It is imperative that the samples be analyzed within the designated time frame since each parameter has a certain "holding" time during which each analysis must be performed.
- *Documentation.* Chain-of-custody forms document the handling and transport history of the sample vessels and coolers from the lab to the site and back to the lab. Strict chain-of-custody documentation must be established and maintained to ensure sample traceability from the time of collection through completion of analysis.
- *Analysis and Reporting.* The samples received by the laboratory must be kept in cold storage until the specified analyses are performed using appropriate, agency-approved analytical methods. Once the data have been tabulated, validation of the data must be conducted to assess any anomalous or aberrant values.

Regulations established under Subtitle D require the statistical analysis of each hazardous constituent in each well. Accordingly, the number of samples collected must be consistent with the chosen statistical procedures. The allowable statistical analyses include (quoted directly from Subtitle D):

- A parametric analysis of variance (ANOVA) followed by multiple comparisons procedures to identify statistically significant evidence of contamination

- An ANOVA based on ranks followed by multiple comparison procedures to identify statistically significant evidence of contamination
- A tolerance or prediction interval procedure in which an interval for each constituent is established from the distribution of the background data
- A control chart approach that gives control limits for each constituent
- Another statistical method that meets agency performance standards and can be justified to meet these standards.

Upon completion of the statistical analysis of groundwater chemistry data, the results for each well must be evaluated to determine if a statistically significant increase above background levels has occurred.

Two types of monitoring programs are included in Subtitle D: detection monitoring and assessment monitoring, as well as corrective measures. *Detection monitoring* is required at all monitoring wells and, at a minimum, must include monitoring for the 62 constituents listed in Appendix I of Subtitle D at least semiannually. However, the director of an EPA-approved site may delete any of the Appendix I constituents from the monitoring requirements if it can be shown that those constituents are not reasonably expected or derived from the waste. Initial background monitoring includes four independent samples from each monitoring well (background and downgradient) to be analyzed for the Appendix I (or approved alternate) constituents during the first semiannual event. The sample interval between successive samples, required to obtain independent samples, depends on the nature of the aquifer. The following minimum intervals are recommended by the U.S. EPA:

Lithologic Unit	Sampling Interval
Gravel	Daily
Sand	Daily
Silty sand	Weekly
Till	Monthly
Sandstone	Weekly
Basalt	Monthly

At least one sample from each well (background and downgradient) must be collected and analyzed at least semiannually during subsequent sampling events for the active life of the facility (including closure) and post-closure period.

If it is determined that a statistically significant increase exceeding Groundwater Protection Standards defined under 40 CFR 255 (h) or (i) for one or more of the monitored constituents has occurred, a notice must be placed in the operating record and the state director notified within 14 days of this finding. An assessment monitoring program must then be established within 90 days unless it has been adequately demonstrated that the contamination was caused by a source other than the landfill or that the statistically significant increase was due to an error in sampling, analysis, statistical evaluation, or natural variation in groundwater quality. Groundwater Protection Standards are:

- Maximum Contaminant Level (MCL) for constituents with an MCL, or
- Background concentration for constituents without an MCL
- For constituents for which the background level is higher than the MCL, an alternative groundwater protection standard may be established using health-based criteria.

Assessment monitoring includes sampling and analysis for the 213 constituents listed in Appendix II of Subtitle D within 90 days of triggering the assessment monitoring program and annually thereafter. At least one sample from each downgradient well must be collected and analyzed during each sampling event. If any Appendix II constituents are detected, a minimum of four independent samples from each well (background and downgradient) must be collected and analyzed to establish background for the constituents. A notice must be placed in the operating record and the state director notified within 14

days of identifying Appendix II constituents. Within 90 days of detection of Appendix II constituents, and on a semiannual basis thereafter, all wells must be resampled and analyzed for Appendix I constituents (or approved alternative list) and for Appendix II constituents previously detected. At least one sample from each well (background and downgradient) must be collected during each sampling event. The facility may return to detection monitoring if all of the detected Appendix II constituents are at or below background levels for two consecutive sampling events.

A groundwater protection standard must be established for each Appendix II constituent detected. If subsequent monitoring indicates concentrations of Appendix II constituents above background but below groundwater protection standards, the facility must continue assessment monitoring. If one or more Appendix II constituents are detected at statistically significant levels above the groundwater protection standard, a notice must be placed in the operating record and the state director and all appropriate local government officials notified within 14 days. Within 90 days, the nature and extent of the release must be characterized; at least one additional monitoring well installed; all adjacent landowners or residents over the contaminant plume notified; and assessment of corrective measures initiated.

During corrective action, alternative corrective measures must be evaluated, the appropriate remedy selected, and the selected corrective remedy implemented. Semiannual assessment monitoring must continue during the corrective action. Corrective action must continue until compliance with the groundwater protection standards is achieved for at least three consecutive years. Once compliance is achieved, the facility may return to detection monitoring.

For Further Information

USEPA. 1986. Office of Solid Waste and Emergency Response. *Technical Guidance Document: Construction Quality Assurance for Hazardous Waste Land Disposal Facilities.* EPA/530-SW-86-031.

USEPA. 1993. Office of Research and Development. *Quality Assurance and Quality Control for Waste Containment Facilities, Technical Guidance Document.* EPA/600/R-93/182.

USEPA. 1993. Office of Research and Development. *Proceedings of the Workshop on Geosynthetic Clay Liners.* EPA/600/R-93/171.

USEPA. 1993. Office of Research and Development. *Report of Workshop on Geomembrane Seaming.* EPA/600/R-93/112.

USEPA. 1982. Office of Solid Waste and Emergency Response. *Evaluating Cover Systems for Solid and Hazardous Waste.* EPA/SW-867.

USEPA. 1988. Risk Reduction Engineering Laboratory. *Guide to Technical Resources for the Design of Land Disposal Facilities.* EPA/625/6-88/018.

References

Fassett, J.B., Leonards, G.A., and Repetto, P.C. 1994. Geotechnical properties of municipal solid wastes and their use in landfill design, Proceedings WasteTech '94, Charleston, SC, National Solid Waste Management Association.

Koerner, R.M. 1994. *Designing with Geosynthetics*, 3rd Edition, Prentice Hall, Englewood Cliffs, NJ.

Koerner, R.M. and Daniel, D.E. 1992. Better cover-ups, *Civil Engineering*, May, 55-57.

Mitchell, R.A. and Mitchell, J.K. 1992. Stability evaluation of waste landfills, Proceedings ASCE Specialty Conference on Stability and Performance of Slopes and Embankments — II, Berkeley, CA, June 28-July 1, 1152-1187.

Mitchell, J.K., Seed, R.B., and Seed, H.B. 1990a. Kettleman Hills waste landfill slope failure. I: Liner system properties, *Journal of Geotechnical Engineering*, ASCE, 116, 4, 647-668.

Mitchell, J.K., Seed, R.B., and Seed, H.B. 1990b. Stability considerations in the design and construction of lined waste repositories, ASTM STP 1070. *Geotechnics of Waste Landfills — Theory and Practice.* A. Landva and G.D. Knowles, Eds. American Society for Testing and Materials, Philadelphia. 207-224.

Repetto, P.C. 1995. *Geo-environment*, in *The Civil Engineering Handbook* CRC Press, Boca Raton, chapter 25.

Repetto, P.C. and Foster, V.E. 1993. Basic considerations for the design of landfills, Proceedings First Annual Great Lakes Geotechnical/Geoenvironmental Conference, The University of Toledo, Toledo, OH.

United States Code of Federal Regulations, Title 40. 1991. Protection of the Environment, Part 258, *EPA Criteria for Municipal Solid Waste Landfills*. Washington, DC.

United States Department of Agriculture. 1986. Soil Conservation Service, Engineering Division, *TR-55: Urban Hydrology for Small Watersheds*. Washington, DC.

United States Environmental Protection Agency. 1982. Office of Solid Waste and Emergency Response. *Evaluating Cover Systems for Solid and Hazardous Waste*, SW-867. Washington, DC.

United States Environmental Protection Agency. 1988. Risk Reduction Engineering Laboratory. *Guide to Technical Resources for the Design of Land Disposal Facilities*. EPA/625/6-88/018. Washington, DC.

United States Environmental Protection Agency. 1993a. Office of Research and Development. *Quality Assurance and Quality Control for Waste Containment Facilities*, Technical Guidance Document EPA/600/R-93/182. Washington, DC.

United States Environmental Protection Agency. 1993b. Office of Solid Waste and Emergency Response. *Solid Waste Disposal Facility Criteria*. Technical Manual EPA530-R-93-017. Washington, DC.

United States Environmental Protection Agency. 1994. Office of Research and Development. *The Hydrological Evaluation of Landfill Performance HELP) Model*. User's Guide for Version 3 (EPA/600/R-94/168a) and Engineering Documentation for Version 3 (EPA/600/R-94/168b). Washington, DC.

Viessman, W. Jr., Lewis, G., and Knapp, J. 1989. *Introduction to Hydrology*, 3rd Edition, Harper & Row Publishers, New York.

Glossary

The definitions presented in this section have been extracted from the Code of Federal Regulations (CFR), 40 CFR 258, *EPA Criteria for Municipal Solid Waste Landfills*, and 40 CFR 261, *Identification and Listing of Hazardous Waste*. Definitions not available in the federal regulations were obtained from the Virginia Solid Waste Management Regulations.

Agricultural Waste   All solid waste produced from farming operations or related commercial preparation of farm products for marketing.

Commercial Solid Waste   All types of solid waste generated by stores, offices, restaurants, warehouses, and other nonmanufacturing activities, excluding residential and industrial wastes.

Construction/Demolition/Debris Landfill   A land burial facility engineered, constructed, and operated to contain and isolate construction waste, demolition waste, debris waste, inert waste, or combinations of the above solid wastes.

Construction Waste   Solid waste which is produced or generated during construction, remodeling, or repair of pavements, houses, commercial buildings, and other structures. Construction wastes include, but are not limited to, lumber, wire, sheetrock, broken brick, shingles, glass, pipes, concrete, paving materials, and metal and plastics if the metal or plastics are part of the materials of construction or empty containers for such materials. Paints, coatings, solvents, asbestos, any liquid, compressed gases or semiliquids, and garbage are not construction wastes.

Cover Material   Compactable soil or other approved material which is used to blanket solid waste in a landfill.

Debris Waste   Wastes resulting from land-clearing operations, including, but not limited to, stumps, wood, brush, leaves, soil, and road spoils.

Demolition Waste   Solid waste which is produced by the destruction of structures and their foundations and includes the same materials as construction wastes.

Garbage   Readily putrescible discarded materials composed of animal, vegetal, or other organic matter.

Hazardous Waste   The definition of hazardous waste is fairly complex and is provided in 40 CFR Part 261, *Identification and Listing of Hazardous Waste*. The reader is referred to this regulation for a complete definition of hazardous waste. A solid waste is classified as hazardous waste if it is not excluded from regulations as hazardous waste; it exhibits characteristics of ignitability, corrosivity, reactivity, or toxicity as specified in the regulations; or it is listed in the regulations (the regulations include two types of hazardous wastes: those from nonspecific sources and those from specific sources).

Household Waste   Any solid waste (including garbage, trash, and sanitary waste in septic tanks) derived from households (including single and multiple residences, hotels, motels, bunkhouses, ranger stations, crew quarters, campgrounds, picnic grounds, and day-use recreation areas).

Industrial Solid Waste   Solid waste generated by a manufacturing or industrial process that is not a hazardous waste regulated under Subtitle C of RCRA. Such waste may include, but is not limited to, waste resulting from the following manufacturing processes: electric power generation; fertilizer/agricultural chemicals; food and related products/by-products; inorganic chemicals; iron and steel manufacturing; leather and leather products; nonferrous metals manufacturing/foundries; organic chemicals; plastics and resins manufacturing; pulp and paper industry; rubber and miscellaneous plastic products; stone, glass, clay, and concrete products; textile manufacturing; transportation equipment; and water treatment. This term does not include mining waste or oil and gas waste.

Industrial Waste Landfill   A solid waste landfill used primarily for the disposal of a specific industrial waste or a waste which is a by-product of a production process.

Inert Waste   Solid waste which is physically, chemically, and biologically stable from further degradation and considered to be nonreactive. Inert wastes include rubble, concrete, broken bricks, whole bricks, and blocks.

Infectious Waste   Solid wastes defined as infectious by the appropriate regulations.

Institutional Waste   All solid waste emanating from institutions such as, but not limited to, hospitals, nursing homes, orphanages, and public or private schools. It can include infectious waste from health-care facilities and research facilities that must be managed as an infectious waste.

Leachate   A liquid that has passed through or emerged from solid waste and contains soluble, suspended, or miscible materials removed from such waste.

Liquid Waste   Any waste material that is determined to contain free liquids.

Liner   A continuous layer of natural or synthetic materials beneath or on the sides of a storage or treatment device, surface impoundment, landfill, or landfill cell that severely restricts or prevents the downward or lateral escape of hazardous waste, hazardous waste constituents, or leachate.

Litter   Any solid waste that is discarded or scattered about a solid waste management facility outside the immediate working area.

Monitoring   All methods, procedures, and techniques used to systematically analyze, inspect, and collect data on operational parameters of the facility or on the quality of air, groundwater, surface water, and soils.

Municipal Solid Waste   Waste which is normally composed of residential, commercial, and institutional solid waste.

Municipal Solid Waste Landfill Unit   A discrete area of land or an excavation that receives household waste, and that is not a land application unit, surface impoundment, injection well, or waste pile. A municipal solid waste landfill unit may also receive other types of RCRA Subtitle D wastes, such as commercial solid waste, nonhazardous sludge, small-quantity generator waste, and industrial solid waste. Such a landfill may be publicly or privately owned.

Putrescible Waste   Solid waste which contains organic material capable of being decomposed by microorganisms and causing odors.

Refuse   All solid waste products having the character of solids rather than liquids and which are composed wholly or partially of materials such as garbage, trash, rubbish, litter, residues from cleanup of spills or contamination, or other discarded materials.

Release Any spilling, leaking, pumping, pouring, emitting, emptying, discharging, injection, escaping, leaching, dumping, or disposing into the environment solid wastes or hazardous constituents of solid wastes (including the abandonment or discarding of barrels, containers, and other closed receptacles containing solid waste). This definition does not include any release which results in exposure to persons solely within a workplace; release of source, by-product, or special nuclear material from a nuclear incident, as those terms are defined by the Atomic Energy Act of 1954; and the normal application of fertilizer.

Rubbish Combustible or slowly putrescible discarded materials which include but are not limited to trees, wood, leaves, trimmings from shrubs or trees, printed matter, plastic and paper products, grass, rags, and other combustible or slowly putrescible materials not included under the term *garbage*.

Sanitary Landfill An engineered land burial facility for the disposal of household waste which is located, designed, constructed, and operated to contain and isolate the waste so that it does not pose a substantial present or potential hazard to human health or the environment. A sanitary landfill also may receive other types of solid wastes, such as commercial solid waste, nonhazardous sludge, hazardous waste from conditionally exempt small-quantity generators, and nonhazardous industrial solid waste.

Sludge Any solid, semisolid, or liquid waste generated from a municipal, commercial, or industrial wastewater treatment plant, water supply treatment plant, or air pollution control facility exclusive of treated effluent from a wastewater treatment plant.

Solid Waste Any garbage (refuse), sludge from a wastewater treatment plant, water supply treatment plant, or air pollution control facility, and other discarded material, including solid, liquid, semisolid, or contained gaseous material resulting from industrial, commercial, mining, and agricultural operations and from community activities. Does not include solid or dissolved materials in domestic sewage, or solid or dissolved materials in irrigation return flows or industrial discharges that are point sources subject to permit under 33 U.S.C. 1342, or source, special nuclear, or by-product material as defined by the Atomic Energy Act of 1954, as amended.

Special Wastes Solid wastes that are difficult to handle, require special precautions because of hazardous properties, or the nature of the waste creates management problems in normal operations.

Trash Combustible and noncombustible discarded materials. Used interchangeably with the term *rubbish*.

Vector A living animal, insect, or other arthropod which transmits an infectious disease from one organism or another.

Washout Carrying away of solid waste by waters of the base flood.

Yard Waste That fraction of municipal solid waste that consists of grass clippings, leaves, brush, and tree prunings arising from general landscape maintenance.

24
Groundwater Monitoring

Michael F. Houlihan and
Patrick C. Lucia
GeoSyntec Consultants

24.1 Fundamentals

24.1.1 Introduction

In this chapter on groundwater monitoring, a summary of approaches and techniques for sampling, analyzing, and evaluating the quality of groundwater is provided. In the remaining portion of this subsection on fundamentals, the purpose of groundwater monitoring is addressed and an overview of the regulatory requirements for groundwater monitoring is provided; also, strategies for developing

groundwater monitoring programs are provided. In other sections of this chapter, descriptions are provided for the design and installation of groundwater monitoring systems, sampling and analysis of groundwater, and evaluation of groundwater monitoring data. The primary focus of this chapter is on groundwater monitoring at sites where waste management activities have impacted, or could impact, groundwater quality.

24.1.2 Purpose of Groundwater Monitoring

Groundwater is used for a variety of purposes, including irrigation, drinking, and manufacturing. Groundwater is also the source of a large percentage of surface water. To verify that groundwater is suited for its purpose, its quality can be evaluated (i.e., monitored) by collecting samples and analyzing them. In simplest terms, the purpose of groundwater monitoring is to define the physical, chemical, and biological characteristics of groundwater. If the characteristics of the groundwater do not meet the requirements of its intended use or if the groundwater could be harmful to human health or the environment, then it may need to be remediated. Groundwater remediation is discussed in Chapter 25 of this handbook.

Typical ranges of physical, chemical, and biological characteristics of uncontaminated groundwater are presented in Table 24.1. As shown in the table, physical characteristics that define groundwater include temperature, the elevation of the potentiometric surface, geologic conditions, the presence of nonaqueous phase liquids, and flow direction. The chemical characteristics that define groundwater include pH, hardness, and the concentrations of contaminants (both naturally occurring and anthropogenic). To minimize the number of tests required to characterize groundwater quality, the presence of chemical constituents is sometimes evaluated by measuring "indicator parameters," including pH, total organic carbon (TOC), total organic halides (TOX), and specific conductance; these parameters may be indicative of the presence of elevated concentrations of organic, chloride-based, or metallic constituents, respectively, in groundwater. The biological characteristics of groundwater include the presence and type of microorganisms (e.g., coliforms, enteric viruses, etc.) and biological oxygen demand. Knowledge of these properties is essential when evaluating the quality and suitability of groundwater for its intended use.

The purpose of a groundwater monitoring program should be defined before monitoring begins so that appropriate procedures, techniques, and analyses can be planned that will meet the specific needs of the project. *Detection* monitoring programs are used to detect an impact to groundwater quality. *Assessment* monitoring programs are used: (1) to assess the nature and extent of contaminants that have been detected in groundwater; and (2) to collect data that may be needed to perform a design for remediation of the groundwater. *Corrective action* monitoring programs are used to assess the impact of a groundwater remedy on contaminant concentrations as a tool in evaluating the success of a remedy. *Performance* monitoring programs are used to evaluate the effectiveness of an element of a groundwater remediation system in meeting the design criteria for the element. These four types of monitoring programs, which are commonly required at facilities regulated under the Resource Conservation and Recovery Act (RCRA), are the focus of this chapter. More detailed descriptions of the requirements of these programs are provided by the United States Environmental Protection Agency (U.S. EPA) (1986a, 1992a, and 1993c) and may also be contained in other applicable state, local, or federal guidance documents. Although the monitoring programs described in this chapter are based on the requirements of RCRA, the elements of these monitoring programs may also be useful and applicable to non-RCRA groundwater monitoring programs, such as groundwater monitoring programs at Superfund sites, at sites that are being voluntarily remediated, and at sites where a legally binding agreement has been made (e.g., a "consent agreement") requiring monitoring of the site. A flow diagram showing the typical sequence of groundwater monitoring activities is presented in Figure 24.1.

TABLE 24.1 Key Physical, Chemical, and Biological Properties of Groundwater

Category	Property	Standard Range of Values in Natural Groundwater
Physical	Temperature	10°–20°C
	Elevation of potentiometric surface, or depth below land surface	Varies widely
	Presence of NAPLs	None
	Total suspended solids	100–500 ppm
	Total dissolved solids	100–1,000 ppm
Chemical	pH	6.5–8.5 standard units
	Dissolved oxygen	2–5 ppm
	Total organic carbon	1–10 ppm
	Specific conductance	100–1,000 μmhos/cm
	Manganese	0–0.1 ppm
	Iron	0.01–10 ppm
	Ammonium	0–2 ppm
	Chloride	2–200 ppm
	Sodium	1–100 ppm
	Volatile organics	<40 ppb
	Sodium, calcium, bicarbonate, magnesium	1,000 to 1,000,000 ppb
	arsenic, cadmium, iron, lead, nickel, selenium, zinc	1 to 1,000 ppb
	Beryllium, Mercury, Silver, Thallium	Less than 1 ppb
Biological	Coliforms, viruses	0 organisms

Data from Driscoll, F.G. 1986. *Groundwater and Wells*. Johnson Division, St. Paul, MN.; Freeze, R.A. and Cherry, J.A. 1979. *Ground Water*. Prentice–Hall, Inc., Englewood Cliffs, NJ.; Hem, J.D. 1989. Study and Interpretation of the Chemical Characteristics of Natural Water. U.S. Geological Survey Water Supply Paper 2254, 3rd Edition.

24.1.3 Regulatory Requirements for Groundwater Monitoring

Federal and state regulations have been established to provide for protection of groundwater resources. An extensive discussion of these regulations is provided in Chapter 22 of this handbook. As described in Chapter 22, these regulations require monitoring of groundwater quality near operations where materials are managed that could have an adverse impact on groundwater quality. The following questions are provided as a guide to identifying applicable groundwater monitoring regulations for a particular site or operation.

- What is the activity (e.g., underground storage of petroleum hydrocarbon products) that could have an adverse impact on groundwater quality? Are there related activities (e.g., transmission of petroleum products to the storage tank through underground pipes) that could also have an adverse impact on groundwater quality?
- What regulations govern the activity or activities, and what portions of those regulations address either groundwater monitoring or potential impacts to groundwater quality?
- What specific requirements do the regulations contain for planning groundwater monitoring events, for sampling groundwater, for analyzing groundwater samples, and for evaluating data from groundwater sampling events?
- Do the regulations contain specific requirements for action in the event that impacts to groundwater quality are identified during the monitoring program?

Once applicable regulatory requirements have been identified, a program can be developed for addressing each applicable regulatory requirement. A useful tool for addressing the requirements is the groundwater monitoring plan. Guidelines for developing a groundwater monitoring plan are presented in Section 24.2.

PREPARE GROUNDWATER MONITORING PLAN
STEP 1. Understand hydrogeological/geochemical setting
STEP 2. Evaluate background conditions
STEP 3. Design groundwater monitoring system
STEP 4. Establish sampling & analysis methods and data quality objectives
STEP 5. Establish data evaluation methods
STEP 6. Develop response criteria and

INSTALL GROUNDWATER MONITORING FEATURES
STEP 1. Select appropriate installation technique
STEP 2. Procure appropriate monitoring feature construction materials
STEP 3. Install monitoring features according to design presented in Groundwater Monitoring Plan
STEP 4. Update conceptual model of site hydrogeology based on findings of installation program

SAMPLE/ANALYZE GROUNDWATER
STEP 1. Inspect wells and area surrounding wells
STEP 2. Measure depth to groundwater and total depth of well
STEP 3. Evacuate groundwater from the well
STEP 4. Collect, preserve, ship groundwater samples
STEP 5. Complete documentation of sampling event
STEP 6. Analyze groundwater samples for presence of contaminants

ANALYZE DATA; PERFORM CONTINGENCY ACTIONS
STEP 1. Select appropriate statistical analysis method
STEP 2. Perform statistical analysis of data
STEP 3. If statistically significant increase is detected, reevaluate validity of analysis
STEP 4. If statistically significant increase is confirmed, then implement appropriate response actions

FIGURE 24.1 Sequence of activities for groundwater monitoring.

24.1.4 Groundwater Monitoring Strategy Considerations

After identifying applicable regulatory requirements and before preparing a groundwater monitoring plan, the overall strategy of the groundwater monitoring program should be defined to guide the development of the plan. In this sense, "strategy" refers to the manner in which a hypothetical release from a regulated unit will be detected or measured. Examples of issues that should be addressed when developing a monitoring strategy include: (1) the type of monitoring data needed; (2) the locations (both horizontal and vertical) from which the samples are to be collected (i.e., definition of "target monitoring zones"); (3) the manner in which the samples will be obtained; and (4) the ability of the monitoring features to rapidly detect a change in groundwater quality. For detection monitoring programs, the types of data needed are usually defined by regulation; for other types of monitoring programs, the types of data needed are typically based on site-specific considerations. Selection of *target monitoring zones* is described in Section 24.2; design of groundwater monitoring features is described in Section 24.3; and the ability of a monitoring system to rapidly detect an impact on groundwater quality is discussed in Section 24.2.2 (see "Phase II" of Step 3 of a groundwater monitoring system design). See Chapter 2 of Nielsen (1991) for additional discussion of groundwater monitoring strategy.

Development of a groundwater monitoring strategy is illustrated in Figures 24.2 and 24.3. As shown in these figures, the potential sources of contamination and the aquifers of concern should be characterized before developing a groundwater monitoring strategy because selection of target monitoring zones cannot be made until the source and the aquifer of concern have been evaluated, usually through a detailed hydrogeologic evaluation of the site. When evaluating the ability of a monitoring system to rapidly detect a release from the potential source, the impact of preferential flowpaths and vertical gradients should be carefully evaluated; a two-dimensional analysis of groundwater elevation may not
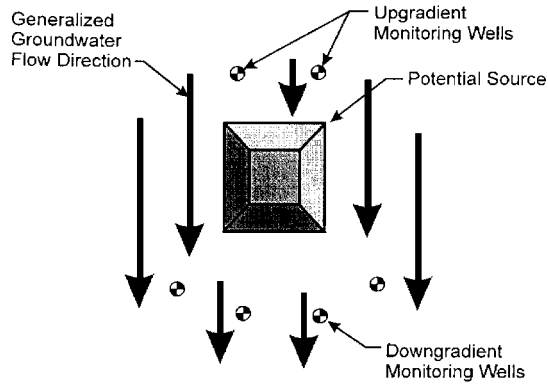
FIGURE 24.2 Plan view of typical unconfined aquifer groundwater monitoring system.
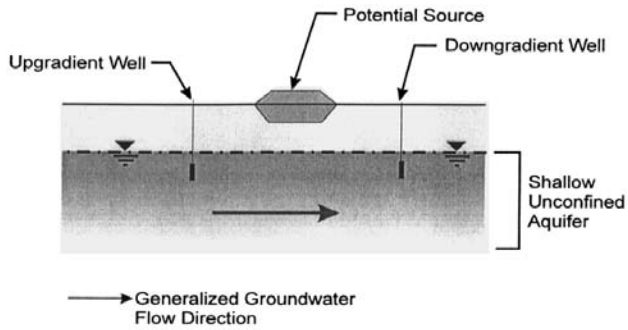
FIGURE 24.3 Section of typical unconfined aquifer groundwater monitoring system.
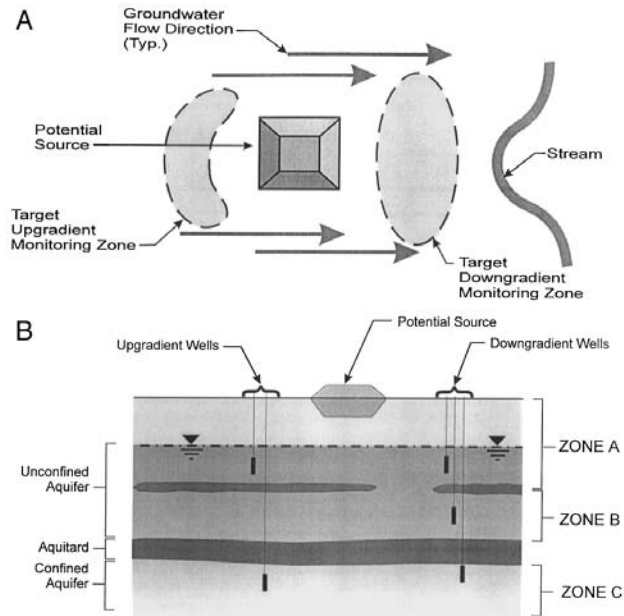
FIGURE 24.4 (A) Horizontal target monitoring zones; (B) vertical target monitoring zones.

reveal actual upgradient or downgradient locations of groundwater flow. The presence of vertical gradients may significantly affect the selection of monitoring locations.

24.2 Groundwater Monitoring Plans

24.2.1 Introduction

Successful implementation of a groundwater monitoring strategy requires a well-planned approach and an understanding of the factors that could affect the quality, validity, or representativeness of groundwater samples. A groundwater monitoring plan is an effective tool for systematically addressing these factors. Groundwater monitoring plans are required on a site-by-site basis for facilities that are regulated under RCRA and the Comprehensive Emergency Response and Compensation Liability Act (CERCLA), and may also be required for facilities that are regulated under various state and local programs. Groundwater monitoring plans are usually prepared to describe each aspect of groundwater monitoring and to control monitoring activities so that they fulfill the overall goals of the groundwater monitoring strategy. A comprehensive groundwater monitoring plan addresses each activity that will occur during sampling, analysis, and data interpretation, as well as response actions to be taken based on the results of monitoring. In this section, a procedure is presented for developing a thorough groundwater monitoring plan. Additional guidance on addressing the elements described in this section is presented by USEPA (1992a) and Nielsen (1991). An example outline for a groundwater monitoring plan is presented in Table 24.2. For brevity, and because the steps are similar for each type of program, the steps are described in significant detail for detection monitoring programs (in Section 24.2.2) but in less detail for other types of programs.

24.2.2 Detection Monitoring Programs

Step 1: Understand Hydrogeologic/Geochemical Setting

A comprehensive understanding of the hydrogeologic and geochemical setting of the site is an essential prerequisite to designing a groundwater monitoring system that can reliably detect a release to groundwater. This understanding should be based on an analysis of hydrogeologic and geochemical data from site explorations, literature reviews, and experience with similar hydrogeologic and geochemical settings. The types of hydrogeologic and geochemical information that should be collected and the general data needs for each type of information are described below, along with recommended means for obtaining the data.

Aquifer Hydrogeologic Parameters. The transmissivity, porosity, and storage properties (i.e., specific yield or specific capacity) of the aquifer should be evaluated. Transmissivity and storage properties are typically calculated based on the results of aquifer tests, as described in Chapter 8 of this handbook and by Freeze and Cherry (1979). Porosity is typically determined in the laboratory or estimated from the literature. Effective porosities can be estimated for bedrock aquifers by assuming that fractured rock acts as an equivalent porous medium.

Geologic Information. Characterization of the geologic units beneath the site is important to developing an understanding of the pathways that contaminants could take if introduced into groundwater (fate and transport of contaminants is described in detail in Chapters 14 and 15 of this handbook and in USEPA [1989c]). Examples of information that should be obtained before installing groundwater monitoring features include the horizontal and vertical limits of the aquifers, horizontal and vertical limits of *confining units* or *aquitards*, presence of interconnections between aquifers, *anisotropies* in aquifer material, presence of discontinuities (e.g., fractures, solution cavities, channel deposits, etc.) within or between stratigraphic units, gradual variations of the stratigraphic units with depth or with horizontal location, and the nature of the earth materials within the aquifers (e.g., particle size, angularity, dispersion-related properties, etc.).

Topographic Information. The surface topography of a site may provide important clues to the presence or extent of geologic/stratigraphic units at the site and, accordingly, may provide an indication of the

TABLE 24.2 Sample Outline — Groundwater Monitoring Plan

potential impact of a release of contaminants to groundwater. An understanding of site topography may provide valuable insight into: (1) the manner in which contaminants may have reached groundwater after being released to the ground surface; (2) past waste management practices at the site; (3) the interconnection between surface water and groundwater; and (4) aspects of the depositional history of the subsurface soils, which may give clues to the subsurface lithology and the manner in which it might control groundwater flow and contaminant transport. A discussion of terrain evaluation approaches and their use in site investigations is provided by Lueder (1959).

Potentiometric Surface Information. Knowledge of the potentiometric surface elevation is essential to understanding the direction of groundwater flow and the fate of contaminants. Potentiometric surface information is also important for evaluating whether an aquifer is unconfined, confined, or perched. Potentiometric surface information should be obtained for the uppermost aquifer and for any underlying aquifer that could be interconnected with the uppermost aquifer, as well as unconnected aquifers that could be impacted by a release from the potential source. Subsurface monitoring data can sometimes be supplemented with surface-water elevation data (e.g., springs, streams, etc.) to provide a better understanding of groundwater flow patterns.

Geochemical Conditions. Contaminants can chemically react with geologic media. Several types of reactions are common in aquifers (e.g., sorption, chelation, complexation, ion exchange, precipitation/dissolution, and biodegradation, which are described in Chapters 14 and 15 of this handbook). Characterizing the geochemical properties of the aquifer materials is important when evaluating the effect of these reactions on fate and transport of contaminants. Detailed discussions of groundwater geochemistry are provided in Chapter 13 of this handbook and by Chappelle (1993).

Step 2: Evaluate Background Conditions

To evaluate the impact of a potential source on groundwater quality, the characteristics of nonimpacted (i.e., background) groundwater must first be evaluated. Background water quality data provide a baseline against which the results of future groundwater monitoring events can be compared. When establishing background conditions, a careful examination of the site should be made to verify that the location selected for the background sample is unaffected by either the potential source or by other sources; for example, locations that could be impacted by incidental spills (such as waste material storage areas or parking areas) should not be selected as background sampling locations.

Step 3: Design Groundwater Monitoring System

After hydrogeologic conditions have been evaluated (see Step 1 above) and background conditions have been defined (see Step 2 above), the groundwater monitoring system can be designed. A three-phased procedure for designing a groundwater monitoring system is described below.

Phase I: Select Monitoring Locations. To rapidly detect a release from the source to groundwater, groundwater monitoring features should be installed within aquifers at locations (i.e., in "Target Monitoring Zones") that could be impacted by a release from the potential source. Selection of target monitoring zones is illustrated in Figure 24.4 and is described in more detail by Nielsen (1991, Chapter 2). As illustrated in Figure 24.4A, the target monitoring zones should be selected to intercept potential contaminant migration pathways, which are described in Step 2 and discussed in detail in Chapters 14 and 15 of this handbook. For example, the presence of LNAPLs (i.e., "floaters") or DNAPLs (i.e., "sinkers") in groundwater would require that monitoring features be placed at the water table or at the top of confining units, respectively. As illustrated in Figure 24.4B, the uppermost aquifer and any interconnected aquifers should be monitored for a release from the potential source. Also, the presence of preferential flow paths (e.g., solution cavities, channel deposits, etc.) must be considered when selecting monitoring feature locations (see Chapter 7).

Phase II: Select Monitoring Features. Appropriate monitoring features should be selected for obtaining the required samples or data from the target monitoring zones. Monitoring features may include groundwater monitoring wells, piezometers, soil gas probes, lysimeters, or sampling probes (as described in Section 24.3). Monitoring programs often incorporate a combination of these features. As an example of selection of monitoring features, Zones A and B in Figure 24.4B could be monitored by installing and sampling monitoring wells, with the wells in Zone B installed at a wider spacing for the purpose of confirming that contaminants have not migrated from the upper aquifer. The features should be installed within the target monitoring zones in locations that are likely to detect a release from the potential source. Also, monitoring wells should be spaced closely enough to the source to minimize the chance that contaminants will migrate undetected beyond the wells. If there are uncertainties about the ability of the proposed monitoring system to detect a release (i.e., the "efficiency" of the monitoring system), then the efficiency can be estimated using a computer program having a particle-tracking capability (e.g., MEMO [Wilson et al., 1992] or MODFLOW [McDonald and Harbaugh, 1988]) to identify zones within the potential source area from which a release would not likely be detected (see Chapter 20).

Phase III: Design the Monitoring Features. Finally, after the monitoring features have been identified, they should be designed as described in Section 24.3 to meet the specific goals of the monitoring program and to provide accurate, representative samples of groundwater.

Step 4: Establish Sampling and Analysis Methods and Data Quality Objectives

Each aspect of groundwater sampling and analysis, from extraction of the sample from the aquifer to laboratory analysis, can affect the results of the monitoring event. Therefore, it is important to specify sampling and analysis procedures that promote collection of samples representative of groundwater quality. Use of appropriate sampling and analysis methods (as described in Sections 24.4 and 24.5) promotes development of consistent, high-quality data and minimizes the possibility of invalid or incomplete data. The analytes should include the appropriate physical and chemical parameters identified in Table 24.1, as well as specific contaminants that could be present in the potential source and their

degradation products. Guidelines for sampling are presented in Section 24.4; guidelines for analysis of groundwater samples are presented in Section 24.5 and summarized in Table 24.3. When selecting sampling and analysis methods, a decision should be made regarding the quality of data needed to support decisions regarding the presence of groundwater contamination. To guide such decisions, data quality objectives (DQOs) can be established in the groundwater monitoring plan. DQOs can be specified as either qualitative (i.e., as a specific set of procedures to follow for collecting data) or quantitative (i.e., as the amount of imprecision or bias error that may be tolerated in data without incurring an unacceptable probability of making incorrect or inappropriate decisions). A more detailed discussion of DQOs is presented by USEPA (1992a, 1987b).

TABLE 24.3 Common Groundwater Analytical Methods

Technique/Instrumentation	Technology Status[a]	Sample Matrix[b]	Contaminant Type[c]	Detection Limit[d]
Chemical Field Measurement Techniques/Sensors				
pH/Alkalinity/Acidity	I/CP	W,S	—	—
Eh	I/CP	W,S	—	—
Dissolved Oxygen	I/CP	W	—	ppm
Temperature	I/CP	W	—	—
Electrical Conductance	I/CP	W,S	—	—
Sample Extraction Procedures				
Vacuum Extraction	I/CP	A	VOC	—
Purge and Trap	I/CP	W	VOC	—
Gaseous Phase Analytical Techniques				
Photo-Ionization Detector	I/CP	A	VOC	ppb–ppm
Flame-Ionization	I/CP	A	VOC	ppb–ppm
Explosimeter	I/CP	A	VOC	%
Luminescence/Spectroscope Techniques				
X-Ray Fluorescence	II/CP,CM, CF	S,W	HM	10s–100s ppm
Infrared Spectroscopy	II/CP	A,W,S	M,VOC,SVO	ppm–1000s ppm
Wet Chemistry				
Chemical Colorimetric Kits	II/CP	W	EA,HM,SVO	ppb–100s ppm
Titrimetry	I/CP	W	EA,HM,TR	ppb–100s ppm

[a] I = Well-established and routinely used field technology; II = Well-established laboratory technology for which experience in field applications is moderate to limited. CP = Commercially available portable instruments; CF = Commercially available field-able instruments; CM = Commercial/custom mobile laboratories available.

[b] A = Air/gaseous matrix; S = Soil/solid matrix; W = Water/aqueous/liquid matrix. Volatiles and semivolatiles in water and solid samples can be extracted for analysis by gaseous phase analytical techniques. Similarly, analytes can be extracted from solid samples for analysis using wet chemistry techniques.

[c] EA = Elemental/ionic analysis; HM = Metals; M = Mineralogy; P = Physical characterization; R = Radioisotopes; SVO = Semivolatile organics; TR = Tracer studies; VOC = Volatile organic compounds.

[d] Ranges for specific instruments and analytes might differ from range shown by orders of magnitude. In general, detection limits for soils will be higher than for groundwater.

From USEPA. 1993a. Office of Research and Development. *Subsurface Characterization and Monitoring Techniques: A Desk Reference Guide, Volume I: Solids and Ground Water, Appendices A and B.* U.S. Environmental Protection Agency, Washington, DC. EPA/625/R-93-003a.

Step 5: Establish Data Evaluation Methods

After the groundwater samples have been analyzed, the data should be evaluated to provide a basis for concluding the presence or absence of groundwater contaminants. Such evaluations typically involve statistical analysis of the data to identify significant differences in groundwater quality. To guide the data evaluation process, the groundwater monitoring plan should contain descriptions of the methods that

will be used to evaluate the monitoring data and the criteria for initiating response actions. Data evaluation methods are described in Section 24.6 of this chapter.

Step 6: Develop Response Criteria and Response Actions

Response criteria are results of groundwater monitoring events for which specific responses must be implemented. For example, in a detection monitoring program, an appropriate response action for the response criteria "*MCL is exceeded at point-of-compliance wells*" could be "*begin assessment monitoring.*" Examples of response criteria for a detection monitoring program and potentially appropriate response actions are provided in Table 24.4. The response actions for a detection monitoring program should be designed to confirm the presence of the detected contaminant or the statistically significant increase; detailed assessment or remediation of the detected contaminants may be delayed until the start of an assessment monitoring program (see Section 24.2.3) or a corrective action monitoring program (see Section 24.2.4). The groundwater monitoring plan should also identify follow-on activities for each possible outcome of a response action (e.g., "*prepare, submit, and implement an assessment monitoring program*" if the presence of contaminants is confirmed, or "*continue detection monitoring program*" if the presence of contaminants is not confirmed).

TABLE 24.4 Example Response Criteria and Corresponding Response Actions for Detection Monitoring Program

Response Criteria	Response Action
• Significant change in groundwater elevation	• Consider seasonal effects on water level, or impacts from groundwater withdrawals (e.g., water supply or irrigation wells)
• Change in total depth of well	• If necessary, remeasure groundwater levels • Redevelop well • If necessary, reconstruct well
• Evidence of intrusion in well • Protective cover damage	• Evaluate possibility of vandalism or tampering • If necessary, reconstruct well
• Statistically significant increase in monitored parameters	• Reevaluate validity of monitoring data • Consider validity of statistical analysis • Evaluate possibility of release from monitored source • Consider possibility of impact by other sources • Resample wells to confirm statistically significant increase
• Statistically significant increase confirmed and attributed to the potential source being monitored	• Notify appropriate regulatory authorities • Develop plan of action with regulators; possible plans could require the following: • Prepare assessment monitoring plan • Submit assessment monitoring plan to regulatory agency for review (if necessary) • Implement assessment monitoring plan
• Statistically significant increase not attributable to source or found to be invalid	• Continue detection monitoring • Revise groundwater monitoring plan to prevent recurrence of problem during future monitoring events.

24.2.3 Assessment Monitoring Programs

For sites where a release to groundwater is known to have occurred, a monitoring program should be implemented to assess the nature and extent of the release. The purpose of assessment monitoring is to evaluate the nature and extent of the release and, if it is found that corrective action is needed, to identify potentially applicable corrective actions. The steps recommended in Section 24.2.2 for preparation of a detection monitoring plan can be followed when developing an assessment monitoring plan. However, if a detection monitoring program has already been implemented at a site, a significant amount of data

should already have been collected, which can be used to focus the scope of assessment monitoring. The assessment monitoring plan should contain criteria for either returning to detection monitoring or proceeding to corrective action. For example, corrective action may be warranted if the presence of contaminants is confirmed and the contaminants are present at concentrations that are greater than either background concentrations or the groundwater protection standard. Typically, the groundwater protection standard is defined as the constituent's MCL (which is defined by the U.S. EPA) or the background concentration, whichever is greater. If the groundwater is discharging to surface water, concentrations protective of aquatic life may also need to be considered. A return to detection monitoring may be warranted for any of the following reasons: (1) through additional sampling and analysis, the detection monitoring results are found to have given a false positive indication of a release from the potential source; (2) none of the parameters on the expanded list of analytes is detected at a statistically significant level above background concentrations; or (3) the concentrations of contaminants do not exceed the groundwater protection standard.

Because the purpose of the assessment monitoring program is to evaluate the nature and extent of contamination, the assessment monitoring field exploration should be focused primarily on providing data to fill gaps in the detection monitoring database or to provide data related to suspected contaminants. Also, if the results of assessment monitoring indicate that corrective action is required, a list of technically feasible corrective measures should be developed. It is useful to develop this list during assessment monitoring so that data can be collected during the assessment monitoring field exploration and subsequently used to perform a preliminary assessment of the feasibility of the candidate remedial alternatives.

24.2.4 Corrective Action Monitoring Programs

If the assessment monitoring data confirm that contaminants are present at concentrations that exceed the groundwater protection standard, then corrective action may be required. Corrective action of contaminated groundwater is described in Chapter 25 of this handbook. During corrective action, a groundwater monitoring program can be implemented to evaluate the effectiveness of the remedy. As an example, for a remedy involving extraction of groundwater to reduce contaminant concentrations in a plume of contaminated groundwater, corrective action monitoring could involve sampling and analysis of groundwater at wells (i.e., "compliance" wells) located within the plume of contamination to evaluate whether or not contaminant concentrations are decreasing at these locations. If it is found that contaminant concentrations in groundwater are not decreasing, then either the remediation system could be adjusted so that it will more likely achieve the remediation goal or it could be demonstrated that groundwater remediation at that site is technically impracticable (as described in the next chapter in Section 25.2.2). Finally, the conditions that trigger the end of corrective action (e.g., attainment of the groundwater protection standard at the compliance wells) or the need for alternative corrective action (e.g., continued exceedance of the groundwater protection standard or significant increase in the concentrations of detected contaminants) should be defined.

24.2.5 Performance Monitoring Programs

Monitoring the performance of a corrective action system is a useful approach for verifying that the features of the system are fulfilling their respective design criteria. As described in Chapter 25, groundwater remediation system features may include groundwater extraction wells, groundwater extraction trenches, or low-permeability subsurface barriers. A performance monitoring program typically involves monitoring to evaluate whether a feature of the remediation system is fulfilling its intended function. An example of performance monitoring for a groundwater extraction system is presented in Figure 24.5. If monitoring results indicate that the design criteria is not being satisfied by the feature, then the design should be reviewed and improvements should be identified that will more likely provide the desired
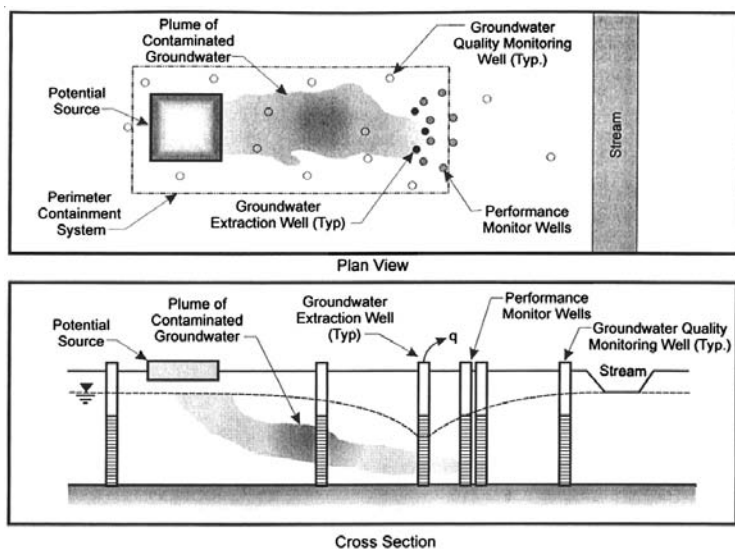
FIGURE 24.5 Performance monitoring of groundwater extraction system. (From USEPA. 1990. *Handbook — Ground Water, Volume I: Ground Water and Contamination.* EPA/625/6-90/016a. Washington, DC.)

result. Performance monitoring is described in more detail by USEPA (1994) and by Rumer and Mitchell (1995).

24.3 Design and Installation of Groundwater Monitoring Features

24.3.1 Overview of Groundwater Monitoring Features

The purpose of this section is to identify and describe installation methods for typical groundwater monitoring features. In Table 24.5, some typical groundwater monitoring features are identified. Each feature has a purpose and a situation for which it is best suited. The most common feature used to monitor groundwater quality is the monitoring well. Installation of groundwater monitoring wells is described in Section 24.3.2, and installation of other types of monitoring features is described in Section 24.3.3.

24.3.2 Groundwater Monitoring Wells

24.3.2.1 General Design of Groundwater Monitoring Wells

The purpose of a groundwater monitoring well is to provide access to the target monitoring zone for collection of a representative sample of groundwater. The representativeness of the sample may be affected by installation of the well or by the materials used to construct the well; design of the well must account for these factors. In this section, groundwater monitoring wells and their applicability are described. The discussion presented in this section should be considered to be a general guide; site-specific conditions and applicable regulatory requirements should be considered over these guidelines when designing a groundwater monitoring well.

Standard approaches for design of groundwater monitoring wells are presented by a number of agencies and organizations, including the U.S. EPA (1986a, 1992a) and the American Society for Testing Materials (ASTM, 1990). Examples of typical groundwater monitoring well designs are presented in Figure 24.6. The design shown in Figure 24.6(a) incorporates several features that minimize the possibility of introducing contaminants into the well (e.g., the protective cover, the bentonite seal, and the well apron). The monitoring well design shown in Figure 24.6(b) is typically used in situations where the well must be

TABLE 24.5 Typical Groundwater Monitoring Features

Category	Primary Use	Advantages and Limitations
Wells	Groundwater sampling	Extensive record of successful use; can be targeted to specific site limitations and contaminants of concern; flexible design.
Piezometers	Groundwater potentiometric surface measurement	Allows measurement of potentiometric surface at a point; can be outfitted to provide continuous monitoring of groundwater elevation; not well suited for groundwater sampling.
Lysimeters	Sampling liquid from vadose zone	Allows characterization of liquid that could eventually impact groundwater quality; may allow loss of VOCs in samples; construction materials (usually ceramic) may affect chemical composition of samples.
Seeps	Surface sampling of groundwater seepage	Low cost for establishing sampling point; gives good indication of quality of water with which humans or organisms may come into contact.
Geophysical techniques	Plume delineation	Large areas can be quickly delineated at the "survey" level, inexpensively; can be used on ground surface or in boreholes; cannot be used to detect presence of specific constituents.
Sampling probes (e.g., HydroPunch®, Geoprobe®)	Plume delineation	Allows quick sample collection without incurring the cost of installing a monitoring well; limited to one-time sampling of groundwater.

installed either: (1) beneath a confining feature (e.g., clay aquitard), where the additional casing prevents leakage from the upper aquifer to the lower aquifer; or (2) through a heavily contaminated soil layer, where the additional casing prevents mixing of drill cuttings from the target monitoring zone with the contaminants in the overlying or underlying portion of the aquifer. These designs can be modified as needed to meet site-specific conditions or regulatory requirements. For example, the number of monitoring points can be increased by installing multiple, discrete sampling points within a well; also, uncased, open boreholes can be used to monitor bedrock aquifers where migration of soil particles into the well is not expected to occur.

Some of the key features of the groundwater monitoring well, shown on Figure 24.6, are the wellscreen, filter pack, bentonite seal, cement grout backfill, concrete apron, and protective cover. The wellscreen should be installed so that the screen section spans only the target monitoring zone. When monitoring the quality of groundwater near the water table in an aquifer having a large seasonal fluctuation in groundwater elevation, a long wellscreen may be needed to ensure that the wellscreen continuously spans the water table. The filter pack is intended to promote formation of a graded filter outside of the well to prevent migration of fine-grained soils into the well (because soil particles are composed of minerals that may be constituents of concern, the presence of fine-grained soils in a well can cause inaccurate groundwater monitoring results, as well as clog the well). The filter pack material should also have a characteristic particle size (i.e., the diameter greater than 85%, by weight, of the soil particles) that is bigger than the wellscreen slot size to prevent clogging of the wellscreen by the filter pack material. Similarly, the filter pack material should be capable of retaining the coarsest 15% of materials in the adjacent geologic formation. The bentonite seal is intended to prevent the cement grout backfill from migrating into the filter pack; the presence of grout in the filter pack could permanently compromise the validity of groundwater samples from the well. The concrete apron is intended to route surface water away from the well and to prevent downward migration of surface water into the wellscreen. The protective cover is intended to prevent unauthorized access to the well and to protect the exposed portion of the riser pipe from damage due to incidental contact.

When installing a groundwater monitoring well, the following potential problems should be anticipated and avoided to the extent possible (after Nielsen, 1991, and USEPA, 1993c):
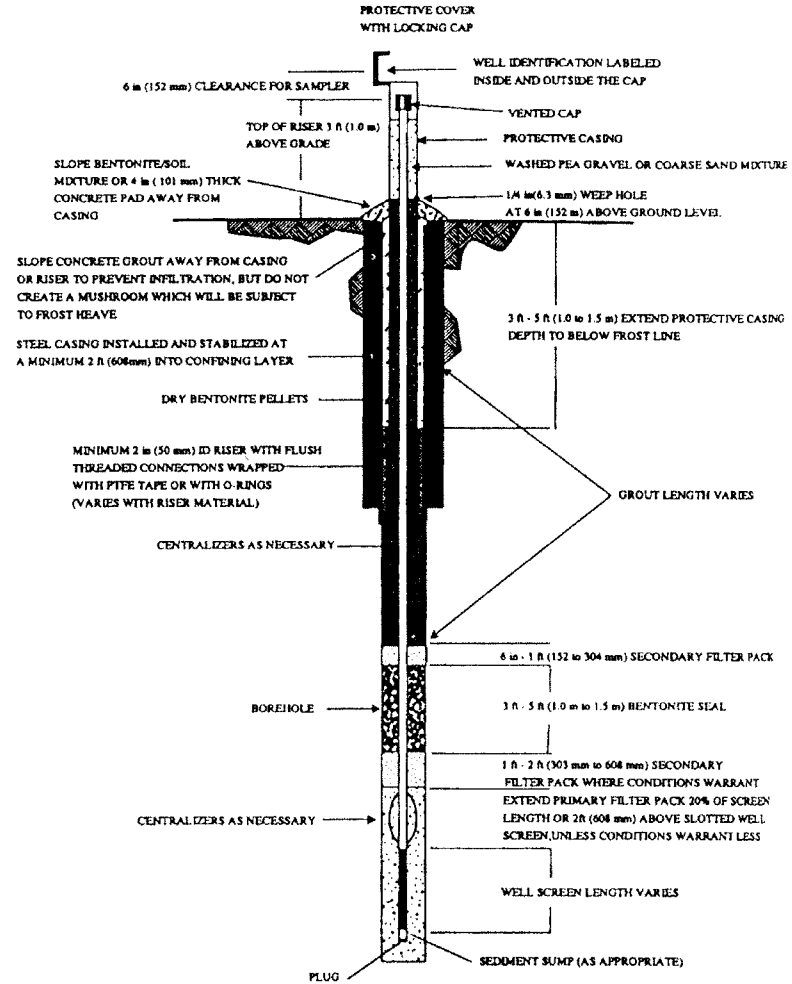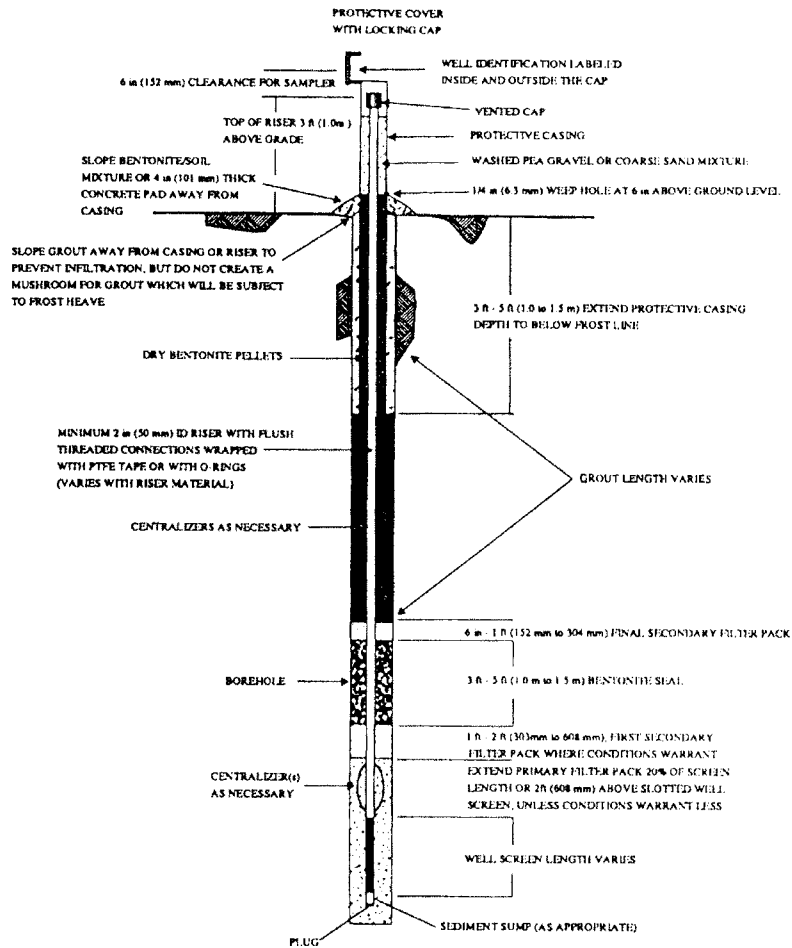
FIGURE 24.6 Monitoring well design.

- Use of well construction materials that are physically or chemically incompatible with either the surrounding natural earth materials or contaminants in the target monitoring zone and strong enough to prevent collapse under the stress applied by the soil
- Improper selection of wellscreen sizes (screen sizes that are too large may allow siltation of the well, and screen sizes that are too small may prevent proper development of a graded filter around the well)
- Placement of the screened interval of the well across stratigraphic zones, if the intent of monitoring is to sample discrete zones of the aquifer (this problem could also limit the use of the well for hydraulic conductivity testing of the aquifer)
- Improper selection or placement of filter pack material, which could cause either siltation of the well or plugging of the wellscreen
- Improper selection or placement of annular seal materials, which can allow plugging of the filter pack, cross-linking of discrete water-bearing units, or migration of grout into the filter pack
- Poor surface-protection measures, which can allow damage to well casing materials or introduction of surface water into the well at the ground surface
- Poor well development and evacuation techniques, which may alter the aquifer formation around the wellscreen, cause excessive siltation of the filter pack and groundwater samples, or compromise well yield.

24.3.2.2 Drilling Methods

There are numerous drilling methods available for installing groundwater monitoring wells, each of which has attributes that may make it suited for a particular drilling location and well type. A summary of drilling methods that are commonly used to install groundwater monitoring wells is presented in Table 24.6. Also, the situations where each method is best suited are identified. More comprehensive discussions of drilling methods are presented by Hvorslev (1949), Driscoll (1986), and USEPA (1993a).

TABLE 24.6 Drilling Methods for Various Geologic Settings

Geologic Setting	Air Rotary[1,2]	Water/Mud Rotary[1]	Cable Tool	Hollow-Stem Continuous Auger[2]	Solid-Stem Continuous Auger[3]	Jet Percussion[2]	Dual-Wall Reverse-Circulation	Driven Wells[2]
Unconsolidated or poorly consolidated materials less than 125 feet deep	●	●	●	●	●	●	●	●
Unconsolidated or poorly consolidated materials more than 125 feet deep	●	●	●				●	
Consolidated rock formations less than 500 feet deep	●	●	●				●	
Consolidated rock formations more than 500 feet deep	●	●	●				●	

[1] Includes conventional and wireline core drilling.
[2] Not recommended for the installation of groundwater monitoring wells.
[3] Above any saturated zones.

Note: Although several methods are suggested as appropriate for similar conditions, one method may be more suitable than others. This determination is based on site-specific conditions and the judgment of the geologist and the driller.

From USEPA. 1992a. *RCRA Ground-Water Monitoring: Draft Technical Guidance*, Office of Solid Waste, EPA/530-4-93-001, PB93-139-350, Washington, DC.

24.3.2.3 Monitoring Well Installation

Careful installation of groundwater monitoring wells is an important step in producing representative samples of groundwater from the target monitoring zone. In this section, a generalized approach to installation of monitoring wells is presented. Further information and guidance on design and installation of monitoring wells is presented by USEPA (1975, 1989d) and ASTM (1990).

Prepare for Field Work. Before beginning monitoring well installation activities in the field, the groundwater field professional should develop as complete an understanding as possible of the conditions that will be encountered. This can be accomplished by reviewing the groundwater monitoring plan, regional and/or site-specific hydrogeologic information, the site health and safety plan (which must be prepared if hazardous conditions might be encountered), and the logs for installation of existing wells or other borings near the proposed well locations. The equipment that will be needed for the field work (e.g., water-level measuring instrument, personal protective equipment, sample jars, field analytical equipment, etc.) should be assembled well in advance of the field work. Monitoring equipment (e.g., water-level measuring equipment, field analytical equipment, etc.) should be calibrated. Well construction materials should be specified in advance of field work; recommendations for selection of well construction materials are provided in Table 24.7. Finally, the groundwater field professional should verify that the well construction materials that are procured by others meet the specific requirements of the groundwater monitoring plan.

TABLE 24.7 Recommendations for Selection of Well Casing Materials

If	Use	Do Not Use
Well depth in range of 225–375 ft (68.6–114 m)	PVC, ABS, SS	PTFE
Well depth in range of 1200–2000 ft (366–610 m)	SS	PVC or ABS
pH < 7.0, DO > 2 ppm, H_2S ≥ 1 ppm, TDS > 1000 ppm, CO_2 > 50 ppm, or Cl^- > 500 ppm	PVC, ABS, or PTFE	SS
A neat PVC solvent/softening agent is present or if the aqueous concentration of the PVC solvent/softening agent exceeds 0.25 times its solubility in water	SS or PTFE	PVC

<div align="center">IN ALL CASES</div>

Do Not Use	Use
Solvent bonded joints for PVC casings	Threaded PVC casings
Welded stainless joints	Threaded SS casings
Any PVC well casing that is not NSF-ASTM approved — D-1785 and F-480	ASTM-NSF approved PVC well casings — D-1785 and F-480
Any stainless steel casing that is not ASTM approved — A312	ASTM approved SS 304 and SS 316 casings — A312
Any ABS well casing that is not ASTM approved	ASTM approved ABS casings — F-480
DO = dissolved oxygen; TDS = total dissolved solids; PVC = polyvinyl chloride; SS = stainless steel; ABS = acrilonitrile-butydiene-styrene; PTFE = polytetrafluoroethylene	

From USEPA. 1992a. *RCRA Ground-Water Monitoring: Draft Technical Guidance*, Office of Solid Waste, EPA/530-4-93-001, PB93-139-350, Washington, DC

Lay Out Location for Monitoring Well. The well location specified in the monitoring plan should be laid out before drilling begins. After the location is marked, the groundwater field professional should verify that the location is suitable for the well (e.g., not in a location where it is likely to be damaged, or in standing water, or where conditions are encountered that were not anticipated in the groundwater monitoring plan). Note that, after the well has been installed, the location should be surveyed by a licensed professional surveyor to provide an exact record of the location of the well.

Decontaminate Drilling Equipment and Well Construction Materials. Equipment or construction materials should be decontaminated before being introduced into the borehole using a high-pressure steam wash or other appropriate decontamination method (see Moberly [1985] for a discussion of decontamination procedures). Decontamination minimizes the possibility of unintentionally introducing contam-

inants into the borehole. Drilling fluids should be inert; the use of oil and grease on drilling equipment should be restricted.

Drill Borehole. The borehole should be drilled using a technique that minimizes disturbance to the aquifer formation and does not impact the ability of the well to produce a representative groundwater sample. The potential impacts of drilling on groundwater samples are illustrated in Table 24.8, and candidate drilling techniques are identified in Table 24.6. During drilling of the borehole, the cuttings should be examined by the driller and the groundwater field professional, and a stratigraphic log should be made of the materials that are encountered; a sample boring log form is shown in Figure 24.7. The drill cuttings, drill fluid, or cores should be examined for unexpected materials (e.g., soils that are not described on logs for nearby wells, contaminants, indications of stratigraphic units at unexpected locations, etc.). The groundwater field professional should verify that the borehole is open throughout its depth, is reasonably straight, and is terminated within the target monitoring zone so that the well can be properly constructed; these can be verified by measuring the total depth of the boring after drilling equipment has been removed from the boring. If downhole geophysical logging is required by the groundwater monitoring plan, then it may be performed at this time. (Note that for some geophysical logging techniques, logging must be performed while the borehole is open, prior to well construction.)

Inspect Well Construction Materials. Well construction materials should be inspected to verify that they meet the requirements of the groundwater monitoring plan. Wellscreen and riser pipes should be clean and undamaged; screen slots should be uniform, open, and continuous throughout the screened interval; threaded connections at ends of pipes should be properly formed and should provide a tight connection. Bentonite seal materials should be delivered to the site in sealed containers and should not be allowed to hydrate before installation in the borehole. Filter pack material should be tested for particle-size distribution to verify that it meets the gradation requirements of the groundwater monitoring plan. Before installation, the length of each section of the riser pipe should be measured to the nearest hundredth of a foot or to the nearest centimeter so that, once installed, the depth to the top and bottom of the wellscreen is known.

Install Wellscreen and Riser Pipe. The riser pipes should be installed carefully to prevent damage during installation. The completed riser pipe should extend about three feet (1 m) above the ground surface, as shown in Figure 24.6. Alternatively, if the presence of an above-ground riser causes an unacceptable obstacle at the site, the riser pipe can be terminated at ground level in a vault that is specially designed to prevent leakage into the well or damage to the well. The wellscreen should be in the borehole, using a centralizer if necessary; if the wellscreen is not centered in the borehole, it might not be possible to install the filter pack material evenly around the pipe, which might prevent proper development of a graded filter around the wellscreen.

Install Filter Pack, Seal, and Backfill Material. The filter pack material should be installed around the wellscreen slowly, to prevent bridging of the material between the wellscreen and the soils in the adjacent aquifer. Bridging, which leaves a portion of the wellscreen unsupported, could result in crushing of the riser pipe or migration of fine-grained soil particles into the well. For deep boreholes, the filter pack material may need to be installed using a tremie pipe, which is used to deposit the material from the bottom up (see Driscoll [1986, p. 477] for additional discussion). The bentonite seal should be installed evenly, above and on all sides of the top of the filter pack using either dry bentonite pellets or bentonite slurry. If dry bentonite is used, then it should be allowed to hydrate in the well for a period of at least one hour (preferably overnight) before the grout backfill material is installed over the seal. The backfill material should be installed carefully to prevent damage to the bentonite seal; backfill materials may include cement grout, drill cuttings (if uncontaminated), or stabilized soil, depending on the requirements of applicable regulations. The depth to the top of the filter pack, seal, and backfill material should be measured and recorded before the overlying component of the well is installed; when making these measurements, the groundwater field professional should confirm that the wellscreen section is located completely within the filter pack. Samples of well construction materials should be retained by the contractor for a period of time designated in the well construction specifications.

TABLE 24.8 Effects of Drilling Methods on Water Samples

Methods	Advantages	Disadvantages
Air rotary	• Drilling fluid is not always used, minimizing contamination and dilution problems	• When more than one water-bearing zone is encountered and hydrostatic pressures are different, flow between zones occurs after drilling is completed but before the hole is cased and grouted • Oil from compressor may be introduced to geologic formation • Use of foam additives containing organic materials can interfere with both organic and inorganic analyses
Mud rotary	• Borehole can be kept open inexpensively • Sample collection is relatively easy	• Drilling fluid that mixes with formation is difficult to remove and can cause contamination • Fluid circulation can cause vertical mixing of contaminants • Drilling fluids and additives can interfere with subsequent water quality analyses • Lubricants may cause contamination
Bucket auger	• No drilling fluid is used, minimizing contamination and dilution problems	• Large-diameter hole makes it difficult to assure adequate grouting • Must continuously add water in soft or loose formations
Solid stem auger	• No drilling fluid is used, minimizing contamination and dilution problems • Can avoid use of lubricants	• Because auger must be removed before well can be set, vertical mixing can occur between water-bearing zones • Can cause vertical mixing of both formation water and geologic materials
Hollow stem auger	• No drilling fluid is used, minimizing contamination and dilution problems • Can avoid use of lubricants • Formation waters can be sampled during drilling • Well can be installed as augers are removed, decreasing interaction with water from higher water-bearing zones	• Can cause vertical mixing of geologic materials • Can cause vertical mixing of formation waters if augers are removed before well is installed
Cable tool	• Little or no drilling fluid required	• Contamination of aquifer is possible if drilling fluid is used • Slight potential for vertical mixing as casing is driven
Jetting	• May be only alternative where drill rig access is poor	• Large quantities of water or drilling fluid are introduced into and above sampled formation • Cannot isolate target monitoring zone with a grout seal

From Driscoll, F.G. 1986. *Groundwater and Wells*. Johnson Division, St. Paul, MN. With permission.

Complete Above-Ground Portion of Well. The above-ground portions of the well should be constructed in a manner that accommodates sampling of the well and protects the well from incidental damage. The actual configuration of the aboveground portion of the well may depend on applicable regulations; a recommended configuration is shown in Figure 24.6. For above-ground risers, the riser should be located at short depth (i.e., about 3 to 6 inches, 75 to 150 mm) beneath the top of the protective casing, and the reference point (for measuring the depth to groundwater and total depth of the well) should be clearly and permanently marked on the riser pipe. The annular space between the protective casing and the riser pipe can be filled with sand or gravel (so that items dropped in the casing can be easily retrieved) and a hole should be drilled in the casing about 6 inches (150 mm) above the ground surface to allow drainage of water that may collect in the casing during sampling. A cement grout pad should be constructed around the protective casing at the ground surface and sloped downward, away from the protective casing,

FIGURE 24.7 Typical boring/well installation form. (From Nielsen, D.M., ed. 1991. *Practical Handbook of Ground-Water Monitoring.* Lewis Publishers, Chelsea, MI. With permission.)

to prevent accumulation of standing water around the well and infiltration of surface water into the well. The volume of grout placed above ground should be minimized to prevent frost heave of the grout pad during freezing weather.

24.3.2.4 Well Development Techniques

The purposes of well development are to remove sediment from the well and surrounding aquifer materials impacted by well installation, and to form a graded filter at the interface between the aquifer materials and the filter pack material. Well development involves removing water from the well after construction is completed. Development prior to initial well sampling is essential if a representative sample is to be obtained. Common well development techniques are described by Driscoll (1986). Hand-bailing techniques are widely used but are time consuming and may alter the geologic formation that surrounds the filter pack, resulting in high values of suspended solids in groundwater samples or even failure of the filter pack. Mechanical development techniques (e.g., swabbing, surging, or low-flow pumping) require less physical effort and can be tailored to the specific limitations of the aquifer and well construction features. Wells should be developed until the pH, specific conductivity, and color (i.e., cloudiness) of the water stabilize; note that, if pH and specific conductivity remain consistently higher in the well than in the aquifer, there may be leakage into the filter pack from the overlying cement grout (grout typically has a pH of about 10 and a specific conductivity of about 1500 (μmhos/cm). Cloudy development water may be an indication that the filter pack is not filtering the aquifer formation material effectively. Such problems indicate that groundwater samples from the well may not be representative of groundwater in the aquifer near the well. Problems encountered during development should be examined carefully and, if necessary, the well should be redeveloped or decommissioned (using techniques described in Section 24.3.2.6) and then reinstalled. Also, a water level inconsistent with surrounding water levels can indicate a problem with well completion.

24.3.2.5 QC/QA of Drilling and Well Installation

To verify that the monitoring well is installed properly, the groundwater professional should implement a program of Quality Control and Quality Assurance (QC and QA) during drilling and installation. In the context of monitoring well installation, QC refers to the procedures used to confirm that the materials and methods used meet the requirements of the groundwater monitoring well design, and QA refers to the activities that are performed to verify that QC procedures were properly implemented. QC and QA procedures are typically specified in the groundwater monitoring plan. QC activities could include testing well construction materials, verifying that appropriate drilling methods are used, observing the performance of the drilling equipment, and documenting drilling and well construction details. QA is typically provided by the groundwater field professional (i.e., either a geologist, an engineer, or an engineering technician) and includes oversight of the drill crew's activities, inspection of the materials of construction, and review of submittals from the driller (e.g., material specification sheets, results of filter pack laboratory gradation tests, etc.). Cooperation between the driller and the groundwater field professional is essential for quick resolution of problems that may arise during well installation.

An essential element of QC/QA is documentation of field activities. Well installation activities should be recorded in a boring log; a sample boring log form is provided in Figure 24.7. The details of well development should also be documented (e.g., pH and specific conductivity measurements for each purged well volume, cloudiness of purge water, time required to purge the well, etc.).

24.3.2.6 Decommissioning of Groundwater Monitoring Wells

In many jurisdictions, groundwater monitoring wells that are not actively used must be decommissioned (also, the terms "destroyed" or "abandoned" are used in some jurisdictions to denote decommissioning). Decommissioning involves removing the well from use so that the well does not act as either a possible conduit for contaminant migration or uncontrolled access to groundwater resources. Decommissioning can be performed either by: (1) installing cement grout within the wellscreen and filter pack and then removing the casing and riser pipe (if possible); or (2) overcoring the well and removing all of the well

construction materials from the ground, then backfilling the borehole with appropriate material (e.g., a bentonite seal in low-permeability zones, soil in unconsolidated soil zones, or cement grout in rock formations).

24.3.3 Piezometers

A piezometer is used to measure the hydraulic head at a specific location within the aquifer. Because piezometers are not designed for collection of groundwater samples, their design may be different from the design of a groundwater monitoring well. Attributes typical of a piezometer include a small diameter borehole, casing, and screen (to minimize the impact of the piezometer on groundwater flow in the aquifer), small screened interval (because piezometers having long screened intervals provide the average head over a large depth, instead of the head at a specific location), small or no filter pack (because the filter pack may impact the hydraulic head near the screened interval of the piezometer), and a very effective seal at the top of the screened interval (to prevent intrusion of water from above the measuring point).

24.3.4 Innovative Monitoring Features

Other groundwater monitoring features identified in Table 24.5 may be used to simplify, accelerate, or economize the collection of groundwater data. These alternative approaches to groundwater monitoring are being used for increasingly quick and cost-effective collection of data. The drawback to using these monitoring features is that, because there is no permanent, fixed feature, there is no basis for collecting samples over a period of time and comparing the results. Some of the more well-developed and commonly used innovative approaches are described below. Additional innovative monitoring features are described by USEPA (1993a).

The Geoprobe® can be used to collect a one-time sample of groundwater, soil, or soil gas. The system includes a soil probing machine and a sampler. First, the soil probing machine is mounted on a vehicle and hydraulically advanced into the ground to the target monitoring zone using static force and percussion; then the sampling device is advanced inside the Geoprobe probe rod beneath the tip of the probe, where the sample is collected and then removed from the borehole. This approach provides a discrete sample without the cost of installing, developing, or maintaining a groundwater monitoring well.

HydroPunch® is used to sample groundwater and involves driving a sampling point into the ground for a one-time sampling of groundwater. The sampler is attached to a cone penetrometer drill rod and, after the probe has been advanced to the target monitoring zone, the outer cylinder of the probe is removed, exposing a perforated stainless steel sample entry barrel. The sample is forced into the Hydro-Punch sampler by hydrostatic pressure and then the sampler is removed to the ground surface. The advantages of this method are low cost, quick collection of samples, and applicability to a wide range of soil deposits. A discussion of the HydroPunch sampler is provided by Edge and Cordry (1989).

Cone penetrometer samplers include a variety of devices that are attached to the end of a cone penetrometer and used for sampling groundwater or soil gas. The devices can also be used to measure hydraulic conductivity and pore-water pressure.

24.4 Groundwater Sampling

24.4.1 General Approach to Sampling Groundwater Monitoring Wells

A general approach to sampling of groundwater monitoring wells is presented in a stepwise fashion in this section. The following procedures are a compilation of groundwater sampling techniques recommended by the U.S. EPA (1986, 1992a, 1993a), Nielsen (1991), Driscoll (1986), and ASTM (1992). In contaminated areas, sampling should proceed from the least contaminated well to the most contaminated well if any sampling equipment is to be used at more than one well.

Step 1: Inspect Wells and Area Surrounding Wells

The well should be examined for evidence of tampering, damage, or other activity that could compromise the ability of the well to produce a representative groundwater sample. The ground surface near the well should be examined for signs of recent activity (including accidental spills of constituents that could impact groundwater quality) and the area being monitored should be examined for changes since the last sampling event (e.g., new construction, new storage activities for waste materials, agricultural activity, etc.); such changes may require reexamination of the adequacy of the monitoring well network layout. Observations should be recorded by the field sampling professional for future reference.

Step 2: Measure Depth to Groundwater and Total Depth of Well

The first activity to be performed after the well is inspected should be measurement of the depth to groundwater and the total depth of the well, using the techniques described in Section 24.4.2. All measurement equipment should be decontaminated (as described by USEPA [1992a]) before introducing it into the well. Measurements of the depth to groundwater that are made shortly after sampling the well may not be representative of the actual depth to groundwater. Measurement of the total depth of a well is important for evaluating changes in the total depth; for example, a decrease in the total depth could be indicative of siltation in the bottom of the well, or a crushed wellscreen or riser pipe.

Step 3: Evacuate Groundwater from the Well

The groundwater in both the well and the filter pack around the wellscreen should be evacuated before sampling begins using the techniques described in Section 24.4.4 or 24.4.5. The purpose of evacuating the well is to cause groundwater from the aquifer formation around the well to flow into the well, allowing collection of a representative groundwater sample from the target monitoring zone. If the well is a groundwater supply or an irrigation well, then evacuation may not be necessary. Evacuated water may need to be collected and treated if there is a chance that it could be contaminated.

Step 4: Collect Groundwater Samples

Groundwater in the well should be sampled using the techniques described in Section 24.4.4, 24.4.5, or 24.4.6. Extreme care should be used when collecting, preserving, storing, and shipping groundwater samples (as described in Section 24.4.7) to provide samples to the laboratory that are representative of groundwater quality. Sample quality is much more likely to be compromised in the field than in the controlled environment of a laboratory.

Step 5: Document the Sampling Event

During the sample event, each step of the sampling process should be documented to allow detailed examination of the sampling procedures in the future. Recommended documentation procedures are described in Section 24.4.8. Good documentation is essential because it provides the basis for verifying the validity of the monitoring event. Also, if monitoring problems cannot be traced to laboratory error or resolved using field logs, then the notes of the field sampling professional may be valuable in identifying the source of the problem.

24.4.2 Evaluations of Groundwater Elevation Measurements or Depths to Groundwater

The elevation of the groundwater potentiometric surface is typically evaluated by measuring the depth to groundwater at several locations, calculating the elevation of groundwater at each location, and then interpolating contours of groundwater elevation for the area. There are many techniques available for measuring the depth to groundwater in a well, as illustrated on Table 24.9, including hand measurement using a wetted tape (which is simple and inexpensive, but somewhat time consuming) and electronic measurement (which is more expensive but less time consuming, and can be configured to provide a continuous record of data, if needed). The groundwater elevation can be calculated for the well location

TABLE 24.9 Summary of Methods for Manual Measurement of Well Water Levels in Nonartesian and Artesian Wells

Measurement Method	Measurement Accuracy (ft)	Major Interference or Disadvantage
	Nonartesian Wells	
Wetted-tape	0.01	Cascading water or casing wall water
Air-line	0.25	Air line or fitting leaks; gage inaccuracies
Electrical	0.02 to 0.1	Cable wear; presence of hydrocarbons on water surface
Transducer	0.01 to 0.1	Temperature changes; electronic drift; blocked capillary
Acoustic probe	0.02	Cascading water; hydrocarbon on well water surface
Ultrasonics	0.02 to 0.1	Temperature changes; well pipes and pumps; casing joints
	Artesian Wells	
Casing extensions	0.1	Limited range; awkward to implement
Manometer/pressure gage	0.1 to 0.5	Gage inaccuracies; calibration required
Transducers	0.02	Temperature changes; electronic drift

From Nielsen, D.M., ed. 1991. *Practical Handbook of Ground-Water Monitoring*. Lewis Publishers, Chelsea, MI. With permission.

by subtracting the depth to groundwater from the surveyed elevation of the reference measuring point on the monitoring well casing.

Measurements of the depth to groundwater can be complicated by the presence of contaminants in the monitoring well or piezometer, as illustrated in Figure 24.8 and as discussed by Blake and Hall (1984). As shown in the figure, the thickness of LNAPLs (e.g., petroleum hydrocarbon products) in an aquifer can be significantly different than the thickness of LNAPL in a well. LNAPLs typically "float" on groundwater at the top of the capillary fringe. LNAPLs flow into the well until the elevation of the top of LNAPL above the capillary fringe equals the elevation of the top of LNAPL inside the well. Because groundwater in the capillary fringe does not flow into the well, the thickness of LNAPL present in the well (T_A) is typically greater than the actual thickness of LNAPL in the aquifer ($H_M + T_M - H_C$). The actual thickness of LNAPL, and a more accurate estimate of depth to groundwater, can be estimated from:

$$T_M = T_A - \left(D_{WT} + H_M\right) \tag{1}$$

where T_M = thickness of mobile LNAPL (ft), T_A = apparent thickness of LNAPL (ft), H_M = distance from bottom of the mobile LNAPL layer to the water table, and D_{WT} = depth of LNAPL in the well below the actual water table (ft). D_{WT} can be calculated as:

$$D_{WT} = H_M \gamma_{LNAPL} / \left(\gamma_w - \gamma_{LNAPL}\right) \tag{2}$$

where γ_{LNAPL} = the specific weight of LNAPL (lb/ft^3), and γ_w = the specific weight of water (lb/ft^3). H_m is approximately equal to the thickness of the capillary fringe, which can be estimated based on the grain-size distribution of the soil (see Lambe and Whitman, 1969).

Once the elevation of groundwater has been calculated for the wells at a site, the configuration of the potentiometric surface can be estimated to provide an indication of the likely direction of groundwater flow and the gradient of the potentiometric surface. An illustration of a potentiometric surface map, as interpreted from groundwater elevations at several wells, is provided in Figure 24.9. As shown in the figure, a potentiometric surface can be produced by interpolating data between monitoring wells; the groundwater professional should carefully consider all available site data when evaluating groundwater flow direction to prevent the type of problem illustrated in Figure 24.9, which shows that a limited subset of the data would imply an incorrect groundwater flow direction. Surface-water features that are connected with groundwater can also provide valuable information regarding the configuration of the potentiometric surface. The gradient of the potentiometric surface is calculated as the ratio of vertical
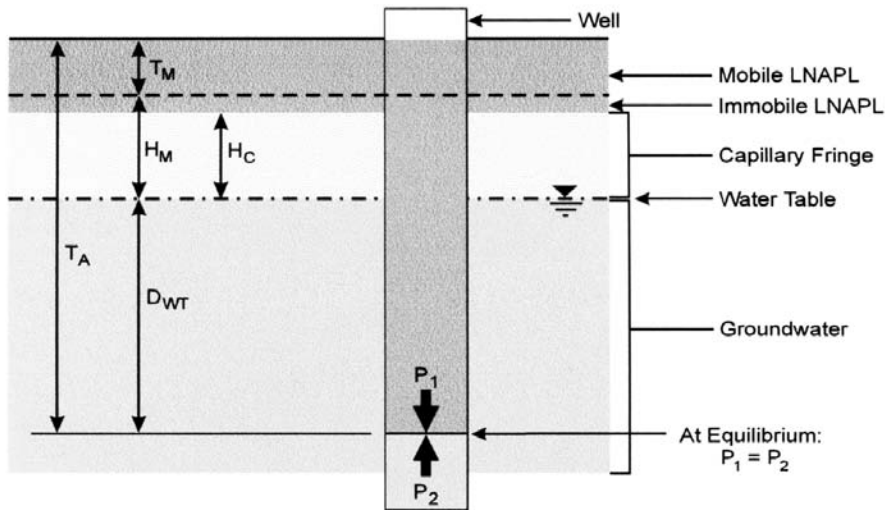
FIGURE 24.8 Distribution of LNAPL and groundwater in a well within an LNAPL-contaminated aquifer.

change of elevation in the potentiometric surface to the distance over which the change occurred. The velocity of groundwater can then be calculated for a given area as (Freeze and Cherry, 1979):

$$v = -\frac{Ki}{n_e} \tag{3}$$

where v = groundwater flow velocity [ft/sec], K = hydraulic conductivity [ft/sec], i = groundwater hydraulic gradient (which is usually a negative number) [–], and n_e = effective porosity of the aquifer material [–]. This velocity can be used as an initial estimate of the speed at which a nonreactive solute might move through the aquifer.

24.4.3 Overview of Sampling Techniques

Sampling involves collection of groundwater for analysis. There are many available sampling techniques, as summarized in Table 24.10. Different techniques may be applicable for different types of wells, different analytes, and different types of hydrogeologic settings, as shown on Table 24.11 and as discussed in USEPA (1993a) and Nielsen (1991). Sampling techniques can be generally categorized as either manual techniques (which are described in Section 24.4.4) or mechanical techniques (which are described in Section 24.4.5).

24.4.4 Manual Sampling Techniques

Manual sampling techniques (i.e., collection of samples using hand-operated equipment) are commonly used for sampling groundwater and are shown under "grab" samples in Table 24.10. Manual sampling devices include open bailers, point-source bailers, and syringe samplers, among others. To sample a well manually, the groundwater field professional must lower the sampling device down the well to the groundwater level, activate the sampling mechanism (if the sampler requires activation), and then remove the sampler from the well. The advantages of manual sampling include simplicity of operation, applicability to a wide range of monitoring situations, relatively low cost, and reliability. The disadvantages of manual sampling include the possibility of operator error, the time required to collect a sample (especially
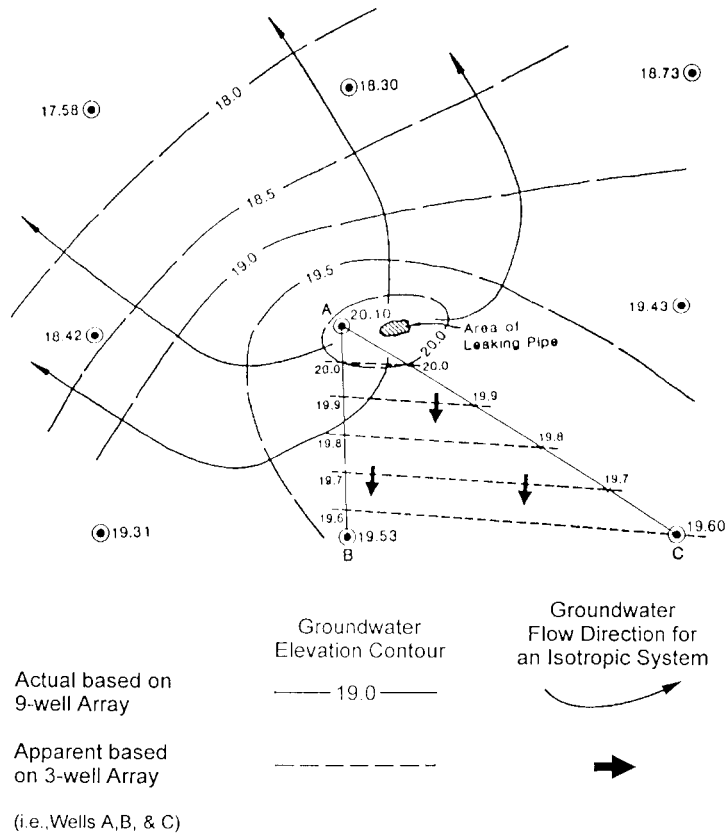
FIGURE 24.9 Potentiometric surface mapping. (From Nielsen, D.M., ed. 1991. *Practical Handbook of Ground-Water Monitoring*. Lewis Publishers, Chelsea, MI. With permission.)

from deep wells), difficulty in obtaining large-volume samples, and potential damage to the filter pack due to agitation caused by introducing the sampling device into the well.

24.4.5 Mechanical Sampling Techniques

Mechanical sampling techniques (i.e., any nonmanual technique) offer several advantages over manual sampling techniques. A significant advantage is that mechanical sampling devices can be configured to prevent some of the impacts to sample integrity that may be caused by manual sampling devices. Also, the time required to collect samples can be greatly reduced using mechanical sampling devices, particularly for deep wells, and mechanical sampling is often the only feasible alternative for deep wells. Mechanical sampling techniques typically involve the use of pumps to lift groundwater to the ground surface, where the sample is collected. Guidelines that apply to each mechanical sampling technique include: (1) sampling at a flowrate at which the aquifer can sustainably produce water (to prevent migration of fines from the aquifer to the well); (2) operating pumps in a manner that does not cause aeration of samples; (3) avoiding contact between sampling equipment and LNAPL or DNAPL product in wells (using, for example, a stilling tube to sample groundwater beneath LNAPL); and (4) using pumps and sampling equipment that are made of a material that will minimize impacts to concentrations of contaminants of concern in the sample.

TABLE 24.10 Generalized Groundwater Sampling Device Matrix

Device	Approx. Maximum Sample Depth	Minimum Well Diameter	Sample Delivery Rate or Volume†	INORGANIC						ORGANIC					RADIOACTIVE		BIO-LOGICAL
				EC	pH	Redox	Major Ions	Trace Metals	Nitrate Fluoride	Dissolved Gases	Non-Volatile	Volatile	TOC	TOX	Radium	Gross Alpha & Beta	Coliform Bacteria
PORTABLE SAMPLING DEVICES																	
GRAB																	
Open bailer	No limit	1/2 in	Variable	•	□	□	•	•	•	□	•	□	□	□	•	□	•
Point-Source Bailer	No limit	1/2 in	Variable	•	•	•	•	•	•	□	•	□	•	•	•	□	•
Syringe	No limit	1–1/2 in	0.01–0.2 gal	•	•	•	•	•	•	□	•	•	□	□	•	•	•
SUBMERSIBLE																	
Gear-drive pump	200 ft	2 in	0–0.5 gpm	•	•	•	•	•	•	•	•	•	•	•	•	•	□
Bladder pump	400 ft	1–1/2 in	0–2 gpm	•	•	•	•	•	•	•	•	•	•	•	•	•	•
Helical rotor pump	160 ft	2 in	0–1.2 gpm	•	•	•	•	•	•	•	•	•	•	•	•	•	□
Gas-driven piston pump	500 ft	1–1/2 in	0–0.5 gpm	•	□	□	•	•	•	□	•	•	□	□	•	•	□
Centrifugal (low-rate) pump	Variable	2 in	Variable	•	•	•	•	•	•	•	•	•	•	•	•	•	•
SUCTION																	
Peristaltic pump	26 ft	1/2 in	0.01–0.3 gpm	•	□	□	•	•	•	□	•	□	□	□	•	□	•

Device	Depth	Diameter	Delivery rate†													
GAS CONTACT																
Gas-lift pump	Variable	1 in	Variable	□	□	□	□	□	□	□	□	□	□	□	□	□	□
Gas-drive pump	150 ft	1 in	0.2 gpm	•	□	□	•	□	•	□	•	□	□	□	•	□	□
IN SITU* SAMPLING DEVICES																
Pneumatic pump	No limit	No limit	0.01–0.13 gpm	•	•	•	•	•	•	□	•	□	□	□	•	•	•

* Sampling devices on this chart are divided into two categories: (1) portable devices for sampling existing monitoring wells, and (2) *in situ* monitoring devices (often multilevel) that are permanently installed. Sampling device construction materials (including tubing, haul lines, etc.) should be evaluated for suitability in analyzing specific groundwater parameters. It is assumed on this chart that existing monitoring wells are properly installed and constructed of materials suitable for detection of the parameters of interest.

† Sample delivery rates and volumes are average ranges based on typical field conditions. Actual delivery rates are a function of diameter of monitoring well, size and capacity of sampling device, hydrogeologic conditions, and depth to sampling point. For all devices, delivery rate should be carefully controlled to prevent aeration and degassing of the sample.

• Indicates device is generally suitable for application (assuming device is cleaned and operated properly and is constructed of suitable materials).

□ Indicates device may be unsuitable or is untested for application.

From USEPA. 1992a. *RCRA Ground-Water Monitoring: Draft Technical Guidance*, Office of Solid Waste, EPA/530-4-93-001, PB93-139-350, Washington, DC.

TABLE 24.11 Sample Preservation Measures

Parameters (Type)	Volume Required (mL) 1 Sample**	Container (Material)	Preservation and Storage Requirements	Maximum Holding Period
Well purging				
pH (grab)	50	T,SS,P,G	None; field det.	<1 hr.***
specific conductivity (grab)	100	T,SS,P,G	None; field det.	<1 hr.***
T (grab)	1,000	T,SS,P,G	None; field det.	None
Eh (grab)	1,000	T,SS,P,G	None; field det.	None
Contamination indicators				
pH, specific conductivity, (grab)	As above	As above	As above	As above
TOC	40	G,T	Dark, 4°C	24 hrs
TOX	500	G,T	Dark, 4°C	5 days
Water quality				
dissolved gases (O_2, CH_4, CO_2)	10 mL minimum	G,S	Dark, 4°C	<24 hrs
Alkalinity/Acidity	100 (filtered)	T,G,P	4°C/None	<6 hrs***/<24 hrs
Fe, Mn, Na+, K+, Ca++, Mg++	All filtered, 1,000 ml	T,P	Field acidified to pH<2 with HNO_3	6 months△
PO_4, Cl, Silicate	@50	(T,P,G glass only)	4°C	24 hrs/7 days; 7 days
NO_3^-	100	T,P,G	4°C	24 hrs
SO_4^-	50	T,P,G	4°C	7 days
NH_4^+	400	T,P,G	4°C/H_2SO_4 to pH <2	24 hrs/7 days
Phenols	500	T,G	4°C/H_3PO_4 to pH <4	24 hrs
Drinking water suitability				
As, Ba, Cd, Cr, Pb, Hg, Se, Ag	Same as above for water quality cations (Fe, Mn, etc.)	Same as above for water quality cations (Fe, Mn, etc.)	Same as above for water quality cations (Fe, Mn, etc.)	6 months
F^-	Same as chloride above	Same as chloride above	Same as chloride above	7 days
Remaining organic parameters	As for TOX/TOC, except where analytical method calls for acidification of sample	24 hrs		

* Modified after Scalf, M.R., McNabb, Dunlap, W.J., Crosby, R.L., and Fryberger, J. 1981. *Manual of Ground-Water Quality Sampling Procedures.* EPA/600/2-81/160 (NTISPB82-103045).

** It is assumed that at each site, for each sampling date, replicates, a field blank, and standards must be taken at equal volume to those of the samples.

*** Temperature correction must be made for reliable reporting. Variations greater than ±10% may result from longer holding period.

△ In the event that HNO_3 cannot be used because of shipping restrictions, the sample should be refrigerated to 4°C, shipped immediately, and acidified on receipt at the laboratory. Container should be rinsed with 1:1 HNO_3 and included with sample.

Note: T = Teflon, SS = stainless steel, P = PVC, polypropylene, polyethylene. G = borosilicate glass

From USEPA. 1993a. Office of Research and Development. *Subsurface Characterization and Monitoring Techniques: A Desk Reference Guide, Volume I: Solids and Ground Water, Appendices A and B.* U.S. Environmental Protection Agency, Washington, DC. EPA/625/R-93-003a.

24.4.6 Innovative Sampling Techniques

Several groundwater sampling techniques have recently been developed that allow sampling in a more cost-effective or timely manner than the traditional techniques described in Sections 24.4.4 and 24.4.5. Some of these techniques are described briefly below.

Geoprobe®, *HydroPunch®*, and *Cone Penetrometer* are innovative sampling techniques that, as described in Section 24.3.4, can be used to collect samples of groundwater. These techniques provide a one-time sample and are typically used during assessments of the nature and extent of groundwater contamination. Their advantages include quick collection of samples and avoidance of the cost of installing a monitoring well. Their disadvantages include the limitation of one-time sample collection and difficulty in sampling at great depths.

Ion-selective probes can be used to detect specific ions in groundwater. The probes are designed to detect specific ions by producing an electric signal that can be compared to a reference signal for a specific ionic constituent. This method, which is in the developmental stages, may be particularly useful for preliminary groundwater characterization.

Fiber optic chemical sensors are also designed to detect a specific chemical in groundwater. The sensor is made of a reagent that is physically confined or chemically immobilized at the end of a fiber optic cable. The cable is inserted into a monitoring well, piezometer, or borehole, and then a signal is transmitted through the cable, which detects the constituent, if it is present. This method, which is also in the developmental stages, can be used on an extremely wide variety of constituents, uses portable equipment, and produces results at very low cost.

24.4.7 Sample Handling and Preservation

After the well has been sampled, the sample must be properly transferred to the sample container and preserved for transport to the laboratory. During sampling, every effort must be made to minimize changes in the chemistry of the sample. To minimize such changes in chemistry, samples should be collected, preserved, and stored correctly. Guidelines for each of these activities are presented by USEPA (1992a) and Nielsen (1991) and are summarized below. A generalized groundwater sampling protocol is presented in Table 24.12.

Sample collection refers to the transfer of the sample from the sampling device to the sample container. The contents of the sampling device should be transferred in a controlled manner that minimizes sample agitation and aeration, which can cause gasification of samples, allow release of volatile organics, or cause oxidation and precipitation of metals in the sample. Groundwater samples should be collected as soon after the well is purged as possible to prevent interaction of the samples with the atmosphere and well casing. In samples that will be tested for volatile organic constituents, there should be no air in the sample containers (which would allow volatilization of the compounds during shipment). Also, the samples should be transferred from the sampling device directly into the sampling container to minimize the opportunity for contamination of samples or changes in sample chemistry.

Filtration of samples is frequently performed to provide samples that do not contain suspended solids. Filtration is usually performed by draining or forcing samples through a filter having openings of 0.45 microns. The advantage of filtering samples is that the constituents adsorbed to the suspended solids can be distinguished from constituents that travel in dissolved phase in the groundwater. Potential disadvantages of filtering include changes in groundwater chemistry during filtration and the removal of some solid particles (e.g., colloids) that may actually travel with groundwater. More detailed discussions of the advantages and disadvantages of filtering groundwater are provided by Puls and Barcelona (1989) and USEPA (1995). In monitoring scenarios where significant quantities of suspended solids exist, it may be useful to analyze both filtered and unfiltered samples to allow a comprehensive evaluation of groundwater quality.

Sample preservation refers to actions taken to minimize changes to the chemistry of the sample after it is removed from the well. Because there are significant changes to the environment of the sample after it is removed from the well (e.g., temperature, light, presence of air, etc.), the sample can experience

TABLE 24.12 Generalized Groundwater Sampling Protocol

Step	Goal	Recommendations
1. Hydrologic measurements	Establish nonpumping water level.	Measure the water level to ±1 cm (±0.01 ft).
2. Well purging	Remove or isolate stagnant H_2O which would otherwise bias representative sample.	Pump water until well purging parameters (e.g., pH, temperature, specific conductance, Eh) stabilize to ± 10% over at least two successive well volumes pumped.
3. Sample collection	Collect samples at land surface or in well-bore with minimal disturbance of sample chemistry.	Pumping rates should be limited to ~100 mL/min for volatile organics and gas-sensitive parameters.
4. Filtration	Filtration permits evaluation of characteristics of soluble constituents. It should be performed in the field as soon as possible after sample collection.	Filter trace metals, inorganic anions/cations, alkalinity. Do not filter samples to be analysed for TOC, TOX, volatile organic compounds, other organic compounds.
5. Preservation	Retard chemical changes that could affect the chemistry of a sample after it is extracted from the well.	See Table 24.11
6. Field determinations	Field analyses of samples will effectively avoid bias in determining parameters/constituents which can change significantly after sample collection (e.g., temperature, alkalinity, pH).	Parameters or constituents which are not preservable should be analyzed in the field if at all possible.
7. Field blanks/standards	These blanks and standards will permit the correction of analytical results for changes which may occur after sample collection, preservation, storage, and transport.	At least one blank and one standard for each sensitive parameter should be made up in the field on each day of sampling. Spiked samples are also recommended for complete QA/QC.
8. Sample storage/transport	Refrigerate and protect samples to minimize their chemical alteration prior to analysis.	Observe maximum sample holding or storage periods recommended by the agency. Actual holding periods should be carefully documented.

Data from USEPA. 1990. *Handbook — Ground Water, Volume I: Ground Water and Contamination*. EPA/625/6-90/016a. Washington, DC.; USEPA. 1992a. *RCRA Ground-Water Monitoring: Draft Technical Guidance*, Office of Solid Waste, EPA/530-4-93-001, PB93-139-350, Washington, DC.

geochemical changes that could render the sample unrepresentative if it is not properly preserved. Sample preservation methods are intended to retard volatile loss and chemical reactions such as oxidation, biodegradation, and sorption. Preservation methods are generally limited to pH adjustment, refrigeration, and protection from light. A list of appropriate preservation measures is presented in Table 24.11.

Sample storage refers to measures used to maintain sample quality during transportation. Samples should be cooled to a temperature of 4°C as soon as possible after they are collected and should be maintained at that temperature until they are received at the laboratory. Samples should be shipped in containers that minimize agitation of the samples and should be accompanied by proper documentation (see Section 24.4.8). Note that most samples must be analyzed within a specified period (i.e., holding time) after the sample is collected; data from samples that are analyzed after the holding time limit has been exceeded are considered to be unreliable. Coordination between the field sampling team and the laboratory can help to prevent exceedance of holding times. Sample storage requirements and maximum recommended holding times are listed in Table 24.11.

24.4.8 Documentation of Sampling Events

Each sampling event should be documented so that the validity of the sample collection, preservation, and storage techniques can be verified in the future. If this information cannot be verified, then (depend-

ing on the data quality objectives specified for the project) the data could be rendered invalid and resampling could be required. An example field sampling documentation log form is presented in Table 24.13. As shown in the table, each aspect of the sampling event should be recorded, including the time of sampling, ambient weather conditions, time required to purge the well and the volume of water purged, purge water characteristics, decontamination procedures, sample equipment calibration, and procedures for preservation of samples. Also, chain-of-custody of documentation should be prepared. The record of sampling should be stored in the project files for future reference.

24.5 Analysis of Groundwater Samples

24.5.1 Field Chemical Analyses

Many analyses can be performed in the field if a field laboratory is properly established; in fact, several parameters must be measured in the field to be relevant. Analyses that should be performed in the field include measurement of temperature, pH, specific conductance, and redox potential. Temperature measurements that are not made in the field immediately after the sample is collected are not representative of *in situ* conditions because the temperature of the sample may change rapidly after the sample is removed from the well. Measurements of pH and redox potential are strongly affected by temperature and, therefore, should also be measured in the field. These analyses can be performed quickly using readily available, relatively inexpensive equipment. A good discussion of field measurement techniques is presented by USEPA (1993a).

Many of the analyses that are typically performed in the laboratory (which are described in Section 24.5.2) can be performed in the field by setting up a field laboratory. Field laboratories have become increasingly common as manufacturers of analytical equipment develop increasingly reliable, portable analytical devices. The advantages of performing analyses in the field include fast return of results, decreased chance of sample disturbance or changes in sample chemistry during shipment, ability to obtain an additional sample quickly if problems occur during analysis, and lower cost per sample (if large quantities of analyses are performed). The disadvantages of performing analytical tests in the field include difficulty in implementing quality assurance measures, increased potential for contamination of samples and equipment, less sophisticated instrumentation (which usually results in higher method detection limits and lower precision and accuracy), and correspondingly less reliable test results. However, in spite of these shortcomings, field analysis can be an extremely valuable and time-saving approach for site characterization studies and for monitoring of remediation activities.

24.5.2 Laboratory Chemical Analyses

Most analyses of groundwater samples are performed in an offsite, permanent laboratory to provide a high level of quality control, precision, and accuracy of analyses. Laboratory analyses may be conducted using a variety of analytical techniques and methods. Analytical "technique" refers to a particular procedure and type of instrument that is used in the laboratory to analyze the sample; an analytical technique may incorporate one or more analytical methods to analyze a sample. Note that appropriate techniques will be a function of required detection limits and regulatory requirements. Laboratories will often recommend an appropriate technique and method. The particular analytical technique depends on the analyte of concern and the particular design of the analytical instrument by the equipment manufacturer. When reporting analytical results, the laboratory may present a qualifier with the data to clarify the validity of the test result. Qualifier types and definitions vary between laboratories, but some qualifiers are used widely by most laboratories. Typical qualifiers are presented below (USEPA, 1992a); the laboratory should be required to submit a list of qualifier definitions with the reports of laboratory analyses.

J = value is estimated (typically below the method detection limit).
U = compound analyzed for but undetected (value presented is the quantitation limit).

TABLE 24.13 Example of a Field Sampling Log (*Source:* USEPA, 1993)

GROUND WATER SAMPLING INFORMATION FORM*

General Information

Location (Site/Facility Name)	Sampling Point (common name)
Project Name/#	Type (mon. well, spring, etc.)
Field Personnel	Field Sample (Event) ID#*
Sampling Organization	Facility ID (for IGWIS data entry)
Weather ☼ ?	Station ID (for IGWIS data entry)

Sampling Station (Well) Details

Read from left to right ☞   top – bottom

Well Depth (ft. below MP)	Casing Diameter (inches)	Open Interval (depth below GS) (0.1 ft)
Static Depth to Water (below MP) (0.01 ft)	Static DTW (ft below GS) (0.1 ft)	Date ____ Time
Water Column Length (L) (ft)	One WC Volume (cu ft)	One WC Volume (gals)
Condition: Securely Locked? Y or N	Station (Well) Damaged? Y or N	Surface Contamination (visible) Y or N

Purging

PID/FID Reading @ Wellhead*	Concentration ____ ppm	Background Conc.	____ ppm
Free Product (order LNAPL of DNAPL)	Detected/Sampled? Y or N / Y or N	Appearance	
Well Purging Equipment	Pump, bailer?	Type*	
Purging Date/Time	Start ⌚ ____/____	Finish ⌚	____/____
Pump/Bailer Intake Set at	Feet below MP	Avg. Purge Rate	____ gpm
Amt. Purged before Sampling	Gals/WC Volumes ____/____	Purge Protocol of ___ WCV's met?	Y or N

Field Water-Quality Measurements and Observations

Date/Time Measurements Began ____/____ Purge Rate for Measurements (gpm) _____
Submersible Pump with direct line to Flow Cell used for all Field Water Quality Measurements? ____ Y or N
All Field Measurement Instruments Calibrated according to Protocol? ____ Y or N
All Field Water Quality Parameters Stabilized according to Protocol Criteria just before filling sample containers? ____ Y or N
The Measurements below Represent: (1) stabilization, (2) sample water collected, (3) both a and 2, (4) other*: _____
Sample Appearance: ☞ _____ Odor: _____

Field Measurement	Value	Military Time	Comments*
Temperature	°C		
Electrical Conductivity	µMhos/cm		= meter reading x magnitude x k
Specific Conductance	µMhos/cm		EC corrected to 25 °C
pH	Standard Units		
Dissolved Oxygen	mg/l		
Eh	mV		
Turbidity	NTU		

Sample Collection

Sampling Device (type of pump/bailer)*	Sample Medium (well water, LNAPL, etc.)*
Permanently Installed Pump? Y or N Dedicated Equipment? Y or N	Used Same Equip. for Purge? Y or N
Pump Intake/Bailer Set at (ft below MP)	Interval Samples Represent (ft below GS) Top - / Bottom -
Date / Time Sampling Began	Date / Time Sampling Finished
Depth to Water (ft below MP)	Depth to Water (ft below MP)
QC Samples Collected? Y or N (see reverse*)	Sample Withdrawal Rate
All Field Protocols were followed with no exceptions (Y,N)	**Enter Protocol Codes* 1.____ 2.____**
Remarks (1)* (include protocol exceptions) _____	
Form Completed by _____ Y or N (see reverse*) Date _____	

Form GWS #?
Revised 9-2-93

60 _____ *See 2 of this form contains definition of abbreviations, protocol codes, additional room for equipment specification, QC sample description and other comments

R = data does not meet laboratory QC requirements and is rejected.

N = compound is tentatively identified (i.e., its presence is likely, based on the judgment of the laboratory, but not conclusively confirmed by the test result).

The most commonly used laboratory analytical methods include chromatography (i.e., detection of constituents based on the rate that they migrate through a stationary medium), spectroscopy (i.e., identification of constituents based on the changes in light spectrum caused by irradiated light as the light passes through the groundwater sample), and photometry (i.e., the measurement of the intensity of light or the relative intensity of different lights as they pass through a groundwater sample) (USEPA, 1993b). Some of the more commonly used laboratory analytical techniques are described below (USEPA, 1993b).

Gaseous-phase analyses are performed using instruments designed to detect constituents in gases or that require conversion of the sample to the gaseous phase before analysis. Gaseous-phase techniques include photo-ionization, flame-ionization, explosimetry, gas chromotography, mass spectrometry, and atomic adsorption spectrometry.

Luminescence/spectroscopic analyses involve exciting a sample (using, for example, X-rays or other means of radiation), and measuring the radiation emitted by the compound, either during excitation or as the electrons in the constituent return to their original state. Luminescence/spectroscopic techniques include X-ray fluorescence (i.e., measurement of the secondary radiation emitted when a sample is excited using X-rays), fluorometry (i.e., measurement of the radiation emitted when electrons in a molecule return to their original state after excitation), or spectrometry (i.e., measurement of the absorption spectra of narrow bandwidths of radiation from the sample).

Wet chemistry analyses include a wide range of colorimetric (e.g., titration, colorimetry, filter photometers, and spectrophotometers), immunochemical (e.g., enzyme immunoassay and fluoroimmunoassay), liquid chromotography, and electrochemical (e.g., voltammetry, polargraphy, pH, Eh, dissolved oxygen, and electrical conductance) techniques. These procedures are relatively straightforward and inexpensive but they require strict application of QA/QC procedures, are time consuming, require different reagents for each analyte of concern, and have limited application for some toxic chemicals.

Other analytical techniques include radiological, gravimetric, magnetic, microscopic, biological, and chemical sensor techniques. Each of these applies to particular constituents. The potential advantages of these techniques include: (1) economy of scale when analyzing for a small list of constituents from a large number of samples; (2) better reliability of results for certain compounds; and (3) (in some cases) simple operation of equipment. Disadvantages could include expensive equipment, difficult QA/QC procedures, and high cost of analyses. Refer to USEPA (1993b) for a discussion of these and other candidate analytical techniques.

24.6 Evaluation of Groundwater Monitoring Data

24.6.1 Data Validation

Data validation refers to the process of verifying that the groundwater sampling and analysis was performed in accordance with the requirements of the groundwater monitoring plan and in accordance with the requirements of applicable QA/QC procedures. Data that does not comply with these specific requirements must be considered to be potentially invalid. Example causes of invalid data include: improper well development techniques or removal of too few well volumes of groundwater during purging; improper sampling, filtration, and preservation techniques; failure to analyze the samples before the recommended holding time is exceeded; the presence of air in VOC sample containers; failure to properly implement chain-of-custody procedures; improper calibration of laboratory analytical equipment; presence of contaminants in field blank, method blank, or trip blank samples; or failure of the analytical laboratory to properly implement or document QA/QC procedures. If data has been found to be potentially invalid, then one of the following may be required: (1) if remaining portions of the samples

exist, then they may be analyzed and substituted for the invalid data; (2) additional samples can be collected and analyzed; or (3) documentation of the sampling event can be reviewed and, if it is found that the improper procedure would not have impacted the analytical results, notations can be made in the project files regarding the occurrence and resolution of the problem, with the analytical result tentatively used in data evaluations.

24.6.2 Statistical Analyses

24.6.2.1 Statistical Evaluation of Data

After data have been validated, they should be evaluated to determine if groundwater quality conditions have changed since the last sampling event. As a first step in this process, the presence of the contaminants should be confirmed beyond a reasonable doubt and an evaluation of the contribution of the contaminants by the potential source should be performed. This first step is typically made by analyzing the data using statistical analyses, which can be used as the basis for concluding whether a contaminant is present in groundwater. An adequate quantity of data is required before a statistical analysis can be performed; often, samples from multiple background wells and from multiple sampling events must be available. A discussion of statistical analysis of groundwater analytical data is presented in USEPA (1989b).

Statistical analyses are based on the probability theory that the likely range of all values (e.g., constituent concentrations at any location within the monitored area) can be estimated based on the distribution of the known values. If a groundwater quality data set is found to be consistent with a standard distribution (e.g., normal, lognormal, etc.), then the probability that the concentration will exceed a certain value (e.g., a groundwater protection standard) can be estimated. Two key parameters that are used in nearly all statistical analyses are the mean and the standard deviation of the data set:

$$\bar{x} = \left(\sum_{i=1}^{n} x_i \right) \Big/ n \tag{4}$$

$$s = \left[\sum_{i=1}^{n} \left(x_i - \bar{x} \right) \Big/ \left(n - 1 \right) \right]^{\frac{1}{2}} \tag{5}$$

where $\bar{x}$ = sample mean; i = observation number; x_i = the value for the ith observation; n = total number of observations; and s = sample standard deviation. Groundwater quality data are usually well-characterized by a normal or lognormal distribution, although some data cannot be adequately characterized by any common distribution (USEPA, 1989b). Because most statistical tests are based on an assumption regarding normality or lognormality, the data must be tested for conformance to either a normal distribution or other type of distribution that is assumed by the particular statistical test. One test for evaluating the fit of a data set to a normal distribution is the coefficient of variation test. To perform this test, the sample mean and standard deviation are first calculated. Then the sample standard deviation is divided by the sample mean; if the result is greater than one, then the data do not fit a normal distribution well. If the data are found not to fit a normal distribution, then a log transformation should be performed on the original data and the coefficient of variance test rerun to see if the data fit a lognormal distribution. If the data fit either a normal or a lognormal distribution, then a parametric approach (described in Section 24.6.2.2) to statistical analyses can be used; otherwise, a nonparametric analysis (also described in Section 24.6.2.2) must be used. Other more accurate but more complicated tests are available to test the fit of the data to a normal or lognormal distribution (e.g., probability plots, coefficient of skewness, the Shapiro-Wilk test, and the probability plot correlation coefficient, all described in USEPA (1989b)).

24.6.2.2 Tests for Statistical Significance of Detected Constituents

To select a statistical approach, the type of comparison being made (e.g., compliance well vs. background, intrawell, etc.) must be selected, and then a method that best suits the approach. Types of comparisons and recommended statistical methods are presented in Table 24.14. As shown in the table, the recommended method of statistical analysis varies based on the type of comparison being made and the number of times the compound of concern was detected. If the constituent was not previously detected, then the analytical results should simply be compared to the quantitation limit of the test apparatus.

TABLE 24.14 Recommended Statistical Methods for Groundwater Monitoring

Compound	Type of Comparison	Recommended Method
Any compound in background	Background vs. compliance well Intrawell	ANOVA tolerance intervals prediction intervals Control charts
ACL/MCL specific	Fixed standard	Confidence intervals tolerance intervals
Many nondetects in database (synthetic)	% Nondetects	Recommended Method
	$0 \leq x \leq 15\%$	Replace nondetect values with one half the method detection limit or quantitation limit (as appropriate) then proceed with any of the following parametric procedures: • ANOVA • Tolerance intervals • Prediction intervals • Control charts If residuals from transformed do not meet parametric ANOVA requirements, use nonparametric approaches.
	$15\% \leq x < 50\%$	Treat nondetect values as ties, then either proceed with nonparametric ANOVA, or use Cohen's adjustment, then proceed with: • Tolerance intervals • Confidence intervals • Control charts
	50% to 99%	Test of proportions
	100%	Compare sample best result values to sample quantitation limits

Key: ND = Nondetect   QL = Quantitation Limit
      MDL = Method Detection Limit   ACL = Alternate Concentration Limit
From USEPA. 1989b. *Statistical Analysis of Ground-Water Monitoring Data at RCRA Facilities, Interim Final Guidance.* Office of Solid Waste, Washington, DC.

One of the most powerful statistical tests for groundwater monitoring is the Analysis of Variances (ANOVA) test, also known as the "general linear model procedure." This method is useful for comparing background data to compliance well data and is useful in situations where constituents are detected in many wells at differing concentrations. The method is used to compare the means of different groups of observations (i.e., from different wells on the same sampling date) and to identify significant differences among the groups; if there are significant differences, then the data can be further analyzed to identify contrasting aspects of the data set. If the compound is not routinely detected in groundwater, then the nondetect data points are replaced with nonzero approximations of data before proceeding with the statistical test. The analysis can be performed using either a parametric or a nonparametric approach. The parametric ANOVA procedure assumes that the differences (called residuals) between the mean of the values and the values themselves are normally distributed with equal variance; this case can be checked as described in Section 24.6.2.1. If this test shows that the residuals do not meet this assumption, then

the nonparametric ANOVA procedure must be performed. For a nonparametric ANOVA, it is first assumed that there is no contaminant present (i.e., all of the data comes from uncontaminated groundwater having the same continuous distribution of constituents) and that, therefore, the median concentration of contaminants must be the same at all wells. The nonparametric ANOVA is performed by comparing the median concentration of hazardous constituents for all wells and then, if the median differs significantly, by comparing the average rank of each of the compliance wells with the average rank of the background well.

Another test for statistical significance that was used extensively in the 1980s is the student's t-test. The student's t-test is based on the t-distribution (which is a statistical distribution analogous to a normal or lognormal distribution) that is used to estimate the likelihood that the mean value of a number of observations will exceed a certain value with a certain confidence. The test is sometimes modified according to Cochran's Approximation to Behren's Fischer's (CABF) t-test (as described in Appendix IV to Title 40 of the *U. S. Code of Federal Regulations*, Part 264), which is simply a method for implementing the t-test. The test is no longer widely used because it compares mean values of sets of data, which is inaccurate because a test result from a specific monitoring well is not a mean value as assumed by the t-test but, rather, a single observation that could fall anywhere within the statistically expected range for that parameter (i.e., probably not at the mean value). Also, the CABF has been found to sometimes provide excessively high occurrences of false-positive or false-negative results.

24.6.3 Database Management

The data sets that are generated during groundwater monitoring events can be very large and difficult to manage. For monitoring programs that involve either a large number of wells or a very long period of time, it may be useful to manage the data in a database management system. Database management systems are computer programs that store and analyze data. Such programs can be used to categorize data according to result (e.g., a listing of all detections of a particular constituent), location (e.g., listing of detections in a certain well), or time (e.g., all detections of a certain constituent over a certain period of time). Also, database management programs can be equipped with statistical analysis packages or geographic information systems to further simplify statistical analysis or presentation of the data.

24.6.4 What to Do If a Statistical Analysis Indicates the Presence of Contaminants

If the results of a statistical analysis indicate that contaminants are present in a particular groundwater monitoring well, then the data should first be reexamined to verify that the test results were not a false-positive indication of contamination. The data should be examined for the following problems that may cause a false-positive indication of groundwater contamination. See also Nielsen (1991), USEPA (1989b), and Miller and Miller (1986) for discussions of false-positive groundwater monitoring results.

- *Natural variation in groundwater quality.* Groundwater quality sometimes varies naturally according to season, temperature, geology, or a number of other factors. If the statistically significant increase is due to a constituent that is present in uncontaminated groundwater (e.g., manganese, sodium, iron, etc.), then the historic data for the site should be reexamined to evaluate the expected range of values for the detected constituent. Constituents of concern may be naturally present in groundwater and may not be the result of the potential source; for example, arsenic and lead are naturally present at levels exceeding MCLs in uncontaminated groundwater in many parts of the United States. These false positives can be identified and resolved by obtaining adequate representative background data.

- *Contaminant source different from the potential source.* The potential source being monitored may not be the actual source of contamination. This could be confirmed by either demonstrating that a hydraulic interconnection exists between an alternative source and the well having the statistically significant increase or by showing that the contaminants present in the well could not have been

derived from the potential source, given the type of waste in the potential source. Another type of false-positive result is the contamination of groundwater samples by condensation of landfill gas in monitoring wells.

- *Well construction problems.* The statistically significant increase may be due to a problem with well construction. For example, high pH or alkalinity values may indicate that grout from the well is "bleeding" down into the filter pack, either through the aquifer formation or through an inadequate bentonite seal. Alternatively, the statistically significant increase may be a result of poor well maintenance; an example of such a problem is infiltration of contaminated surface water ponded around the well at the ground surface.

- *Analytical error.* The statistically significant increase may be the result of laboratory error. For example, methylene chloride and acetone (which are commonly used in the laboratory to clean test devices) are routinely detected in groundwater samples but are often a false-positive indication of contamination. Such analytical errors can usually be detected by preparing and analyzing laboratory, trip, and method blanks. Contaminants present in laboratory or trip blanks that are also present in groundwater samples are not indicative of groundwater contamination. Good QA/QC procedures and documentation by the laboratory can usually lead to resolution of analytical errors.

- *Statistical error.* These errors are typically mathematical errors or inappropriate application of statistical procedures (e.g., improper consideration of nondetects in sample data sets). If such a problem is expected, then the results of the statistical analysis should be reviewed by an expert in statistical analysis.

- *Sampling error.* These errors include problems in the field that occur during sampling, such as improper filtering of samples, improper evacuation of wells before sampling, agitation of wells during sampling, agitation of samples during transfer to sample containers, and improper preservation of samples. This source of error is typically the most difficult to evaluate because, usually, very little documentation is generated during actual handling of the sample (because the sample technician's hands are occupied), making it difficult to isolate the source of the problem. The best method of preventing such problems is to use properly trained professional field technicians and to routinely implement good QA/QC procedures during sampling. The use of good field forms is an effective way of ensuring that key observations are made and recorded.

If no source can be identified for the statistically significant increase, then the statistically significant increase should be attributed to the potential source being monitored and subsequent measures (such as those described in Section 24.2.2, Step 6) should be implemented.

For Further Information

Nielsen (1991) gives a good, broad description of groundwater monitoring techniques. Driscoll (1986) gives a thorough discussion of groundwater well construction techniques and materials of construction for groundwater wells. USEPA (1993a,b) gives a detailed description of monitoring and analytical techniques for subsurface media. Complete citations for these works can be found in References.

References

American Society for Testing and Materials Standard D5092-90. 1990. *Standard Practice for Design and Installation of Ground Water Monitoring Wells in Aquifers.* Philadelphia.

American Society for Testing Materials (ASTM) 1992. *Standard Guide for Sampling Groundwater Monitoring Wells, Standard D4448-85a; ASTM Standards on Environmental Sampling.* 1995. Philadelphia, 220-233.

Barcelona, M.J., Gibb, J.P., Helfrich, J.A., and Garske, E.E. 1985. *Practical Guide for Ground-Water Sampling.* EPA/600/2-85/104.

Blake, S.B. and Hall, R.A. 1984. Monitoring petroleum spills with wells: some problems and solutions. Proceedings of the Fourth National Symposium and Exposition on Aquifer Restoration and Ground Water Monitoring. National Water Well Association. Columbus, OH. 305-310.

Chappelle, F.H. 1993. *Groundwater Microbiology and Geochemistry*. John Wiley & Sons, New York.

Cohen, R.M. and Mercer, J.W. 1993. *DNAPL Site Evaluation*. C. K. Smoley, Boca Raton, FL.

Driscoll, F.G. 1986. *Groundwater and Wells*. Johnson Division, St. Paul, MN.

Edge, R.W. and Cordry, K. 1989. The HydroPunch®: An *in situ* sampling tool for collecting ground water from unconsolidated sediments. *Ground Water Monitoring Review*, 177-183.

Fetter, C.W. 1993. *Contaminant Hydrogeology*. Macmillan Publishing Co., New York.

Freeze, R.A. and Cherry, J.A. 1979. *Ground Water*. Prentice-Hall, Inc., Englewood Cliffs, NJ.

Hem, J.D. 1989. Study and Interpretation of the Chemical Characteristics of Natural Water. U.S. Geological Survey Water Supply Paper 2254, 3rd Edition.

Hvorslev, M.J. 1949. *Subsurface Exploration and Sampling of Soils for Civil Engineering Purposes*. U.S. Waterways Experiment Station, Vicksburg, MS. (Reprinted by Engineering Foundation, New York.)

Lambe, T.W. and Whitman, R.W. 1969. *Soil Mechanics*. John Wiley & Sons, New York.

Lueder, D.R. 1959. *Aerial Photographic Interpretation*. McGraw-Hill Book Company, New York.

McDonald, M.G. and Harbaugh, A.W. 1988. *A Modular Three-Dimensional Finite-Difference Groundwater Flow Model*. USGS Techniques of Water-Resources Investigations, Book 6, Chapter A1, USGS, Reston, VA.

Miller, J.C. and Miller, J.N. 1986. *Statistics for Analytical Chemistry*. John Wiley & Sons, New York.

Moberly, R.L. 1985. Equipment Decontamination. *Ground Water Age*. 19(8):36-39.

Nielsen, D.M., ed. 1991. *Practical Handbook of Ground-Water Monitoring*. Lewis Publishers, Chelsea, MI.

Puls, R.W. and Barcelona, M.J. 1989. Filtration of ground water samples for metals analysis. *Hazardous Waste and Hazardous Materials*. 6(4).

Rumer, R.R. and Mitchell, J.K. 1995. *Assessment of Barrier Containment Technologies — A Comprehensive Treatment for Environmental Remediation Applications*. National Technical Information Services, Springfield, VA.

Scalf, M.R., McNabb, J. F., Dunlap, W.J., Crosby, R.L., and Fryberger, J. 1981. *Manual of Ground-Water Quality Sampling Procedures*. EPA/600/2-81/160 (NTISPB82-103045).

USEPA. 1975. *Manual of Water Well Construction Practices*. USEPA Office of Water Supply, Report No. EPA-5709-75-001. Washington, DC.

USEPA. 1986a. *RCRA Ground-Water Monitoring Technical Enforcement Guidance Document*. Office of Waste Programs Enforcement and Office of Solid Waste and Emergency Response, OSWER-9950.1. Washington, DC.

USEPA. 1986b. *Test Methods for Evaluating Solid Wastes*. Office of Solid Waste and Emergency Response, Washington, DC, SW-846.

USEPA. 1987a. *The Use of Models in Managing Ground-Water Protection Programs*. EPA 600/8-87/003. Washington, DC.

USEPA. 1987b. *Data Quality Objectives for Remedial Response Activities, Development Process*. EPA 840/G-87/003. Washington, DC.

USEPA. 1989a. *Risk Assessment Guidance for Superfund: Interim Final Guidance*. Office of Emergency and Remedial Response (EPA/540/1-89/002). Washington, DC.

USEPA. 1989b. *Statistical Analysis of Ground-Water Monitoring Data at RCRA Facilities, Interim Final Guidance*. Office of Solid Waste, Washington, DC.

USEPA. September 1989c. *Seminar Publication — Transport and Fate of Contaminants in the Subsurface*. EPA/625/4-89/019. Washington, DC.

USEPA. 1989d. *Handbook of Suggested Practices for the Design and Installation of Ground-Water Monitoring Wells*. PB90-159-807. Washington, DC.

USEPA. 1990. *Handbook — Ground Water, Volume I: Ground Water and Contamination*. EPA/625/6-90/016a. Washington, DC.

USEPA. 1991. *Handbook — Ground Water, Volume II: Methodology*. EPA/625/6-90/016b. Washington, DC.

USEPA. 1992a. *RCRA Ground-Water Monitoring: Draft Technical Guidance*, Office of Solid Waste, EPA/530-4-93-001, PB93-139-350, Washington, DC.

USEPA. 1992b. *Statistical Analysis of Ground-Water Monitoring Data at RCRA Facilities, Addendum to Interim Final Guidance*, Office of Solid Waste, Washington, DC.

USEPA. 1993a. Office of Research and Development. *Subsurface Characterization and Monitoring Techniques: A Desk Reference Guide, Volume I: Solids and Ground Water, Appendices A and B*. U.S. Environmental Protection Agency, Washington, DC. EPA/625/R-93-003a.

USEPA. 1993b. Office of Research and Development. *Subsurface Characterization and Monitoring Techniques: A Desk Reference Guide, Volume II: The Vadose Zone, Field Screening and Analytical Methods; Appendices C and D.* U.S. Environmental Protection Agency, Washington, DC. EPA/625/R-93/003b.

USEPA. 1993c. *Technical Manual — Solid Waste Disposal Facility Criteria*. EPA530-R-93-017, PB94-100-450.

USEPA. 1994. *Handbook — Groundwater and Wellhead Protection*. EPA/625/R-94/001.

USEPA. 1995. *Ground Water Sampling — A Workshop Summary*, November 30 – December 2, 1995, Dallas, Texas. EPA/600/R-94/205.

Wilson, C.R., Einberger, C.M., Jackson, R.L., and Mercer, R.B. 1992. Design of ground-water monitoring networks using the monitoring efficiency model (MEMO). *Ground Water Age*, 30(6):965-970.

Glossary

Anisotropy   Variation in geologic formations in the vertical or horizontal direction.

Anthropogenic   Resulting from, or caused by, human activities.

Aquitard   A lithologic unit that impedes, but does not completely prevent, groundwater movement.

Background Sampling   Evaluation of initial water quality conditions.

Bridging   The development of gaps caused by obstructions in either grout or filter pack materials   Also refers to blockage of particles in natural formation materials or artificial filter pack materials that may occur during well development.

Casing   An impervious, durable pipe placed in a borehole to prevent the walls of the borehole from caving and to prevent flow of undesirable groundwater, surface water, gas, or other fluids into the well.

Chain of Custody   A method for documenting the possession history of a sample from the time of its collection through its final disposition.

Confining Unit   A relatively low-permeability material stratigraphically adjacent to one or more aquifers.

Contaminant   A substance that is not normally present in groundwater or that is present at an unusually high concentration.

Cross-Contamination   The movement of contaminants between aquifers or water-bearing zones through an unsealed or improperly sealed borehole or other conduit.

Detection Limit   The lowest concentration of a chemical that can be reliably detected by an analytical device.

Development   The act of removing materials introduced during drilling from a well and adjacent aquifer formation.

DNAPL   A dense nonaqueous phase liquid that is relatively immiscible and has a density greater than that of water. Also known as "free product" or a "sinker."

Downgradient   In the direction of decreasing hydrostatic head.

Drilling Fluid   A fluid (liquid or gas) that is used in drilling operations to remove cuttings from the borehole, to clean and cool the drill bit, and to maintain the integrity of the borehole during drilling.

Equipotential Surface   A surface, in a three-dimensional groundwater flow system, for which the total hydraulic head is the same at every point on the surface.

False Negative   A condition for which contamination is present, but the results of sample analyses fail to indicate its presence.

False Positive   A condition for which no contamination is present, but the results of sample analyses indicate presence of contamination.

Field Blank   A laboratory-prepared sample of known properties that is transported to the sampling site for use in validating field sampling procedures.

Filter Pack   A clean silica-sand or sand and gravel mixture that is installed in the annular space between the borehole wall and the wellscreen for the purpose of retaining and stabilizing the particles from the adjacent strata.

Groundwater Protection Standard   The acceptable quantity of contaminants in groundwater; the groundwater protection standard could be an MCL or another value.

LNAPL   A light nonaqueous-phase liquid. Also known as "free product" or a "floater."

Maximum Contaminant Level (MCL)   The concentration at which the excess cancer risk of exposure to humans is estimated to be 1×10^{-6}, as defined by the U.S. EPA.

Nonaqueous   Fluids that are relatively insoluble.

Parts Per Billion (ppb), or Parts Per Million (ppm)   Unit weight of solute per billion (or per million, respectively) unit weights of solution (solute, plus solvent).

Piezometer   A well that is used only to measure the elevation of the water table or the potentiometric surface.

Potentiometric Surface   A surface that represents the level to which water will rise in a tightly cased well.   The "water table" is the potentiometric surface for an unconfined aquifer.

Preservation   Measures taken during sample storage to minimize the change in concentration of a constituent of interest until analyses can be performed (e.g., storage of sample in acidic solution).

Quantitation Limit   The lowest concentration at which a chemical can be accurately and reproducibly quantitated. Usually equal to the instrument detection limit times a factor of three to five.

Quality Assurance   A management function used to establish and monitor quality control protocols and to evaluate their outcomes.

Quality Control   Technical and operational procedures that are used to investigate and confirm the proper conduct of field, sample transportation, and laboratory activities necessary to assure accuracy and precision in the data.

Statistically Significant   Exceedance of a certain level of probability, based on the results of a statistical analysis.

Target Monitoring Zone   The portion of an aquifer for which there is a reasonable likelihood that a vertically placed well will intercept migrating contaminants.

Upgradient   In the direction of increasing hydrostatic head.

Volatile Organic Compounds (VOCs)   Compounds which will partition relatively easily into the air phase, sometimes defined as compounds with Henry's law constant greater than around 10^{-6} atm-m^3/mol.

Wellscreen   A filtering device used to retain the filter pack and aquifer materials while allowing groundwater to enter the well, commonly constructed of slotted or perforated casing material.

25

Remediation of Contaminated Groundwater

Michael F. Houlihan and
Patrick C. Lucia
GeoSyntec Consultants

25.1 Fundamentals

25.1.1 Introduction

In this chapter, techniques are described for remediating contaminated groundwater. The purpose of this chapter is to present a comprehensive approach for developing and implementing a groundwater remediation program; examples of such approaches and references for further information on such approaches are presented. In the remaining portion of this subsection on fundamentals, the purpose of, and typical goals for, groundwater remediation programs are addressed, and the role of risk assessment in groundwater remediation programs is described. Approaches to groundwater remediation are described based on the current requirements of the United States Environmental Protection Agency (U.S. EPA) for remediating sites that are regulated under the Resource Conservation and Recovery Act (RCRA). Other requirements, such as local regulations (as described in Chapter 22) and the need to mitigate human

health concerns posed by the contaminated groundwater may also define the needs of a groundwater remediation program.

25.1.2 Groundwater Remediation Goals

Groundwater is an extremely important source of water. In the United States, about 50% of all drinking water is obtained from groundwater sources. Further, a significant majority of Americans live near industrialized population centers, which are typically located near groundwater supplies. In the past several decades, several instances of groundwater contamination have illustrated the adverse impacts that contaminated groundwater can have on human health. For all of these reasons, protecting the integrity of groundwater supplies is crucial to the protection of human health and the environment.

The need for remediation of contaminated groundwater is typically established based on the results of a groundwater assessment or monitoring program (see Chapter 24). In general, if the results of groundwater monitoring indicate that groundwater contains contaminants at concentrations that make it a threat to human health or the environment, then the groundwater may need to be remediated. Remediation, in this sense, is a broad term that refers to the reduction of risk caused by exposure to contaminated groundwater.

One of the most important steps in a remediation program is defining the goals of the program. There are many different goals that can be defined for a groundwater remediation program, including the following (see NRC, 1994):

- *Complete restoration*, which involves removal of all contaminants from the contaminated aquifer
- *Nondegradation*, which involves removal of contaminants that exceed either the detection limits of available analytical equipment or background concentrations
- *Remediation to health-based standards*, which involves removal of contaminants that are present at a concentration that could cause adverse health effects (some examples of health-based standards are maximum contaminant levels (MCLs), alternate concentration limits (ACLs; see Section 25.1.3), and local or federal drinking-water standards)
- *Remediation to the limits of technology-based standards*, which involves use of the best available technology to remove as much of the contaminants as possible
- *Partial-use restrictions* (or *institutional controls*), such as legal restrictions on the use of groundwater in areas where groundwater has been contaminated, or physical barriers (e.g., fences) to prevent access to contaminated media
- *Containment*, which involves the use of engineered systems for preventing migration of the contaminants to locations where receptors could be exposed to the contaminants.

A brief summary of the advantages and disadvantages of each of these remediation goals is presented in Table 25.1. When selecting a remediation goal, the degree to which groundwater can actually be remediated should be examined. For example, for the past twenty years, pump-and-treat has been considered to be one of the best technologies currently available for restoring groundwater quality. However, based on recent studies by the U.S. EPA, the actual success rate for pump-and-treat remedies is extremely low. A summary of the U.S. EPA's findings, as presented by the National Research Council (NRC) (1994), is provided in Table 25.2. Although there were various reasons for the failure of the systems to meet their goals, the data strongly suggest that, for most circumstances, best available technologies are not capable of restoring groundwater to its original quality. However, returning groundwater to its useful purpose or eliminating health and environmental risks does not always require restoration to original quality. Accordingly, a growing number of remediation systems are being designed with the goal of either containment or remediation to health-based standards. A typical approach for achieving this remediation goal is, first, to estimate the risk associated with exposure to the contaminated groundwater and then design a remediation system that will improve groundwater quality to a level that is protective

TABLE 25.1 Advantages and Disadvantages of Cleanup Goals

Goal	Advantages	Disadvantages
Complete restoration	Eliminates all risk	Likely impossible
Nondegradation	Reduction of contaminants to lowest level measurable	Extremely difficult, expensive, and time-consuming for many contaminants and hydrogeologic settings
Health-based standards	Designed to prevent measurable impacts to human health or environment	Health-based standards are difficult to define and may not accurately address all possible health impacts of exposure to contaminated groundwater
Technology-based standards	Allows treatment to the best capabilities of current technology	May not reduce risk to a level that is protective of human health and the environment
Partial use restrictions	Prevents contact between contaminants and receptors in a cost-effective manner	Leaves contaminants that could cause risk if partial use restrictions are ineffective
Containment	Relatively predictable and reliable; typically less costly than other remediation approaches	Leaves contamination that could migrate if containment system fails

(Left margin, vertical arrows: Increasing Flexiblity ↑, Increasing Cost ↑, Increasing Protection ↑)

Adapted from National Research Council (NRC), Water Science and Technology Board. 1994. *Alternatives for Groundwater Cleanup.* Washington, D.C. With permission.

TABLE 25.2 Summary of Pump and Treat System Performance Data

Category[1]	No. of Sites	No. of Sites Containment Achieved	No. Sites Cleanup Goal Achieved	No. Sites Goal Not Achieved
1	2	1	1	1
2	13	8	4	9
3	19	12	4	15
4	36	18	0	36

Note: [1] Indicates relative ease of cleanup, with 1 easiest and 4 hardest, as defined by NRC (1994) and as illustrated in Table 25.8.

This summary is based on information provided by NRC (1994) for sites where data were available regarding attainment of the cleanup goal.

Adapted from National Research Council (NRC), Water Science and Technology Board. 1994. *Alternatives for Groundwater Cleanup.* Washington, D.C. With permission.

of human health and the environment (i.e., health-based standards). An approach for deriving an achievable, health-based remediation goal is presented in Section 25.1.3.

After a remediation goal has been selected, a remedy must be implemented to achieve the goal. In general, groundwater can be remediated in one of the following three manners.

- *Containment.* Containment of a plume of contaminated groundwater involves preventing the plume from migrating to a location where receptors can be exposed to it. Techniques for containing contaminated groundwater are described in Section 25.3.

- *Extraction.* For some remediation approaches, groundwater must be extracted from the ground, either to allow *ex situ* treatment or to provide hydraulic control of contaminated groundwater. Groundwater extraction approaches are described in Section 25.4.

- *Treatment.* Contaminated groundwater can be treated, either *in situ* or *ex situ*, using a variety of physical, chemical, or biological methods. Techniques for treatment of contaminated groundwater are described in Section 25.5.

Before beginning a program of groundwater remediation, it is recommended that the groundwater professional define a stepwise approach to developing and implementing the program. Recommended steps for developing and implementing a plan for groundwater remediation are presented in Section 25.2.

25.1.3 Risks Associated with Contaminated Groundwater

Contaminated groundwater may pose a risk to human health or the environment if it comes into contact with receptors and if the groundwater contains concentrations of constituents that could affect the health of the receptors. Risk can be defined as the probability of occurrence of human health or environmental impacts due to a release of contaminants to the environment. Risk can be evaluated by performing a baseline risk assessment, as described by U.S. EPA (1989a) and as summarized in Figure 25.1. The goal of performing a baseline risk assessment is to estimate the current and future risks associated with contaminated groundwater and to compare the estimated risk to the acceptable threshold of risk (definition of an "acceptable" level of risk is addressed at the end of this subsection); if the acceptable threshold level of risk is exceeded, groundwater remediation is likely warranted. Risk assessment techniques can also be used to estimate the risk that would remain after a remediation program has been implemented. Estimating the resulting future risk for several candidate remedies can provide a basis for selecting a remedy that reduces risk most effectively. Based on the results of a risk assessment, a non-zero concentration limit for contaminants in groundwater could be proposed that results in an acceptable risk to potential future receptors. For several contaminants, U.S. EPA has defined exposure concentrations that result in a risk of 10^{-6}; these concentrations are defined as "Maximum Contaminant Levels" (or MCLs). U.S. EPA also allows owners of contaminated sites to propose, on a case-by-case basis, Alternate Contaminant Levels (ACLs) that result in a site-specific exposure level (after considering the fate and transport of the contaminants) of 10^{-6}.

There are numerous sources of uncertainty in currently available risk-assessment methods (NRC, 1994). First, there is very little data to establish a conclusive link between human exposure to specific contaminants and specific health or environmental impacts. The accuracy of epidemiologic studies performed on humans to date is limited, with significant uncertainties related to: (1) exposure quantity and duration; (2) latency of observable impacts; (3) the relatively small size of the study populations; (4) inadequate control over comparison (i.e., "control") groups; (5) inability to attribute observed health problems to only the known exposure; and (6) the synergistic effects in humans of exposure to some contaminants. Similarly, the accuracy of epidemiologic studies on animals is limited by: (1) the need to extrapolate the results of large-dose impacts on animals to estimate impacts of small doses on humans; and (2) the focus of such studies on single-contaminant exposures. Finally, because the concentrations of contaminants in groundwater may change over time due to factors such as treatment or attenuation over the exposure period (i.e., a lifetime), the quantity of contaminants to which a receptor may be exposed (as calculated in Step 2, *Exposure Assessment*) could change during the period under consideration. Because of these uncertainties, risk assessment results have a degree of uncertainty. To account for this degree of uncertainty, very conservative assumptions are typically used when estimating toxicity and exposure quantity to minimize the possibility of underestimating the actual risk. Therefore, the estimated risk may be extremely conservative.

After the risk of exposure to contaminated groundwater has been estimated, the groundwater professional must decide whether or not the risk is high enough to warrant remediation. This decision process, which is referred to as risk management (a process that is separate from risk assessment), is typically performed based on input from both regulatory agencies and the population likely to be exposed to the risk. For example, a decision has been made by the U.S. Congress, as presented in the National Contingency Plan, that a risk level of 10^{-4} to 10^{-6} is acceptable for contaminants at Superfund sites. The U.S. EPA typically considers a similar range of risk to be acceptable at RCRA sites. The acceptable level of risk may vary from location to location and is a balance between the benefits to the public of a clean environment and the public cost of attaining a clean environment. The benefit of a clean environment is difficult to determine, but can be estimated in terms of dollars that the public would be willing to
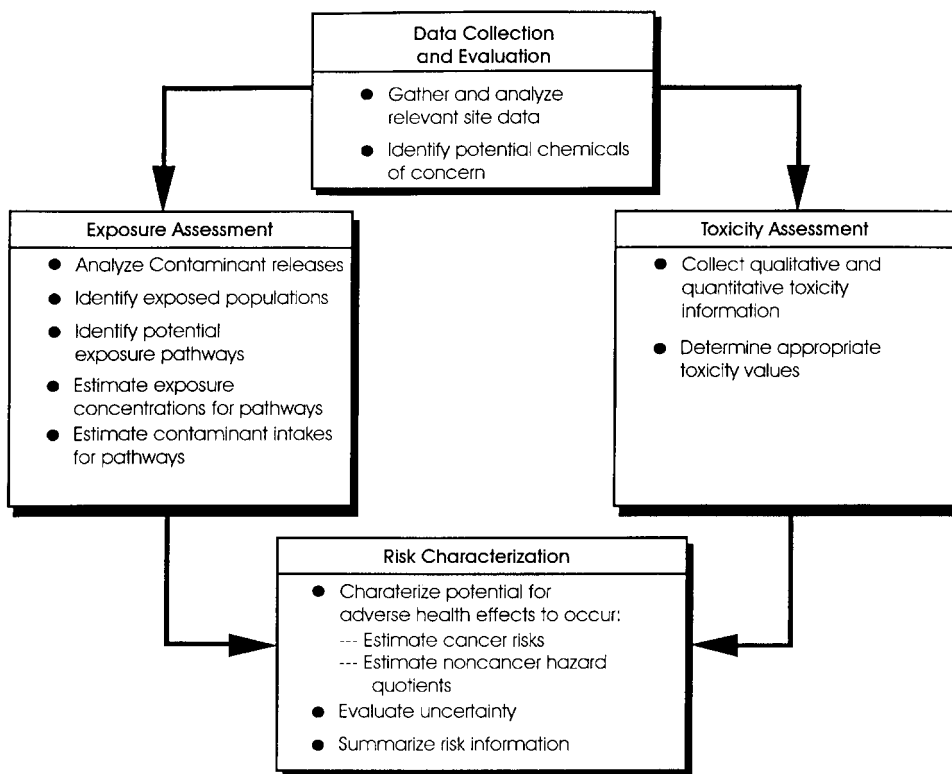
FIGURE 25.1 Baseline risk assessment. (From USEPA. 1989a. *Risk Assessment Guidance for Superfund: Interim Final Guidance*. Office of Emergency and Remedial Response (EPA/540/1-89/002). Washington, DC.)

spend to clean up the environment (i.e., the "contingent valuation" approach); however, this approach is abstract and contains flaws that are currently being evaluated (NRC, 1994). Because of difficulties in obtaining an accurate estimate of the value of a clean environment, the commonly-used value for acceptable risk (e.g., 10^{-6}) has been arbitrarily defined.

25.2 Groundwater Remediation System Design

25.2.1 Introduction

The design of an efficient, cost-effective groundwater remediation system requires a comprehensive understanding of the nature and extent of contamination, the remediation objectives, and a careful evaluation of remedial technologies and their abilities to meet the objectives. When preparing a remediation system design, the use of a stepwise approach can simplify the design process and minimize the chances of failure of the system to achieve the remediation goal. In this section, recommended steps are presented for preparing a design to remediate contaminated groundwater. In subsequent sections, approaches to the design of some typical remediation systems are described.

25.2.2 Step 1. Define the Problem

The first step in the design of the groundwater remediation system is to define the problem in adequate detail to allow the design of an efficient and cost-effective groundwater remediation system. The problem should be defined in terms of the extent of contamination, the risk associated with the contaminated groundwater, the regulations that apply to the problem, the degree of remediation required, and the

subsurface conditions. The extent of contamination may be defined using procedures described in Chapter 24 of this handbook. If the vertical and horizontal extent of contamination have not been adequately evaluated during the site exploration or groundwater monitoring program, then additional field exploration data should be collected to complete this evaluation. The risk associated with contaminated groundwater can be estimated as described in Section 25.1.3 of this chapter. The regulations that apply may be identified based on the information presented in Chapter 22 of this handbook, and based on input from legal counsel, if necessary. Subsurface conditions may be defined as described in Chapter 1. The various manners in which groundwater can become contaminated are illustrated in Figure 25.2; hydrogeologic considerations for groundwater remediation are summarized in Table 25.3 and 25.4. As illustrated in Table 25.5, the hydrogeology of the site (as well as other factors) has a strong bearing on the type of remediation system that can be considered for a site.
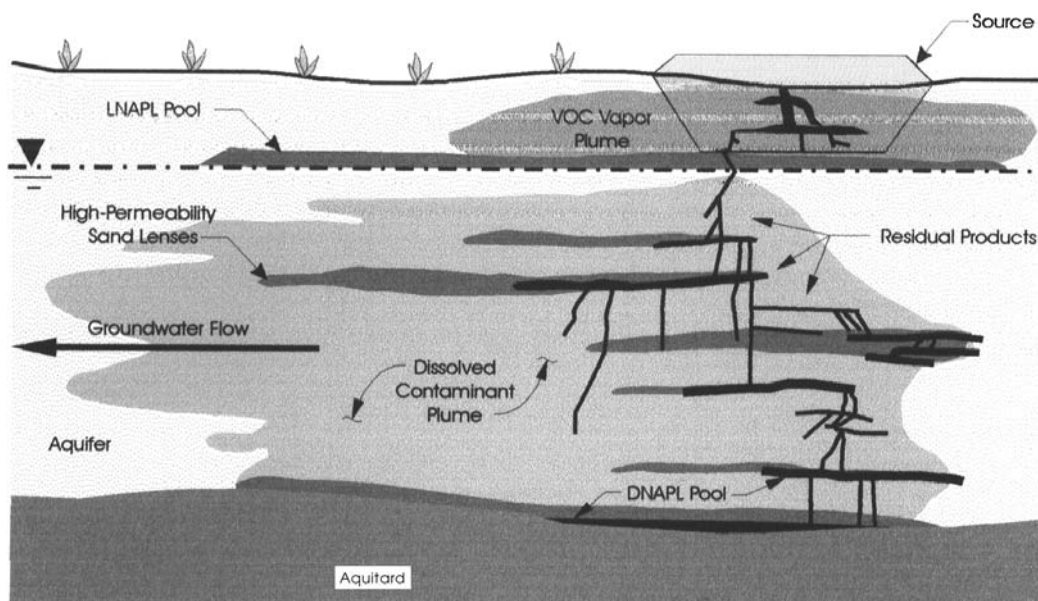
FIGURE 25.2 Groundwater contamination by dissolved vapor and nonaqueous phase constituents.

25.2.3 Step 2. Define the Goal of Groundwater Remediation

The next step in designing a groundwater remediation system is to define the goal of remediation. This step involves: (1) identifying remediation technologies that could possibly be used to remediate the contaminated groundwater; then (2) screening the list to identify remediation technologies that could be successfully implemented at the site; and finally (3) selecting an appropriate remediation goal based on the remedies that can be successfully implemented. A summary of remediation technologies that could be used is presented in Table 25.6. Candidate goals for remediating contaminated groundwater are described in Section 25.1.2; candidate remediation approaches that correspond to these goals are presented in Table 25.7.

For sites where it is demonstrated that remediation is needed, the remediation goal should be developed with as complete an understanding as possible of the goals that are actually attainable. This begins with an understanding of the difficulty of remediating or restoring groundwater quality in certain hydrogeologic settings, as summarized in Table 25.5. Then, the possibility of achieving the remediation goal should be evaluated based on past experience with similar problems in similar hydrogeologic settings. For example, the likelihood of success of groundwater restoration (based on the profession's documented experience in achieving health-based remediation goals) is illustrated in Table 25.8, which was developed based on an analysis of the information presented in Table 25.2. As shown in Table 25.8, it is currently

TABLE 25.3 Hydraulic Properties of Aquifers Important for Groundwater Cleanup

Property	Description	Importance for Groundwater Cleanup
Porosity	Volume of pore space relative to the total volume	Pores store water and contaminants
Effective porosity	Interconnected pore space that can transmit fluid	Water and contaminants flow through interconnected pores
Groundwater velocity	Rate of fluid movement	Influences the direction and velocity of dissolved contaminant movement
Hydraulic gradient	Elevation and pressure differences that cause fluids to flow	Influences the direction of contaminant movement
Hydraulic conductivity	Ease with which water can move through a geologic formation	Influences the rate at which fluid can be pumped for treatment, as well as the rate and direction of groundwater migration
Transmissivity	Product of formation thickness and hydraulic conductivity	Influences the rate at which a plume of contaminants migrates and the rate at which fluid can be pumped for treatment
Storage coefficient	Volume released by pressure changes per unit area during pumping in a confined aquifer	Influences the quantity of fluid that can be obtained by pumping
Specific yield	Fraction of total pore volume released as water by gravity drainage during pumping of an unconfined aquifer	Influences the quantity of fluid that can be obtained by pumping
Specific retention	Fraction of total aquifer volume retained as water above the water table after pumping an unconfined formation	Influences the quantity of contaminant that remains in the subsurface after pumping

Adapted from National Research Council (NRC), Water Science and Technology Board. 1994. *Alternatives for Groundwater Cleanup.* Washington, D.C. With permission.

TABLE 25.4 Summary of the Mechanisms Influencing the Fate of Contaminants in the Environment

Process	Environmental Conditions	Contaminant
Movement (i.e., advection)	Water flow rate	Amount of material
	Formation permeability	Physical state
	Water motion	Solubility
	Gravity	Viscosity
	Surface tension	
Retention (physical processes affecting transport)	Soil/Sediment	Type solubility
	Organic matter content	Ionic character
	Sorptive capacity	
Reaction (chemical processes affecting transport)	pH	Chemical transformation
	Redox status	Biodegradability
	Microbial communities	

From USEPA. September 1989b. Seminar Publication — *Transport and Fate of Contaminants in the Subsurface*. EPA/625/4-89/019. Washington, DC.

feasible to remediate some types of sites to health-based standards (e.g., homogeneous aquifers contaminated with dissolved, mobile contaminants, as shown in "Group A" in Table 25.8), but not other types of sites (e.g., heterogeneous or fractured rock aquifers contaminated with non-aqueous phase liquid contaminants, "Group C" in Table 25.8). As a further consideration, restoration of groundwater to original quality represents a tremendous financial burden to the nation, as illustrated in Table 25.9. Regardless

TABLE 25.5 Examples of Factors Affecting Groundwater Remediation

Characteristic	Generalized Remediation Difficulty Scale		
	Less Difficult ⟶		⟶ More Difficult

Nature of release	Small Volume ⟶		Large Volume
	Short Duration ⟶		Long Duration
	Slug Release ⟶		Continual

Chemical Properties

Biotic/Abiotic decay potential	High ⟶		Low
Volatility	High ⟶		Low
Contaminant retardation (sorption) potential	High ⟶		Low

Contaminant Distribution

Contaminant phase	Aqueous, Gaseous ⟶ Sorbed ⟶ LNAPLs ⟶ DNAPLs		
Volume of Contaminated media	Small ⟶		Large
Contaminant depth	Shallow ⟶		Deep

Geology

Stratigraphy	Simple Geology, ⟶ e.g., Planar Bedding	Complex Geology, e.g., Interbedded and Discontinuous Strata, fractured media
Texture of unconsolidated deposits	Sand ⟶	Clay
Degree of heterogeneity	Homogeneous (e.g., well-sorted sand) ⟶	Heterogeneous (e.g., interbedded sand and silts, clays, fractured media, karst)

Hydraulic/Flow

Hydraulic conductivity	High (10^{-2} cm/sec) ⟶	Low (10^{-4} cm/sec)
Temporal variation	Little/None ⟶	High
Vertical flow	Little ⟶	Large Downward Flow Component

From USEPA. 1993b. *Guidance for Evaluating the Technical Impracticability of Ground-Water Restoration.* EPA-540-R-93-080. Washington, DC.

of the goal selected, it should be noted that, as stated by the National Research Council (1994), "The attainment of zero contaminant concentration as an outcome for groundwater remediation should be recognized as an unattainable goal no matter how far cleanup technologies advance in the future." The difficulty of restoring groundwater quality at some sites is also addressed by the U.S. EPA in its document entitled, *Guidance for Evaluating the Technical Impracticability of Ground-water Restoration* (1993b). In this document, the U.S. EPA identifies several hydrogeologic settings (based on information in Table 25.5) in which, based on the U.S. EPA's experience, groundwater quality likely cannot be restored to background conditions.

Once the remediation goal has been selected and feasible remedies have been identified that could be implemented at the site, the groundwater professional should screen the list of candidate remedies to those that, based on experience, can be expected to achieve the remediation goal. These remedies should then be further evaluated during Step 3.

TABLE 25.6 Summary of Commonly Used Groundwater Remediation Technologies

Alternative Technology	Residual Groundwater Concentration	Residual Sorbed Concentration in Source Area	Cleanup Time[a]	Number of Peer-Reviewed Publications[b]
Source Remediation				
Conventional pump and treat	Low to medium	Medium to high	Long	Some
Vacuum extraction and bioventing	NA	Low to medium	Short	Some
Air sparging (vertical or horizontal wells)	Low to medium	Low to medium	Short to medium	Limited
In situ bioremediation-hydrocarbons	Low to medium	Low to high	Short to medium	Some
In situ bioremediation-chlorinated solvents	Low to medium	low to high	Medium to long	Some
Cosolvent and surfactant flushing	Low to medium	Low to medium	Short to medium	Some
Steam stripping	Low to medium	Low to medium	Short	Some
In situ thermal desorption	Low to medium	Low to medium	Short	Some
In situ chemical oxidation	Medium (?)	Medium to high (?)	Medium (?)	Limited
In situ bioremediation-metals	Low to medium	Low to high	Medium to high	Limited
Intrinsic bioremediation	Low to medium	Low to high	Long	Limited
Plume Remediation				
Conventional pump and treat	Low	Medium to high	Long	Many
Air sparging (vertical or horizontal wells)	Low to medium	Low to medium	Medium to long	Limited
In situ bioremediation-hydrocarbons	Low to medium	Low to high	Medium to long	Some
In situ bioremediation-chlorinated solvents	Low to medium	Low to high	Medium to long	Some
In situ reactive barriers	Low	NA	Long	Limited
Intrinsic bioremediation	Low to medium	Low to medium	Long	Limited

Note: A "low" residual concentration and "medium" cleanup time reflect relatively good performance, while a "high" residual concentration and "long" cleanup time reflect much less effective performance. "NA" denotes that the technology is not applicable to this situation. Cost of the technology and feasibility for different situations are not addressed on this table but should be carefully evaluated before selecting a technology.

[a] Because few cases of achieving cleanup goals have been reported, these qualitative assessments reflect the judgment of NRC.

[b] "Limited" indicates that very little information about this technology is available in peer-reviewed publications, while "some" indicates a greater availability of information.

Adapted from National Research Council (NRC), Water Science and Technology Board. 1994. *Alternatives for Groundwater Cleanup.* Washington, D.C. With permission.

25.2.4 Step 3. Screen Candidate Remedies

After a list of feasible candidate remediation technologies has been prepared, the remedies should be screened to identify the remedy that can most efficiently and cost effectively achieve the remediation goal. Screening involves evaluating the likely effectiveness and cost of the feasible remediation technologies and selecting a remedy for detailed design based on the results of this evaluation. Three valuable steps in performing the evaluation are described below.

- *Prepare Preliminary Conceptual Design for Each Candidate Remedy.* By preparing a conceptual design, the effectiveness and cost of the remedy can be evaluated, as described below. Developing and implementing design criteria can be an effective way to focus the conceptual design effort on the keys to success for the project. Design criteria are guidelines to be used during design that, if achieved, will likely result in successful performance of the remediation system. These criteria are

TABLE 25.7 Candidate Groundwater Remediation Goals and Corresponding Remediation Approaches

Remediation Goal	Examples of Candidate Remediation Approaches
Groundwater restoration	Natural (intrinsic) remediation
	Bioremediation
	Combinations of conventional remediation approaches
Nondegradation	Source removal
	Conventional pump-and-treat
	Bioremediation
	SVE/AS
Return aquifer to health-based standards	Source removal
	Conventional pump-and-treat
	Bioremediation
	SVE/AS
Apply technology-based standards	Phytoremediation
	Electrokinetics
	Solvent extraction
	Thermal desorption
Implement restricted-use policies	Deed restrictions
	Local ordinances
Containment	Physical barriers
	Hydraulic barriers
	Capping

Adapted from National Research Council (NRC), Water Science and Technology Board. 1994. *Alternatives for Groundwater Cleanup.* Washington, D.C. With permission.

TABLE 25.8 Likelihood of Success of Groundwater Restoration

	Contaminant Chemistry					
Hydrogeology	Mobile, Dissolved (degrades/ volatizes)	Mobile, Dissolved	Strongly Sorbed, Dissolved (degrades/ volatizes)	Strongly Sorbed, Dissolved	Separate Phase LNAPL	Separate Phase DNAPL
Homogeneous, single layer	A (1)	A (1-2)	B (2)	B (2-3)	B (2-3)	B (3)
Homogeneous, multiple layers	A (1)	A (1-2)	B (2)	B (2-3)	B (2-3)	B (3)
Heterogeneous, single layer	B (2)	B (2)	B (3)	B (3)	B (3)	C (4)
Heterogeneous, multiple layers	B (2)	B (2)	B (3)	B (3)	B (3)	C (4)
Fractured	B (3)	B (3)	B (3)	B (3)	C (4)	C (4)

Note: Group A represents types of sites for which cleanup of the full site to health-based standards should be feasible with current technology. Group C represents types of sites for which full cleanup of the source areas to health-based standards will likely be technically infeasible. Group B represents sites for which the technical feasibility of complete cleanup is likely to be uncertain. The numerical ratings indicate the relative ease of cleanup, where 1 is easiest and 4 is most difficult.

Adapted from National Research Council (NRC), Water Science and Technology Board. 1994. *Alternatives for Groundwater Cleanup.* Washington, D.C. With permission.

defined by the owner, engineer, and the operator. Design criteria can be developed that address constructibility concerns, performance limitations of the technology, budget constraints, operation and maintenance requirements, and regulatory requirements. The candidate remedies can be compared by examining the results of these evaluations and selecting the remedy that can most

TABLE 25.9 Costs of Various National Policies for Hazardous Waste Site Remediation

National Policy	Present Value of Resource Cost (billions of 1991 U.S. dollars)		
	Lower Bound	Best Guess	Upper Bound
Current policy[a]	180	280	390
Less stringent policy [b]	140	180	260
More stringent policy[c]	360	440	630

Note: This table converts the figures in Russell et al. (1991) to present value by assuming that costs are prorated equally each year for a 30-year time period and that the discount rate is 4%.

[a] According to Russell et al. (1991), current policy means "the set of principles and practices for hazardous waste remediation that are inferred to be in place in the period 1988-91 when the experience base and data for this study were collected."

[b] According to Russell et al. (1991), less stringent policy means relying more on containment and less on full cleanup.

[c] According to Russell et al. (1991), more stringent policy means application of more intensive treatment technologies and reduced burden on future generations.

Adapted from National Research Council (NRC), Water Science and Technology Board. 1994. *Alternatives for Groundwater Cleanup.* Washington, D.C. With permission.

cost effectively achieve the remediation goal. Conceptual designs should then be prepared that address the design criteria that are relevant at the conceptual design stage.

- *Evaluate Effectiveness of Conceptual Designs.* The effectiveness of the designs in meeting the remediation goal should be evaluated next. This evaluation can be performed objectively (for example, by modeling the performance of the remediation system as described in Chapters 14 and 15 of this handbook) or subjectively (using the information on remedies presented in Table 25.6 and the information on remediation feasibility presented in Table 25.8). Also, the time required to achieve the remediation goal should be estimated for use in the subsequent cost evaluation.

- *Evaluate Cost of Remedies.* The cost of each remedy should be estimated. A good reference for remediation cost estimating techniques is presented in USEPA (1985). In particular, the present-worth value of each remedy should be calculated, considering both capital and long-term operation and maintenance costs, to provide a consistent basis for comparing the costs of the candidate remedies.

25.2.5 Step 4. Prepare Detailed Design

After the preferred remedy has been selected, a detailed design of the approach should be prepared that is suitable for construction of the system. Recommendations for design of specific elements of some of the more commonly used groundwater remediation systems are presented in Sections 25.3, 25.4, and 25.5. It is recommended that detailed calculations be performed to demonstrate that each of the design criteria is satisfied. Also, it is recommended that the ability of the design to mitigate adverse impacts to human health or the environment be demonstrated; ideally, this is done by performing a risk assessment (as described in Section 25.1.3) for the conditions that will exist during and after operation of the remediation system. An effective remedy will achieve the remediation goal, mitigate the risk of exposure to contaminated groundwater, and provide reliable service throughout the design life of the system.

Like all engineering design activities, remedial design is an iterative process. Therefore, the design should be critiqued several times to identify improvements that can be made to enhance the performance or reduce the cost of the system. Such an evaluation could be performed by subjecting the preliminary design to a detailed review by: (1) a peer reviewer who is qualified in the technology being considered and who has not been an integral part of the design process (to provide an unbiased, critical evaluation of the design); (2) a contractor (to provide an evaluation of the constructibility of the design); (3) a risk assessment professional (to give an indication of the ability of the design to meet the remedial action objectives for mitigation of risks to human health or the environment); or (4) an operator of a similar system (for a value-engineering review or an evaluation of the potential costs or difficulties in operating

the proposed system). Although such reviews may appear to be somewhat time consuming, they frequently result in valuable input that can improve the likelihood that the design criteria will be met. Also, because such reviews reduce the possibility of failure of the remediation system, they can be an effective component of a liability and loss prevention program.

25.2.6 Step 5. Implement the Design

After the design has been completed and reviewed, it must be implemented, which involves constructing and operating the system. The first step in constructing the system is deciding how and by whom the system will be constructed. There are several methods of arranging to construct a remediation system, including: (1) construction by the owner; (2) soliciting competitive bids from independent contractors qualified to do the work; (3) negotiating a contract with an independent contractor; or (4) retaining a design/build contractor. For the last three approaches to construction, a detailed set of bid documents should be prepared that can be used by the bidder to prepare a responsive bid. The advantages of procuring a contractor through a competitive bid process include: (1) a generally lower construction cost, resulting from the competitive nature of the bid process; (2) well-defined roles of each party to the construction contract and well-defined terms of the construction contract; and (3) documentation of the information that the contractor used to prepare its bid, which can be valuable if additions or subtractions to the contract scope of work and budget are required. The disadvantages of competitively bidding work are: (1) the formality required, which involves a preparation of detailed plans and legal documents, sometimes at a significant cost; and (2) difficulty in factoring in "subjective" criteria in the bid evaluation, such as the contractor's qualifications for performing the work and the contractor's past performance on similar projects.

25.2.7 Step 6. Confirm the Effectiveness of the Design

The final step of a remediation program is confirmation of the effectiveness of the system in achieving the remediation goal. The effectiveness of the design can be evaluated simply by reviewing the data from a performance monitoring program (as described in Chapter 24) and assessing whether or not the remediation goals and the performance criteria are being satisfied. The effectiveness of the design should be evaluated both during the startup of operations (if the system contains components that require operation and maintenance) and at specified time intervals after startup. As an example of confirming the effectiveness of a groundwater pump-and-treat system design that is intended to prevent migration of contaminants beyond a line of extraction wells, the system could be monitored to ensure that the groundwater downgradient of the extraction wells does not exceed design cleanup criteria; similarly, if the system is designed to remove contaminants in the aquifer, then groundwater collected by the removal wells could be tested to confirm that groundwater quality is improving as predicted. As a subsequent step after evaluating the performance of the remediation system, the system should be modified as needed to improve the system's ability to achieve the remediation goal and performance criteria. As described in Section 25.6, the performance monitoring program should contain criteria for "shutting off" the system (as verified using data collected during performance monitoring); also, the performance monitoring program should contain criteria for reevaluating the adequacy of the design if the system has failed to meet the remedial design criteria.

25.3 Hydraulic Containment of Groundwater

25.3.1 Overview

There are many alternatives for providing hydraulic containment of groundwater. The most widely used features for containing groundwater are physical barriers and hydraulic barriers. Therefore, this section is focused on physical barriers (Section 25.3.2) and hydraulic barriers (Section 25.3.3). Also, innovative

hydraulic containment methods are described in Section 25.3.4. Recommended methods for performance monitoring of hydraulic containment features are summarized in Section 25.6.

25.3.2 Physical Barriers

Physical barriers are vertical features in the ground that provide a barrier to groundwater flow. Physical barriers can be formed using a variety of materials and can be either nonselective (i.e., a barrier to the flow of all groundwater) or selective (i.e., a barrier to only the migration of target contaminants). In this section, only nonselective physical barriers are discussed; selective physical barriers (i.e., "permeable reaction barriers") are described in Section 25.5.2.5. A list of several of the most commonly used physical barriers is presented in Table 25.10, along with their typical dimensions, range of unit costs, and approximate production (i.e., construction) rates.

TABLE 25.10 Types of Physical Barriers

Barrier Type	Width (Ft)	Maximum Depth (apx) (Ft)	Unit Cost ($/SF)	Production Rate per 10 hrs (SF)
Soil bentonite	2–3	80	2–8	2,500–15,000
Cement bentonite	2–3	80	5–18	1,000–8,000
Biopolymer drain	2–3	70	7–25	1,500–5,000
Deep mixing	2.5	90	6–15	1,000–8,000
DM structural	2.5	90	15–30	1,000–3,000
Jet grouting	1.5–3	200	30–80	300–2,500
Grout curtain	one row	200	40–100	200–1,000
Sheet piling	one sheet	150	15–40	2,000–15,000
Geochemical barrier	varies	hundreds	8–400	100–1,000

From Rumer, R.R. and Mitchell, J.K. 1995. *Assessment of Barrier Containment Technologies — A Comprehensive Treatment for Environmental Remediation Applications.* National Technical Information Services, Springfield, VA.

The oldest and most commonly used types of physical barriers are cutoff walls; several types are illustrated in Figure 25.3. Cutoff walls limit migration of groundwater by forming a physical impediment to groundwater flow. The effectiveness of the wall is a function of its continuity, its resistance to degradation by contaminants in groundwater, and its resistance to physical degradation. Note that, because it impedes groundwater flow, the elevation of groundwater may rise (or "mound") on the upgradient side of the barrier; to prevent overtopping or flanking of the barrier by contaminated groundwater, other controls may be needed, such as: (1) groundwater extraction wells or trenches adjacent to the barrier to route mounded groundwater to a discharge point; or (2) permeable segments in the barrier to allow water to pass through controlled locations in the wall.

Soil-bentonite slurry walls are, by far, the most common type of physical barrier used in geotechnical and environmental remediation projects. Slurry walls have been used in the U.S. since the 1940s. Construction of a slurry wall consists of installing a mixture of soil (or other material) and bentonite clay into a vertical trench to form a very low-permeability, vertical barrier to groundwater flow. A general approach for slurry wall construction is illustrated in Figure 25.4 and includes: (1) excavation of a narrow (i.e., about 2 to 3 feet, or 0.6 m to 1 m, wide) vertical trench (if the unsupported trench could cave, then the trench should be temporarily supported using a bentonite slurry); and (2) placing the soil-bentonite mix in the slurry-filled trench in a manner that displaces the bentonite slurry and forms a continuous vertical wall of low-permeability soil-bentonite material. A typical soil-bentonite mix contains about 2 to 4% bentonite clay, 20 to 35% water, 15 to 40% soil finer than the No. 200 U.S. standard sieve, and about 40 to 60% material coarser than the No. 200 U.S. standard sieve (Inyang, 1992). In many cases, the material excavated from the trench can be used as an economical source of the soil component of the soil-bentonite mix. Factors that must be considered in selecting a source for soil to be mixed with the bentonite include: (1) the grain-size distribution of the soil; (2) treatment and/or disposal costs for

Plan views of
upgradient slurry
walls for partial
containment
of buried waste
and/or plume
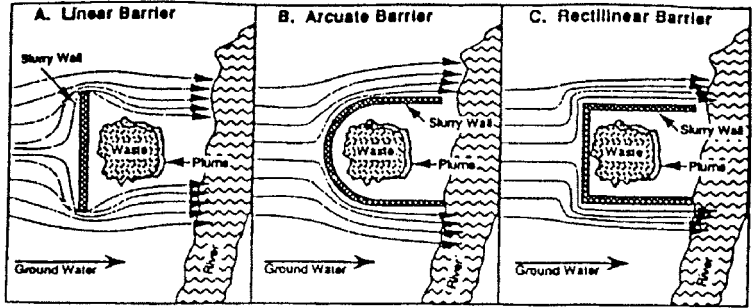
Plan views of
downgradient
slurry walls
for partial
containment
of buried waste
and/or plume
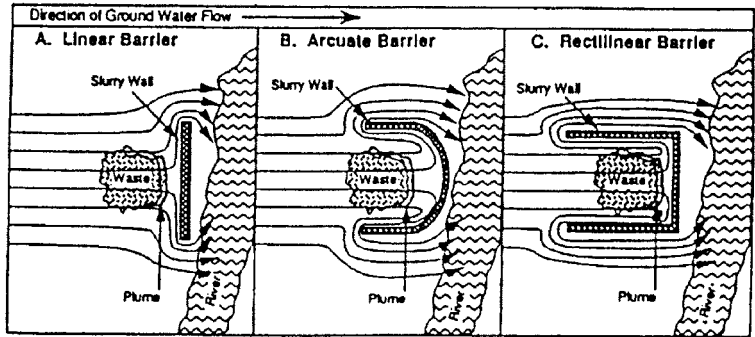
Plan views of
waste totally
enclosed by
slurry walls
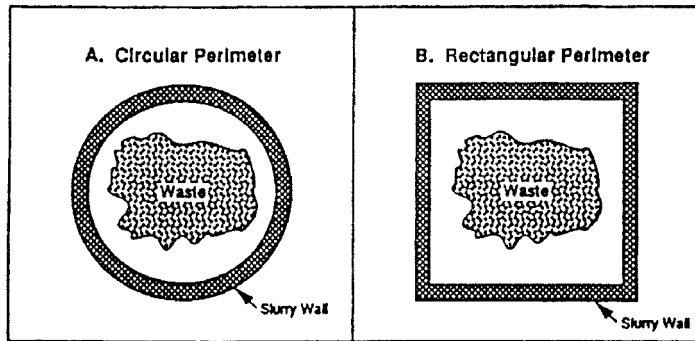
FIGURE 25.3 Examples of physical barrier configurations. (From GeoSyntec Consultants. 1994. *Subsurface Barriers.* Report to the United States Navy, Contract SBIR Topic N93-130. Atlanta, Georgia. 80 pages.)
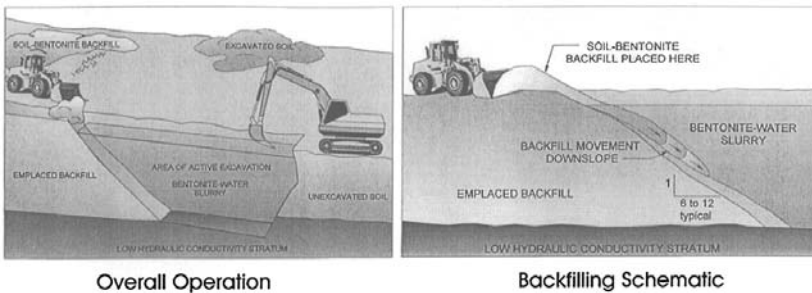
FIGURE 25.4 Soil-bentonite slurry wall construction. (From GeoSyntec Consultants. 1994. *Subsurface Barriers.* Report to the United States Navy, Contract SBIR Topic N93-130. Atlanta, Georgia. 80 pages.)

excavated trench soils that are contaminated; and (3) the chemical composition of the soil (e.g., soil that is contaminated with metals or VOCs could inhibit the ability of the bentonite to hydrate and, therefore, compromise the permeability of the wall). An excellent reference for design and construction of slurry walls is provided by Xanthakos (1979).

The effectiveness of a soil-bentonite slurry wall in preventing flow through the wall is a function of: (1) the environment of the wall (i.e., the nature of the chemicals in groundwater to which the soil-bentonite mix will be exposed); (2) the percentage of bentonite clay used in the soil-bentonite mix (as illustrated in Figure 25.5); (3) the type and gradation of soil used typically measured as a percentage of the soil that is finer than the No. 200 U.S. standard sieve (as illustrated in Figure 25.6); and (4) the quality of wall construction. When designing a soil-bentonite slurry wall, laboratory hydraulic conductivity tests are typically performed on candidate mixtures of soil and bentonite clay, and the test results are used to select the proportions of soil and bentonite that will produce the desired *in situ* hydraulic conductivity. Based on widespread experience with soil-bentonite slurry walls over the past half century, it is generally recognized that the lowest in-place hydraulic conductivity that can reasonably be achieved for a soil-bentonite slurry wall is typically between 10^{-6} cm/sec and 10^{-7} cm/sec, with hydraulic conductivities in the lower end of this range attainable only under ideal circumstances and using excellent construction practices (GeoSyntec, 1994). Finally, note that the performance of a soil-bentonite slurry wall can be adversely affected by exposure of the wall to VOCs, aqueous salts (such as those contained in sea water), or NAPLs. Sai and Anderson (1992) reported an increase of two to three orders of magnitude in hydraulic conductivity for soil-bentonite slurries that were permeated with the VOCs xylene and methanol. To improve the resistance of slurry walls to permeation by VOCs, some researchers are experimenting with the addition of activated carbon to the slurry.
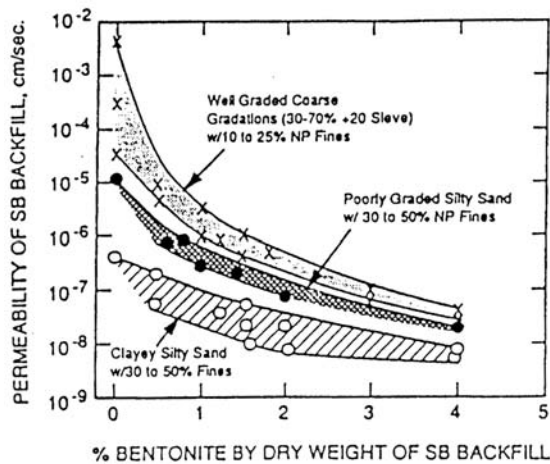
FIGURE 25.5 Permeability as a function of percent bentonite. (From GeoSyntec Consultants. 1994. *Subsurface Barriers.* Report to the United States Navy, Contract SBIR Topic N93-130. Atlanta, Georgia. 80 pages.)

25.3.3 Hydraulic Barriers

The term *hydraulic barrier* is used to describe a feature that causes a depression in the piezometric surface of groundwater and acts as a barrier beyond which groundwater within the zone of influence of the barrier should not flow. Hydraulic barriers may be formed using trenches, wells, or other features that remove groundwater, thus depressing the potentiometric surface. Groundwater collection trenches and groundwater extraction wells, which are the most common types of hydraulic barriers, are described in this section.

Groundwater Collection Trenches. Trenches can be installed in an aquifer to produce an induced hydraulic barrier by depressing the potentiometric surface of the aquifer. A typical groundwater extraction
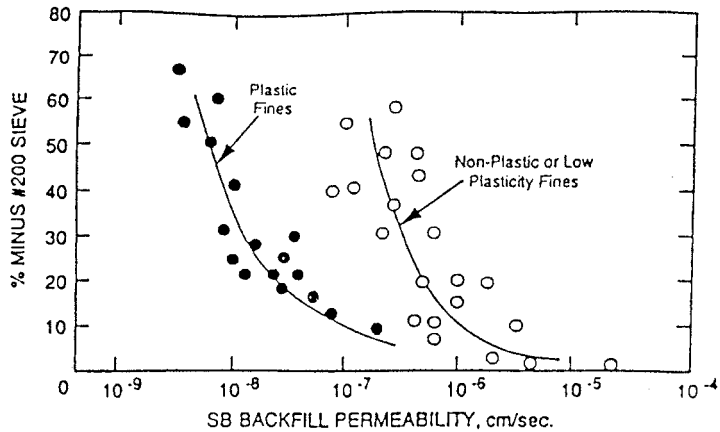
FIGURE 25.6 Permeability as a function of soil fines content. (From GeoSyntec Consultants. 1994. *Subsurface Barriers.* Report to the United States Navy, Contract SBIR Topic N93-130. Atlanta, Georgia. 80 pages.)
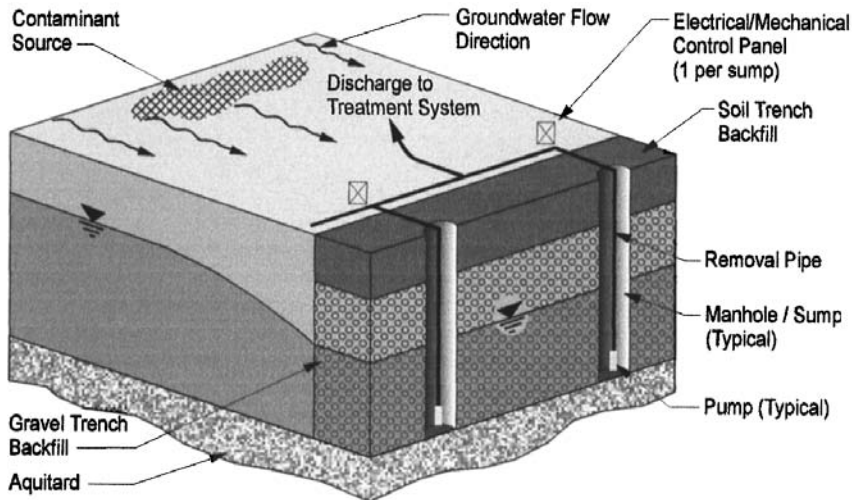
FIGURE 25.7 Typical groundwater extraction trench system.

trench system is illustrated in Figure 25.7. During design, the designer must identify the depth to which the trench must be constructed to induce a hydraulic barrier. The depth of the trench should be carefully selected based on consideration of the seasonal variations in the elevation of groundwater, the location of contaminants (i.e., "floating" LNAPLs, dissolved constituents, or "sinking" DNAPLs), and the depth of contaminants in the aquifer. Trenches may be more cost-effective than wells for extracting groundwater if contaminants are located at a shallow depth below ground (i.e., less than about 50 ft [15m]) and the aquifer materials are relatively easy to excavate (e.g., no excavation of bedrock). The depth of influence of a groundwater collection trench in an isotropic, homogenoeus aquifer is illustrated in Figure 25.8 and can be estimated using the equation (Zheng et. al., 1988):

$$D = \left(\frac{2S}{\pi}\right)^{1/2} \tag{1}$$

where
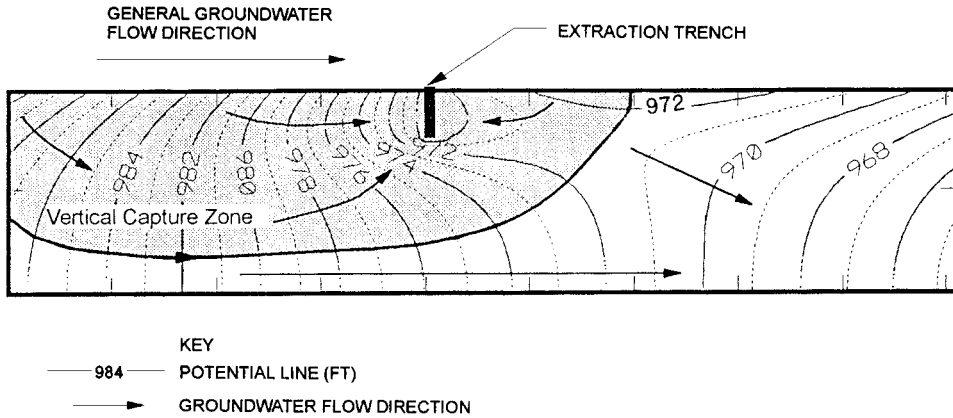
FIGURE 25.8 Vertical zone of influence of partially penetrating groundwater extraction trench.

$$S = 2a\left(h_o - h_d\right) \tag{2}$$

and D is the depth of influence of the ditch (ft), a is equal to one half the width of the ditch (ft), h_o is the head in the aquifer beneath the ditch (ft), h_d is the head in the ditch (ft), and i is the uniform water table gradient in the aquifer. The flow rate of groundwater to the trench per unit width of trench can be estimated using the following equation:

$$Q = KiD \tag{3}$$

where Q is the volumetric flow to the trench (ft^3/sec), K is the hydraulic conductivity of the aquifer material, i is the uniform slope of the potentiometric surface in the aquifer, and D is the depth of influence of the trench calculated from Equation 3. Equations 1 and 3 apply only to isotropic, homogeneous aquifers; for aquifers that contain vertical or horizontal variations in stratigraphy (e.g., clay lenses, sand seams, coarsening-with-depth deposits, etc.), or anisotropies, a more complicated model (see for examples Chapters 14 and 15) may be needed. In such environments, the designer must be aware of the presence of such features and, if necessary, extend the depth or lateral extent of the trench to intercept contaminated groundwater that could flow through these features.

Groundwater Extraction Wells. Wells can also be installed in an aquifer to induce a hydraulic barrier to groundwater flow. A typical groundwater extraction well is illustrated in Figure 25.9. The radius of influence and depth of influence of a single well can be calculated based on the properties of the aquifer, the geometry of the well (i.e., depth and radius), and the pumping rate of the well. Analytical solutions are provided for several combinations of these variables as summarized by Cohen and Miller (1983). The zone of influence of a single well in a homogenous, isotropic, confined aquifer having a non-zero gradient is illustrated in Figure 25.10 and can be estimated as (USEPA, 1994a):

$$W = \frac{Q}{Ti} \tag{4}$$

where W is the width of the zone of influence (ft), Q is the pumping rate (cfs) of the extraction well, T is the transmissivity of the confined aquifer (or the equivalent transmissivity of the unconfined aquifer) (sec^{-1}), and i is the hydraulic gradient (dimensionless). Zones of influence can be estimated for wells in unconfined aquifers using this equation by substituting Kb (where b is the full thickness of the unconfined aquifer at the pumping well) for T if drawdown is small compared to the thickness of the aquifer. More accurate estimates of capture zone geometry in unconfined aquifers can be made using numerical models.
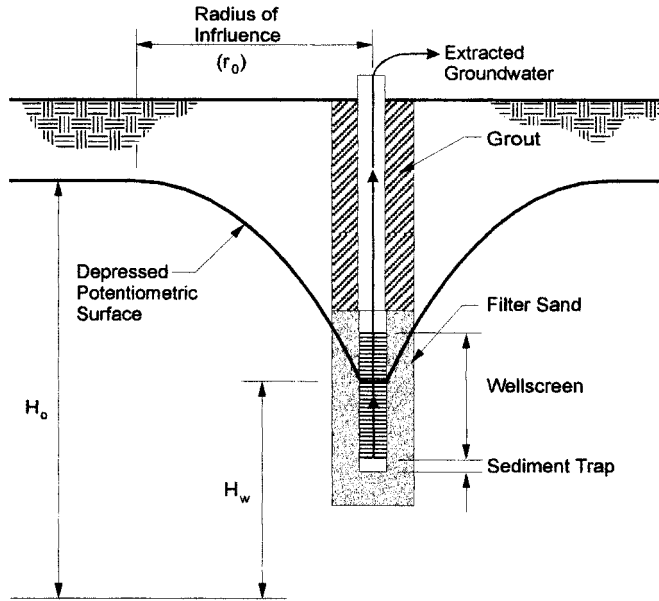
FIGURE 25.9 Groundwater extraction well.
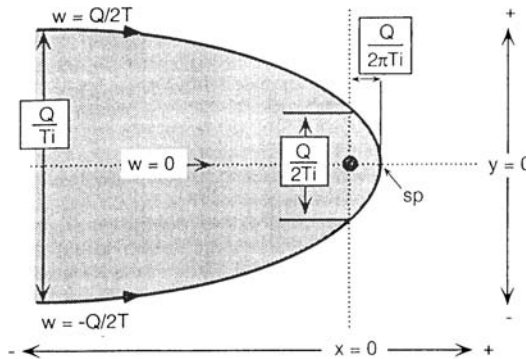
sp = Stagnation Point

FIGURE 25.10 Single-well capture zone. (Data from USEPA. 1994a. *Monitoring the Performance of Groundwater Pump and Treat Systems.* EPA/600/R-94/123. Washington, DC.)

The value Q in Equation 4 can be estimated for fully penetrating wells in an unconfined, zero gradient aquifer, where drawdown is small compared to the depth of the aquifer, as (Bear, 1979):

$$Q = \pi K \left[\frac{H_o^2 - H_w^2}{\ln\left(\dfrac{r_o}{r_w}\right)} \right] \tag{5}$$

where Q is the pumping rate (cfs), K is the hydraulic conductivity of the aquifer (ft/s), H_o is the hydraulic head at the limit of the zone of influence, H_w is the hydraulic head at the well (ft), r_o is the distance to the limit of the zone of influence (ft), and r_w is the radius of the pumping well screen (ft). The relationship

SINGLE-WELL CAPTURE-ZONE TYPE CURVES

(a) Q/BU = 2000 m
1600
1200
800
400
Regional Flow

Meters

1000.
500.
0.
-500.
-1000.

-500.   0.   500. 1000. 1500. 2000. 2500.

DOUBLE-WELL CAPTURE-ZONE TYPE CURVES

(b) Q/BU = 2000 m
1600
1200
800
400
Regional Flow

1500.
1000.
500.
0.
-500.
-1000.
-1500.

-500.   0.   500. 1000. 1500. 2000. 2500. 3000.

THREE-WELL CAPTURE-ZONE TYPE CURVES

(c) Q/BU = 1000 m
800
600
400
200
Regional Flow

Meters

1500.
1000.
500.
0.
-500.
-1000.
-1500.

-500.   0.   500. 1000. 1500. 2000. 2500. 3000.

FOUR-WELL CAPTURE-ZONE TYPE CURVES

(d) Q/BU = 1000 m
800
600
400
200
Regional Flow

1500.
1000.
500.
0.
-500.
-1000.
-1500.

-500.   0.   500. 1000. 1500. 2000. 2500. 3000.

FIGURE 25.11 Capture zone of extraction well network. (From USEPA. 1994a. *Monitoring the Performance of Groundwater Pump and Treat Systems.* EPA/600/R-94/123. Washington, DC.)

between Q, H_o, H_w, r_o, and r_w is best estimated based on the results of a pump test, performed as described in Chapters 8 and 9.

When evaluating the combined zone of influence of a network of extraction wells, the location of the stagnation point and the overlap of adjacent zones of influence should be carefully examined. For aquifers where the hydraulic gradient is steep, the path taken by water flowing toward the network of wells could pass between and beyond the network of wells. To minimize this problem, zones of influence should be overlapped significantly (e.g., wells should be spaced at a distance of approximately $Q/(\pi Ti)$ for two extraction wells and slightly greater for increasing numbers of extraction wells [USEPA, 1994a]). The capture zone of a group of extraction wells is illustrated in Figure 25.11. To perform a more accurate estimate of the dimensions of the capture zones or to more accurately evaluate the possibility of groundwater flow past the induced barrier, a particle tracking groundwater model can be used, as described in Chapters 14 and 15.

25.3.4 Other Options for Hydraulic Containment

Numerous alternative approaches to hydraulic containment have been developed in the past several years. Some promising approaches are described below.

- *Horizontal drains* (see USEPA, 1985, and Cohen and Miller, 1983) induce a barrier to groundwater flow in a manner similar to trenches. With recent advances in horizontal "directional" drilling technologies, the accuracy of placement for horizontal wells has improved significantly. Horizontal

drains may be particularly useful in locations where access for excavation of a trench is not available (e.g., beneath a building or roadway).

- *Geomembranes* (see Rumer and Mitchell, 1995) can be incorporated into vertical barriers (as shown in Figure 25.12) to provide an extremely low-permeability barrier to groundwater flow. Geomembranes offer excellent compatibility with subsurface contaminants and have highly uniform properties. However, for the barrier to perform properly, joints between adjacent geomembrane panels must be carefully made to prevent leaks at these locations. Some commonly used installation methods for geomembrane barrier walls are summarized in Table 25.11.

- *Wellpoints* (see Driscoll, 1986) are discrete groundwater extraction devices that are typically jetted or driven into the ground. Wellpoints are useful in aquifer formations where a filter pack is not needed and where groundwater is at a relatively shallow depth. Typically, groundwater   is extracted by suction or by using air-lift techniques from wellpoints; pumps are typically not used in wellpoint systems.

A good discussion of other innovative remediation techniques is presented in USEPA (1994b), and a market evaluation of innovative technologies is presented in USEPA (1993a).
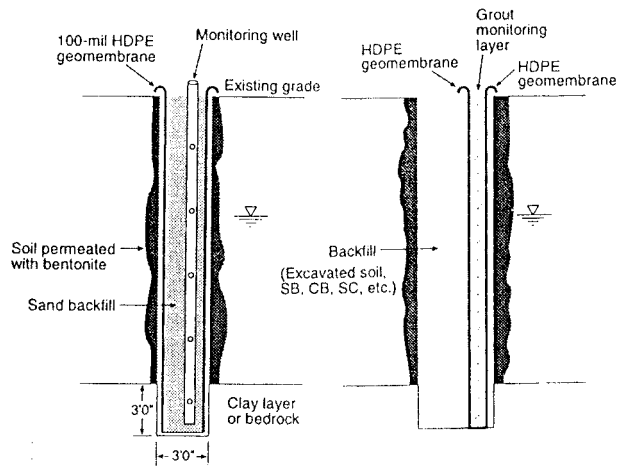
FIGURE 25.12 Geomembrane lined cutoff trench. (From Rumer, R.R. and Mitchell, J.K. 1995. *Assessment of Barrier Containment Technologies — A Comprehensive Treatment for Environmental Remediation Applications.* National Technical Information Services, Springfield, VA.)

25.4 Design of Groundwater Extraction Systems

25.4.1 Introduction

In Section 25.3, the design of groundwater extraction features was addressed, particularly groundwater extraction wells and trenches. In this section, the design of the other features of groundwater extraction systems (e.g., pumps, controls, etc.) is presented. Groundwater extraction systems are a combination of both subsurface features (i.e., wells or trenches, which provide access to the contaminated groundwater) and above-ground features (which are used to regulate and monitor the extraction process). In Sections 25.4.2 and 25.4.3, features of groundwater extraction well and trench systems are described. Then, in Section 25.4.4, methods for estimating the volume of water and time required to remove a plume of contaminants from the subsurface by extraction are presented and critiqued. Finally, innovative groundwater extraction approaches are described in Section 25.4.5. Note that groundwater extraction is only one step of the two-step extraction/treatment process; treatment processes are described in Section 25.5.

TABLE 25.11 Installation Methods for Geomembrane Barrier Walls

Method No.	Method or Technique	Geomembrane Configuration	Trench Support	Typ. Trench Width mm (in.)	Typ. Trench Depth m (ft.)	Typ. Backfill Type
1	Trenching machine	Continuous	None	300–600 (12–24)	1.5–4.5 (5–15)	Sand or native soil
2	Vibrated insertion plate	Panels	None	100–15 (4–6)	1.5–6.0 (5–20)	Native soil
3	Slurry supported	Panels	Slurry	600–900 (24–36)	No limit, except for trench stability	SB, SC, CB, SCB, sand or native soil
4	Segmented trench box	Panels or continuous	None	900–1200 (36–48)	3.0–9.0 (10–30)	Sand or native soil
5	Vibrating beam	Panels	Slurry	150–220 (6–9)	No limit	SB, SC, CB, SCB slurry

Key: SB = soil bentonite; SC = soil cement; CB = cement bentonite; SCB = soil cement bentonite

From Rumer, R.R. and Mitchell, J.K. 1995. *Assessment of Barrier Containment Technologies — A Comprehensive Treatment for Environmental Remediation Applications*. National Technical Information Services, Springfield, VA.

25.4.2 Extraction Well Systems

Groundwater extraction wells are typically more cost-effective for removing groundwater than trenches if the depth of contamination is great or if construction of the trench to the required depth in the aquifer is complex. In addition to extraction wells, extraction well systems contain the pumps, power sources, and controls that are needed to actually remove the contaminated groundwater from the aquifer and transmit it to the treatment or disposal system. A schematic diagram of a typical groundwater extraction well system is presented in Figure 25.13. The typical features of the system are briefly described below.

- *Well.* The well provides access to the groundwater in the aquifer. The well may be designed using techniques as described in Section 24.3.2.1, except the design of an extraction well should be different from the design of a groundwater monitoring well to improve the long-term performance of the extraction well and minimize the need for maintenance. Design of groundwater extraction well components (i.e., wellscreen, filter pack, etc.) is addressed by Driscoll (1986).

- *Pump and Motor.* The pump raises the groundwater from the aquifer to the above-ground wellhead assembly. The pump and motor should be selected according to the specific operating conditions that will exist during groundwater extraction; for example, some pumps are very well suited for operation under low-flow conditions, and other pumps are better suited for operation under high-flow conditions. In Table 25.12, information is presented for several types of pumps that are commonly used in groundwater extraction wells.

- *Wellhead.* The wellhead is the above-ground assembly that contains the features used to transfer groundwater from the extraction well to the transmission piping. Typical components of the wellhead assembly include valves, fittings, flow meters, pressure gauges, and sampling ports. The wellhead can be designed to be either in an above-ground shed or in a below-ground vault, depending on the use of the area surrounding the well.

- *Transmission Piping.* The transmission piping is used to transmit the extracted groundwater from the wellhead to the treatment or discharge location. The following should be considered when designing the piping: (1) the required quantity of flow (as described by Corbitt [1990]); (2) damage due to overburden stresses (using analysis techniques described by Merritt (1989)), including stress caused by vehicles if appropriate); (3) resistance of the piping to degradation by the contaminants in groundwater; and (4) constructibility of the piping system (note that difficult-to-construct systems are less likely to be well constructed).
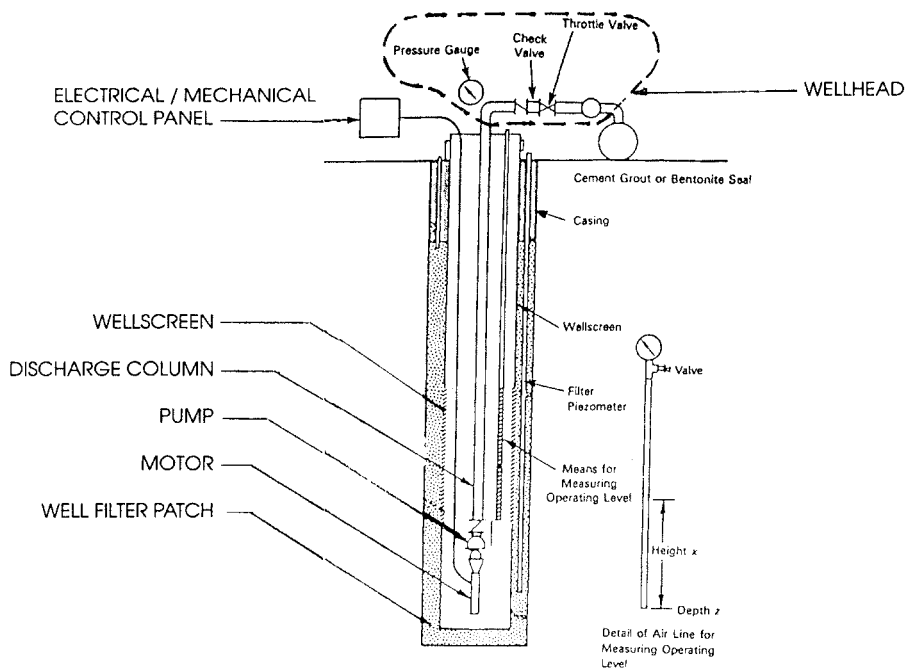
FIGURE 25.13 Typical groundwater extraction system. (From USEPA. 1995a. *Ground-Water and Leachate Treatment Systems.* EPA/625/R-94/005. Washington, DC.)

- *Power Source.* A source of power is typically needed for operation of the extraction system components. Power can be provided by electricity, compressed air, solar energy, or by natural gas. Local building codes should be checked to identify any specific requirements for design of electrical components of the system.
- *Electrical/Mechanical Control Panel.* The control panel contains the electrical controls (e.g., alarms, programmable logic controls, etc.) for the extraction system. Functions that are typically controlled at a panel include pump cycle time, pump temperature, monitoring for alarm conditions, and transmittal of the data collected at the wellhead (e.g., groundwater elevation, pump flow rate, etc.) to a central processing unit.
- *Discharge Point.* The discharge point is the feature at which the extracted groundwater is transferred to either the treatment system or the ultimate discharge point.

It is recommended that the following factors should be considered when designing a groundwater extraction system (Driscoll, 1986):

- Because of the high quantity of groundwater that flows through an extraction well, very durable materials (such as stainless steel wellscreen) should be specified, and the open area of the well should be maximized, but carefully balanced against the strength requirements of the wellscreen.
- The chemical compatibility of each component of the extraction, transmission, and pump systems with the contaminated groundwater should be carefully examined before final selection of the extraction system components and materials.
- The health and safety of the well installation crew should be considered when developing the plan for installation of the groundwater extraction wells (the well installer should be required to prepare a health and safety plan for well installation activities.
- The impact of naturally occurring compounds in extracted groundwater on the durability and performance of the extraction system components should be evaluated, including the effects of low and high pH (a pH significantly lower than 7 indicates acidic, possibly corrosive conditions),

TABLE 25.12 Groundwater Removal Pump Information

Type of Pump	Practical Suction Lift* ft (m)	Usual Well-Pumping Depth ft (m)	Usual Pressure Heads ft (m)	Advantages	Disadvantages
1. Positive displacement	22–25 (6.7–7.6)	22–25 (6.7–7.6)	100–2,200 (30.5–670. 6)	1. Pumps water containing sand and silt 2. Especially adapted to low capacity and high lifts	1. Subject to vibration and noise 2. Maintenance cost may be high
2. Ejector		15–20 (4.57–6.10) below ejector	80–150 (24.4–45.7)	1. High capacity at low heads 2. Simple operation	1. Capacity reduces as lift increases 2. Air in suction or return line will stop pumping
a. Shallow well	15–20 (4.6–6.1) below ejector				
b. Deep well	15–20 (4.6–6.1) below ejector	25–120 (7.62–36.58) below ejector	80–150 (24.4–45.7)	1. Same as shallow well jet 2. Well straightness not critical	1. Same as shallow well jet 2. Lower efficiency, especially at greater lifts
3. Centrifugal Shallow well		10–20 (3 –6.1)	100–150 (30.5–45.7)	1. Pumps water containing sand and silt	1. Loses prime easily
a. Straight centrifugal (single stage)	20 ft (6.10) max.			2. Usually reliable and good service life	2. Efficiency depends on operating under design heads and speed
b. Regenerative vane turbine type (single impeller)	28 (8.5) max.	28 (8.5)	100–200 (30.48–45.72)	1. Same as straight centrifugal except not suitable for pumping water containing sand or silt 2. Self–priming	1. Same as straight centrifugal except maintains priming easily
Deep well		50–300 (15.2–91.4)	100–800 (30.5–243.8)	1. Same as shallow well turbine 2. All electrical components are accessible, above ground	1. Efficiency depends on operating under design head and speed 2. Requires large, straight well
a. Vertical line shaft turbine (multistage)	N/A (impellers submerged)				
b. Submersible turbine (multistage)	N/A (pump and motor submerged)	50–400 (15.2–121.9)	50–400 (15.2–121.9)	1. Same as shallow well turbine 2. Short pump shaft to motor 3. Quiet operation	1. Repair requires pulling from well 2. Sealing of electrical equipment from water vapor critical 3. Abrasion from sand

* Practical suction lift at sea level. Reduce lift 1 ft for each 1000 ft above sea level.

From Corbitt, R.A., editor. 1990. *Standard Handbook of Environmental Engineering*. McGraw–Hill, New York. With permission.

dissolved oxygen (greater than 2µg/l indicates possibly corrosive conditions), iron (as little as 0.25 mg/L can cause fouling of a wellscreen or groundwater treatment system), hydrogen sulfide (as little as 1 mg/L can be corrosive), dissolved solids (greater than 1,000 mg/L can cause electrical corrosion), carbon dioxide (greater than 50 mg/L can produce carbonic acid and result in corrosive conditions), and encrusting conditions (which are typically caused by conditions of pH > 7.5, hardness (as $CaCO_3$) > 300 mg/L, iron > 0.5 mg/L, or manganese > 0.2 mg/L).

25.4.3 Extraction Trench Systems

Groundwater extraction trenches will likely be more cost-effective than extraction wells if the depth of contamination is relatively shallow and/or when it is relatively simple to install a trench to the required depth. Before beginning detailed design of the extraction trench system, the groundwater professional should have already identified the depth of the trench and the location where the trench must be installed (using techniques described in Section 25.3.3). In addition to an extraction trench, an extraction trench system consists of the pumps, power sources, and controls that are needed to actually remove the contaminated groundwater from the trench. A schematic diagram of a typical groundwater extraction trench system is presented in Figure 25.7. The typical features of the system are briefly described below.

- *Trench.* Trenches used for extraction of groundwater may be installed using any of a number of methods (e.g., backhoe, slurry trench backfilled with aggregate, biodegradable slurry trenches, trench machines that install the pipe in a single operation, etc. Backfill materials used in most types of groundwater extraction trenches include collection piping and aggregate; these materials should be designed to withstand degradation by the chemicals in groundwater that they will be exposed to and to resist damage caused by the stresses that they will be exposed to. Design of plastic pipes (which are typically used in extraction trenches because of their resistance to chemical degradation) is addressed by Merritt (1989).
- *Pump.* The pump should be designed using the criteria presented in Section 25.4.2 for extraction well pumps. Because groundwater extraction trenches are typically several feet wide, large-diameter sumps can be placed in the trench, which can accommodate relatively large pumps. Large pumps can be designed to be resistant to clogging by soil particles, typically have a much longer service life, and are typically more reliable and less costly to maintain than smaller pumps. A sump should be located at each low point in the trench.
- *Transmission Piping, Power Source, Control Panel, Discharge Point.* These features may be designed using the recommendations presented for extraction well systems in Section 25.4.2.

25.4.4 Time Required to Extract a Plume of Contaminated Groundwater

Over the past two decades, one of the primary purposes of extracting contaminated groundwater from an aquifer was to restore the aquifer to its useful purpose by actually removing the contaminated groundwater from the aquifer. In this section, the concept of restoring an aquifer to its useful purpose and the reasonableness of restoration as a cleanup goal are addressed. Aquifer restoration is based on the concept that the quantity of groundwater that must be removed to restore the aquifer can be calculated based on the physical properties of the aquifer. If there were no continuing source of contamination and if contaminated groundwater flowed uniformly unimpeded to the extraction wells or trench, then the time required to "restore" the aquifer could be estimated simply by dividing: (1) the distance from the farthest part of the plume of contaminated groundwater to the extraction wells or trench (ft) by (2) the average linear groundwater flow velocity (ft/day). The volume of groundwater pumped under this scenario would be equal to the cleanup time (as calculated above) times the combined pumping rate of all of the groundwater extraction features.

Experience has shown (NRC, 1994) that, in nearly all cases, other factors affect the time and the quantity of groundwater that must be extracted to remediate a contaminated aquifer. Some of the more significant of these other factors are sorption, hydrolysis, cosolvation/ionization, volatization/dissolution,

complexation, oxidation/reduction, and biodegradation, each of which may increase or decrease the time required for remediation. A concise discussion of these processes and their impact on remediating a contaminated aquifer is presented in USEPA (1989b). Of these factors, sorption typically has the greatest impact on remediation, although other factors could have a greater impact depending on the specific site features and the particular contaminants that are present in groundwater. Sorption is described in Chapter 14. In general, sorption is the dual process of: (1) adsorption of dissolved contaminants in groundwater to the surface of soil particles, and (2) subsequent desorption of the contaminants from the soil particles when equilibrium conditions in the aquifer change. As equilibrium conditions change in the aquifer, "partitioning" of contaminants between the dissolved phase and the solid (i.e., adsorbed) phase changes, causing contaminants to either adsorb to soil particles from groundwater or to desorb from soil particles to groundwater. Contaminants become adsorbed to soil particles as a result of electrochemical forces (i.e., either ion exchange reactions or van der Waals forces) between the soil particles and the dissolved contaminants. Partitioning coefficients for some common organic contaminants are presented in Table 25.13. Ions that have a high valence charge (e.g., Cr^{+6}, Al^{+3}) and molecules that are either large (e.g., chlorine-based compounds, such as TCE) or long-chained (e.g., polynuclear hydrocarbons) tend to adsorb strongly to soil particles. Because desorbed contaminants can adversely impact groundwater quality, aquifer materials having adsorbed contaminants act as a continuing "source" of groundwater contamination, allowing uncontaminated groundwater that flows past the adsorbed contaminants to become contaminated until all of the adsorbed material has been desorbed.

TABLE 25.13 Partitioning Characteristics for Selected Chemicals

Chemical	Water Solubility (mg/L)	Vapor Pressure (atm)	Henry's Law Constant, K_H (atm-m³/mol)	Organic Carbon Partition Coefficient, K_{OC} (L/kg)
Aldrin	1.80×10^{-1}	7.89×10^{-9}	1.60×10^{-5}	9.60×10^{4}
Benzene	1.75×10^{3}	1.25×10^{-1}	5.59×10^{-3}	8.30×10^{1}
Chloroethane	5.74×10^{3}	1.32×10^{0}	6.15×10^{-4}	1.70×10^{1}
DDT	5.00×10^{-3}	7.24×10^{-9}	5.13×10^{-4}	2.43×10^{5}
Tetrachloroethene (PERC)	1.50×10^{2}	2.30×10^{-2}	2.59×10^{-2}	3.64×10^{2}
Trichloroethene (TCE)	1.10×10^{3}	7.60×10^{-2}	9.10×10^{-3}	1.26×10^{2}
Vinyl chloride	2.67×10^{3}	3.50×10^{0}	8.19×10^{-2}	5.70×10^{1}

From USEPA. 1996a. *Superfund Chemical Data Matrix*. EPA/540-R-94-009. Washington, DC.

There are numerous models available for evaluating the time required to remediate an aquifer for which the effects of sorption are significant. The models range from simple (e.g., the batch-flush model using simple hydraulic "pore-volume" approaches, or advective-dispersive approaches) to complex (e.g., finite-element models incorporating data from bench-scale tests, pilot tests, etc.). A simple and reasonably accurate analytical method for estimating the time and quantity of extracted groundwater required to remediate an aquifer is the batch-flush model. The batch-flush model is based on the ideal assumptions that: (1) upon becoming contaminated, groundwater is perfectly mixed within the aquifer, and (2) sorption is linear, reversible, and rapid. The analysis does not account for nonhomogeneities or anisotropies within the aquifer, dispersion, or any other mechanism of contaminant transport described in Chapters 14 and 15. The batch-flush approach is based on the explicit, finite-difference approximation of the one-dimensional differential equation for groundwater transport, which is:

$$V_o \frac{dC}{dt} + \rho \frac{V_o}{n} \frac{dS}{dt} = -QC \qquad (6)$$

where C is the concentration of the contaminant in water (mg/L), S is the sorbed-phase concentration of the contaminant (mg/L), V_o is the volume of contaminated water (i.e., the pore-volume water) (m³), ρ is the bulk density of the porous media (g/m³), n is the effective porosity (dimensionless), Q is the

volumetric groundwater extraction rate (m³/sec), and t is time (sec). Because sorption is linear, reversible, and rapid, S can be substituted with $K_d C$, where K_d is the soil-water partitioning coefficient. Finally, substituting and rearranging gives the equations:

$$T_t = R\left[-\ln\frac{C_t}{C_o}\right] \tag{7}$$

and

$$t_t = \frac{V_o}{Q}T_t \tag{8}$$

where C_o is the initial concentration of contaminant in the aquifer, T_t is the number of pore volumes that must be extracted to "restore" the aquifer to the desired cleanup concentration (C_t), and t_t is the time required to achieve C_t. In Equation 9, the term R is known as the retardation factor and is defined as:

$$R = 1 + \frac{\rho}{n}K_d \tag{9}$$

where ρ, n, and K_d are as previously defined. Values of K_d may be measured in the laboratory; however, because the value of K_d is primarily a function of sorption to organic material, a useful approximation of K_d can be found using the equation:

$$K_d = K_{oc} \times f_{oc} \tag{10}$$

where K_{oc} is a proportionality constant (which can be estimated based on the octanol-water partitioning coefficient, $K_{ow,}$ of the contaminant) and f_{oc} is the fraction of organic carbon in the aquifer soils. Estimates of K_{oc} for several common organic contaminants are provided in Table 25.13. Because desorption is a function of the ratio of the concentration of sorbed contaminant to the concentration of dissolved contaminant, sorbed contaminants could be expected to desorb indefinitely, which (depending on the desired cleanup concentration, C_t) could make it impossible to restore groundwater quality in a reasonable period of time, if at all. For the same reason, groundwater extraction systems that are operated at a constant extraction rate may not remove an appreciable amount of contaminants in the later stages of
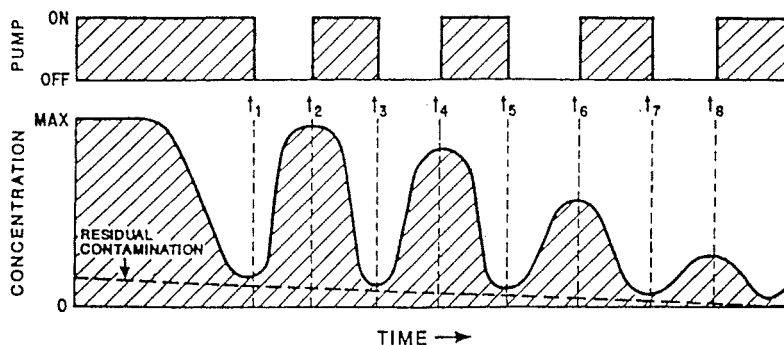
FIGURE 25.14 Use of a pulsed pumping scheme to maximize reduction of contaminant concentration. (From Keely, J.F. 1989. *Performance Evaluations of Pump-and-Treat Remediations*, EPA Superfund Groundwater Issue, EPA/540/4-89/005. Washington, DC.)
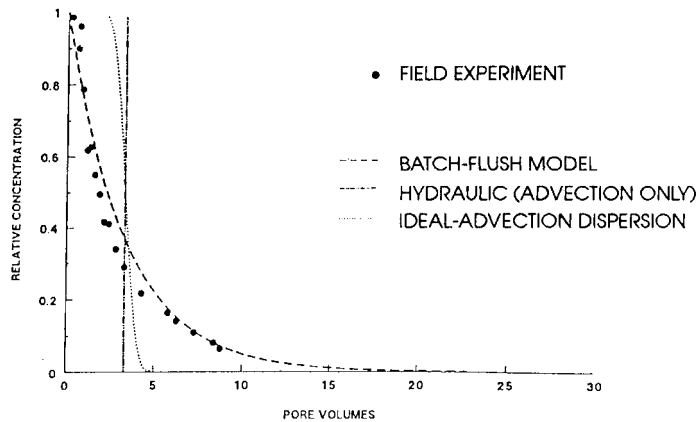
FIGURE 25.15 Elution curve for diethylether. (From Brusseau, M. 1996. *Evaluation of Simple Methods of Estimating Contaminant Removal by Flushing Groundwater,* p. 19-22. With permission.)

TABLE 25.14 Impact of Cleanup Goal on Cost of a Conventional Pump-and-Treat System

Percent Removal Required	Calculated Years to Achieve Goal	Present Worth
80	15	$ 2,800,000
90	21	$ 3,250,000
99	42	$ 4,750,000
99.9	63	$ 5,600,000
99.99	84	$ 6,000,000

Adapted from National Research Council (NRC), Water Science and Technology Board. 1994. *Alternatives for Groundwater Cleanup.* Washington, D.C. With permission.

remediation; in such cases, pulsed pumping (as illustrated in Figure 25.14) approaches may optimize the remediation process by minimizing the total quantity of groundwater pumped and treated.

Batch flush models are useful for estimating cleanup times for simple aquifer systems in which sorption is nearly linear. However, the model typically underestimates cleanup times because it does not account for any of the following: (1) heterogeneities; (2) continuous releases from a source of contamination; (3) the presence of separate-phase contaminants; or (4) the effects of nonlinear sorption. These limitations also apply to models that are more complex than the batch flush model. Because of the effects of sorption, the cost to clean up an aquifer to a very low level of contamination can be significantly greater than the cost to achieve a slightly higher level of contamination; an example of this problem is shown in Figure 25.15 and in Table 25.14, which illustrates groundwater quality as a function of the number of pore volumes removed for a simple pump-and-treat groundwater remediation system in a homogeneous sand aquifer. Remediation cost can be expected to increase substantially as the retardation factor increases (as shown in Figure 25.16), and a substantial additional cost can be expected for achieving very high cleanup efficiencies (as shown in Figure 25.17 and Table 25.14. If it is found that the groundwater remediation goal cannot be achieved through extraction, then either a different groundwater remediation goal should be selected or a complementing remediation technique should be considered to enhance remediation of groundwater.

25.4.5 Other Groundwater Extraction Approaches

There are many groundwater extraction approaches available to the groundwater professional in addition to vertical wells and trenches, including both innovative approaches for removing groundwater and approaches that enhance the extraction of contaminants from the subsurface. Other groundwater extraction techniques include the use of horizontal wells and wellpoints. Some of the more promising innovative
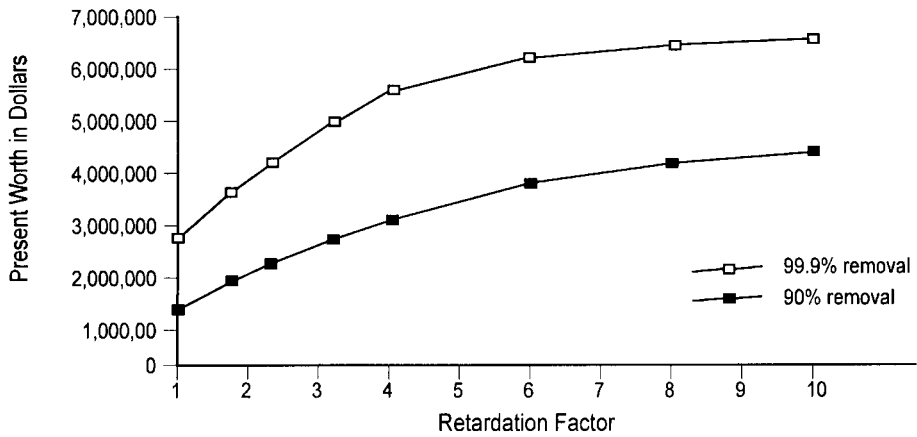
FIGURE 25.16 Cost of pump-and-treat system as a function of the retardation factor of the contaminant. (Adapted from National Research Council (NRC), Water Science and Technology Board. 1994. *Alternatives for Groundwater Cleanup*. Washington, D.C. With permission.)
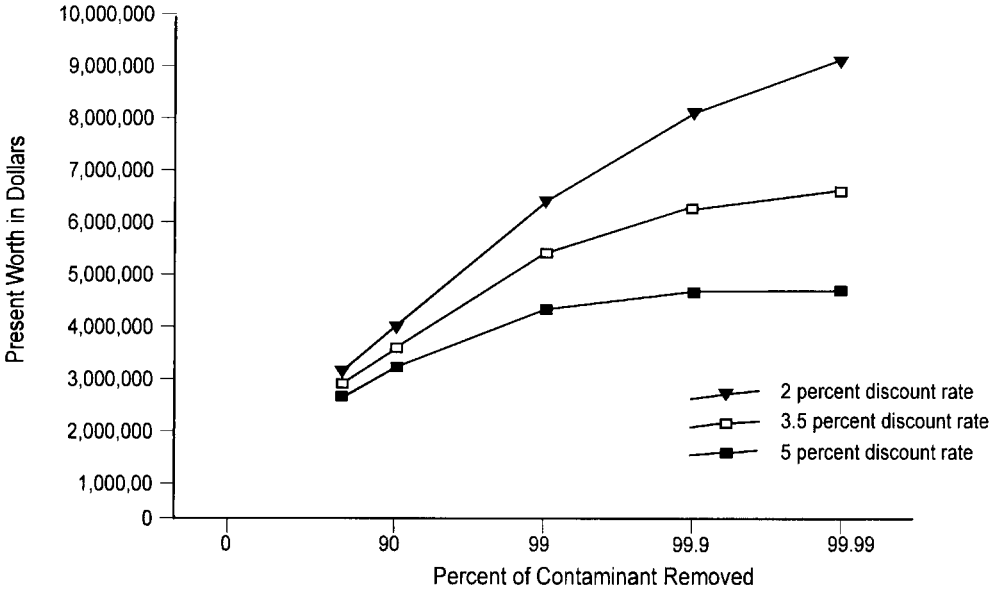
FIGURE 25.17 Cost of pump-and-treat system as a function of the cleanup goal. (Adapted from National Research Council (NRC), Water Science and Technology Board. 1994. *Alternatives for Groundwater Cleanup*. Washington, D.C. With permission.)

approaches to enhance extraction efficiency include hydraulic fracturing (i.e., fracturing subsurface media using water pressure or explosives to increase the permeability of the media and the groundwater extraction efficiency), flushing (i.e., reinjecting treated groundwater upgradient of the source area to increase the hydraulic gradient toward the extraction feature), surfactant flushing or chemical extraction (i.e., flushing using surfactants or other chemicals that are designed to increase the mobility or solubility of the contaminants in the aquifer), electroacoustical decontamination (i.e., use of electrokinetic and acoustical waves to reduce surface tension and viscosity of separate-phase contaminants, increasing their mobility and the opportunity to extract them), and steam extraction (i.e., injection of steam into the aquifer to either decrease the viscosity of a contaminant or to volatize contaminants). When evaluating
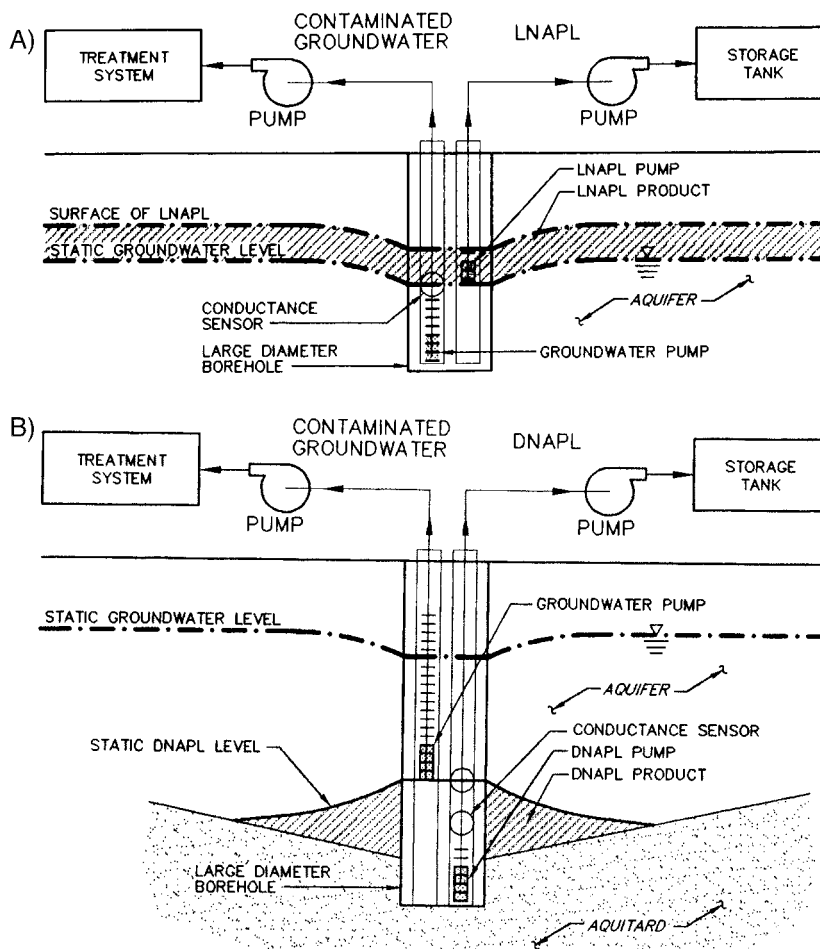
FIGURE 25.18 (A) LNAPL dual-extraction system; (B) DNAPL dual-extraction system.

reinjection techniques that involve the use of chemicals, the groundwater professional should be careful to evaluate the potential adverse impact of the chemicals, which may be toxic, on human health and the environment.

When evaluating groundwater extraction approaches, separate attention should be given to extraction of NAPLs. Because they represent a significant source of potential contamination, removal of NAPLs can significantly decrease the amount of time required to achieve a groundwater remediation goal using pump-and-treat techniques. Dual-phase pumping approaches can be effective for removing both LNAPL and DNAPL products; examples of dual-extraction systems for LNAPL and DNAPL products are illustrated in Figure 25.18.

25.5 Treatment of Contaminated Groundwater

25.5.1 Introduction

There are many techniques available for treatment of contaminated groundwater. Over the past 10 years, an extremely large number of groundwater treatment techniques has been developed; these techniques are described by the U.S. EPA and the U.S. Air Force (1993) and in USEPA (1995a). In this section, brief descriptions of the most widely used standard treatment techniques and the most promising innovative

treatment techniques are presented. Candidate treatment technologies that are addressed in this section are presented in Table 25.15. For the purposes of this section, these technologies are categorized as either *in situ* or *ex situ* techniques. *In situ* techniques are generally intended to either render a contaminant nontoxic through treatment (e.g., bioremediation) or to enhance extraction of the contaminant from the aquifer (e.g., air sparging). *Ex situ* techniques can be used only to treat groundwater that has been extracted from the aquifer. *In situ* treatment techniques are described in Section 25.5.2, and *ex situ* treatment techniques are described in Section 25.5.3.

TABLE 25.15 Summary of Candidate Treatment Technologies

In Situ	Process Description
Bioremediation	Biological degradation of contaminants using naturally occurring microbes in soil
Soil vapor extraction	Volatization of contaminants that are present in the vadose zone
Air sparging	Volatization of contaminants in the saturated zone
Permeable reaction barriers	Physical or chemical treatment in a trench
Vacuum vapor extraction	Volatization, within a well, of contaminants from saturated zone
Density driven convection	Enhanced bioremediation using single-well driven convection system in aquifer.

Ex Situ	Process Description
Bioreactor	Biological degradation of contaminants (activated sludge, fixed-film biological reactor, biophysical treatment)
Slurry-phase biological treatment	Variation on bioreactor in which contaminants are treated in a slurry form
Air stripping	Volatization of contaminants
Carbon adsorption	Adsorption of contaminants to activated carbon
Ion exchange	Exchange-type attachment of contaminants to ion-exchange resin
Alkaline precipitation	Alteration of water quality (usually pH adjustment) to conditions in which concentration exceeds the compound's solubility limit, causing precipitation
Membrane	Separation of solids from water using membranes (reverse osmosis, ultrafiltration)
Wetlands treatment	Uptake of contaminants by wetland features
Electrokinetic decontamination	Desorption of contaminants by "acidic front" of groundwater caused by hydrolysis of the groundwater

25.5.2 *In Situ* Treatment

25.5.2.1 Introduction

In situ treatment techniques are described in this section in three categories. First, techniques that incorporate biological processes (e.g., bioventing, bioremediation) are described in Section 25.5.2.2. Next, techniques that incorporate volatization processes (e.g., air sparging, soil vapor extraction) are described in Section 25.5.2.3. Then, techniques that incorporate a chemical or physical process (e.g., permeable reaction barriers) are described in Section 25.5.2.4. Finally, some promising innovative *in situ* treatment approaches are described in Section 25.5.2.5. In some cases, *in situ* treatment technologies may offer a viable alternative to more conventional remediation approaches (e.g., pump-and-treat) and may improve the chances of achieving the ultimate remediation goal (in some cases, even aquifer restoration).

25.5.2.2 Biological Remediation

Biological remediation (more commonly referred to as bioremediation) involves the promotion of conditions in an aquifer that are conducive to the growth of microbes that can degrade contaminants. In this section, a brief overview of biological remediation is presented; a more detailed discussion of bioremediation is presented in Chapter 26 (i.e., Bioremediation). Numerous types of microbes are naturally present in aquifers. Each microbe tends to degrade certain types of chemicals; the microbes that are most likely to be present in the subsurface are those that have evolved on a diet of the most abundant, naturally occurring organic material in the aquifer. Contaminants are degraded by the process of microbial metabolism, in which the microbe uses the contaminant as an electron donor to complete

heterotrophic microbial respiration. The by-products of complete biodegradation are carbon dioxide, water, and biomass. Biodegradation of complex compounds may occur in a series of several steps, in which a contaminant is degraded into several intermediate products before it is completely degraded; for example, tetrachloroethylene biodegrades to trichloroethylene, which biodegrades to dichloroethylene, which biodegrades to vinyl chloride, which biodegrades (albeit much more slowly than the previous degradation steps) into carbon dioxide and water.

Bioremediation is usually most successful in aquifers that are contaminated with chemicals that are easily metabolized by the indigenous microbes; in general, the more closely a contaminant resembles a naturally occurring compound, the more likely it is that there exists a microorganism in the aquifer that is capable of biodegrading the contaminant. An illustration of the relative biodegradability of organic contaminants is shown in Figure 25.19. Note that inorganic compounds are not generally biodegradable (however, inorganic compounds may be bioaccumulated in microbes) and some inorganic compounds (e.g., metals) may actually be toxic to microbes. Also, note that the ambient conditions within the aquifer have a significant effect on the potential for successful bioremediation; a summary of the environmental factors that affect bioremediation and the range of optimum values for each factor is presented in Table 25.16.

The rate of biodegradation (and, generally, the consumption of any material by biological activity) is expressed by the rate equation defined below:

$$\ln\left(\frac{C_t}{C_o}\right) = -kt \tag{11}$$
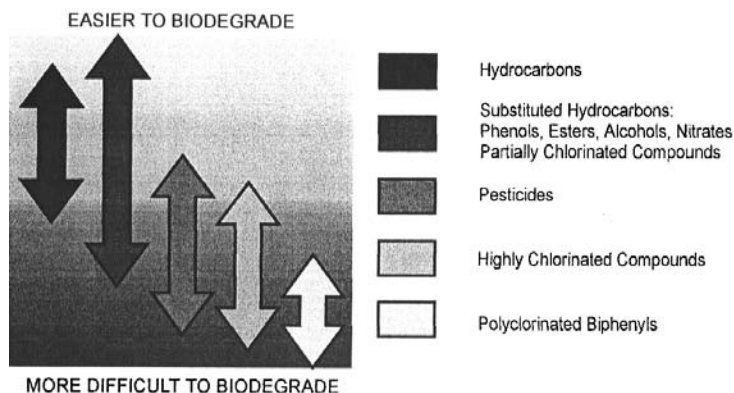
FIGURE 25.19 Relative biodegradability of organic compounds.

TABLE 25.16 Environmental Factors that Affect Biodegradability

Environmental Factor	Optimum Levels
Available soil water	25–85% of water holding capacity; –0.01 Mpa
Oxygen	Aerobic metabolism: greater than 0.2 mg/L dissolved oxygen, minimum air-filled pore space of 10% by volume
	Anaerobic metabolism: oxygen concentrations less than 1% by volume
Redox potential	Obligate facultative anaerobes: greater than 50 millivolts
	Obligate anaerobes: less than 50 millivolts
pH	pH values of 5.5–8.5
Nutrients	Sufficient nitrogen, phosphorus, and other nutrients so as to not limit microbial growth (Suggested C:N:P ratio of 120:10:1)
Temperature	15-45°C (Mesophiles)

From USEPA. July 1991. *Handbook — Ground Water, Volume II: Methodology.* EPA/625/6-90/016b. Washington, DC.

where k is the first-order reaction rate constant (sec^{-1}), C_t is the concentration of constituent remaining at time t (mg/L), C_o is the initial concentration of the constituent (mg/L), and t is time (sec). Typically, C_o is measured in an analytical laboratory and k is estimated based on the laboratory-measured rate of biodegradation of contaminants exposed to actual microbes from the aquifer. The time required for degradation of a particular plume of contamination can be estimated by solving Equation 11 for t.

The design effort for a bioremediation system is typically focused on identifying the stimuli (e.g., nutrients, oxygen, etc.) that will promote biodegradation of the contaminants that are present in groundwater and adsorbed to soils in the aquifer. To accelerate the rate of biodegradation, for example, the system could be designed to optimize the parameters in Table 25.16 that impact biological activity; for example, if oxygen content is low, then oxygen could be delivered to the aquifer through wells in the form of hydrogen peroxide or air. The impact of such inputs on the rate of biodegradation (including detrimental impacts, such as precipitation of dissolved iron caused by addition of hydrogen peroxide) are typically estimated based on the results of laboratory treatability tests.

25.5.2.3 Volatization Processes

Volatization is the transfer of a chemical from the liquid state to the gaseous state. Remediation techniques that employ volatization processes promote the change of contaminants from liquid phase (including dissolved) to the vapor phase. Vapor-phase contaminants are typically easier to remediate than liquid-phase contaminants because aquifer materials are significantly more permeable to vapors than to liquids, making vapors easier to remove from the aquifer (see discussion of intrinsic permeability in Chapter 1). Remediation technologies that employ volatization processes include both soil vapor extraction (which is applicable to contaminants adsorbed to soil in the vadose zone) and air sparging (which is applicable to both dissolved and adsorbed contaminants in the saturated zone). Both processes are based on the principle of molecular diffusion of contaminants from the dissolved or nonaqueous phase to the vapor phase, which is described by Fick's Law for diffusion. The principles of soil vapor extraction (SVE) and air sparging (AS) are described below.

Soil Vapor Extraction. An example of an SVE system is illustrated in Figure 25.20. The concentration of a particular contaminant in the extracted vapor can be estimated using Fick's Law as follows:

$$C_{est} = \sum_i \frac{x_i P_i^v M_w, i}{RT} \tag{12}$$

where C_{est} is the estimated concentration of contaminant in the vapor (mg/L), x_i is the mole fraction of component i in the liquid-phase contaminants, P_i^v is the vapor pressure of component i at temperature T (atm), M_w is the molecular weight of component i (mg/mole), R is the gas constant (i.e., 0.0821 l-atm/mole °K), and T is the absolute temperature of the contaminant residual (K). Note that contaminant concentrations will decline with time as the concentration of the contaminant adsorbed to the subsurface soils decreases; therefore, the contaminant vapor concentration will change over time. SVE is potentially suitable for contaminants that have a vapor pressure conducive to volatization (usually greater than about 1×10^{-4} atm at the average subsurface temperature); a summary of vapor pressures for some typical contaminants is presented in Table 25.13. Using the value of C_{est}, the removal rate can be estimated by multiplying C_{est} by the expected air flow rate Q. The achievable air flow rate can be estimated using the following equation (Johnson, et al., 1990):

$$\frac{Q}{H} = \pi \frac{k}{\mu} P_w \frac{\left[1 - \left(P_{atm}/P_w\right)^2\right]}{\ln\left(R_w/R_l\right)} \tag{13}$$

where Q/H is the flow rate per unit thickness of wellscreen (cm^3/sec), k is the soil permeability to air flow (cm^2), μ is the viscosity of air (1.8×10^{-4} g/cm-s, or 0.018 cp), P_w is the absolute pressure at the extraction well (atm), P_{atm} the absolute atmospheric ambient pressure (i.e., 1.01×10^6 g/cm-s^2, or 1 atm),
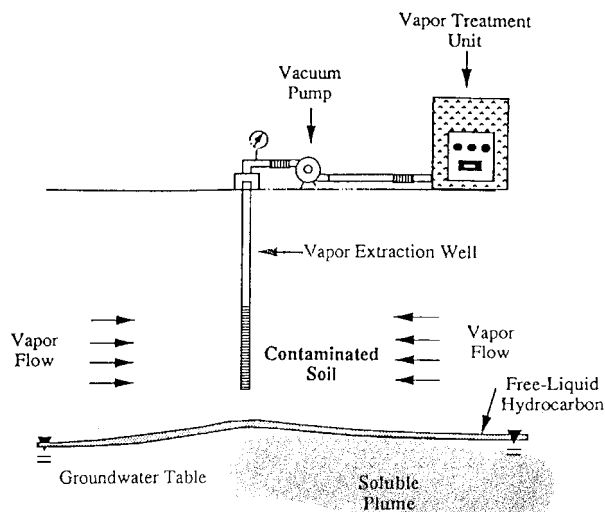
FIGURE 25.20 "Basic" soil vapor extraction system. (From Johnson, P.C., Stanley, C.C., Kemblowski, M.W., Byers, D.L., and Colthart, J.D. 1990. A practical approach to the design, operation, and monitoring of in-situ soil-venting systems. *Groundwater Monitoring Review*, Spring 1990, 159-178. With permission.)

R_W is the radius of the extraction well (cm), and R_I is the radius of influence of the extraction well (cm). The reasonableness of the expected gas flow rates for the permeability of the site soils may be confirmed using Figure 25.21. In some cases, there may be a significant quantity of residual contamination adsorbed to soil particles located a short depth below the water table as a result of changes in groundwater elevation since contaminants were released. In this case, it may be efficient to implement a "dual extraction" (i.e., air and groundwater) system similar to the dual extraction system illustrated in Figure 25.18A for groundwater and LNAPLs. By depressing the groundwater table below the residual contamination, the residual contamination can be extracted using SVE techniques, which are much more efficient than groundwater extraction techniques for removing adsorbed or residual contamination.

Air Sparging/Bioventing. An illustration of an air sparging (AS) system is presented in Figure 25.22. Air sparging involves introducing air into the aquifer and allowing the air to rise through contaminated groundwater to the ground surface. Because sparging introduces oxygen into the subsurface, it stimulates aerobic biodegradation; therefore, the technique is sometimes also referred to as "bioventing" to convey its value to the bioremediation processes. In general, the physics of air movement in the saturated zone are currently not well understood, which limits the predictability of air sparge system performance; in fact, the decisions regarding sparge well location and spacing are usually based on experience rather than calculation. Still, based on preliminary studies (USEPA, 1996) and because it has a relatively low cost, air sparging can be a cost-effective approach to groundwater remediation. As presently understood, air discharged to groundwater by a sparge well typically flows upward to the vadose zone through air channels (as illustrated in Figure 25.22) that are established at first operation of the system. These air channels directly impact only a very small quantity of groundwater and soil. Contaminants partition between the sparge air, groundwater, and soils; the portion that partitions to vapor in the sparge air rises with the sparge air to the ground surface. For air sparging to be effective, adsorbed contaminant mass must be transported by diffusion from the surface of soils through the groundwater to the air-water interface of the air channels, which is a slow process. To maximize the effectiveness of this technique, it is recommended that AS systems be used in conjunction with SVE systems (which capture the sparged air after it rises from groundwater into the vadose zone), as shown in Figure 25.22, to prevent offsite migration of contaminated sparge air.

The minimum air pressure required to operate a sparge system is provided by the equation (Johnson, et al., 1993):

Predicted steady-state flow rates figure with axes. Labels include:

R_w = 5.1 cm (2")
R_1 = 12 m (40')

P_w = 0.40 atm = 20.3 ft H₂O
P_w = 0.60 atm = 13.6 ft H₂O
P_w = 0.30 atm = 6.8 ft H₂O
P_w = 0.90 atm = 3.4 ft H₂O
P_w = 0.95 atm = 1.7 ft H₂O

(m³/m-min)

Vapor Flowrate (scfm/ft)

clayey sands   fine sands   medium sands   coarse sands

Soil Permeabilty (darcy)

{ft H₂O} denote vacuums expressed as equivalent water column heights
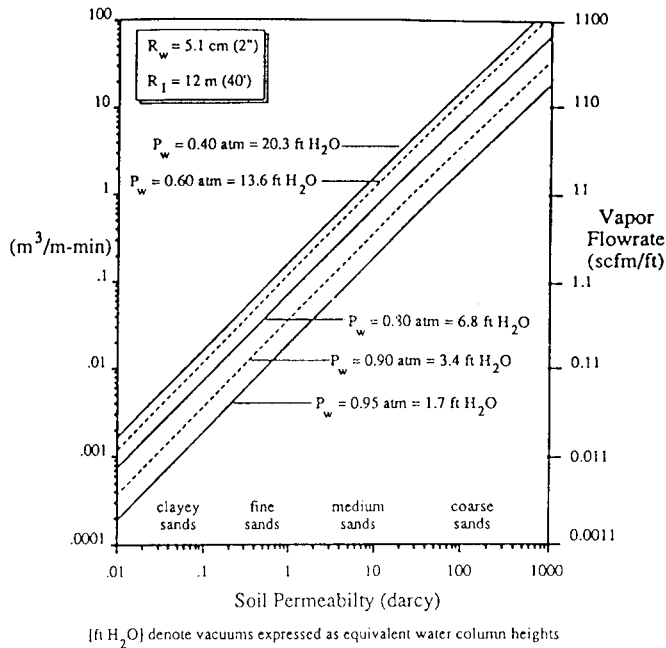
FIGURE 25.21 Predicted steady-state flow rates (per unit wellscreen thickness) for a range of soil permeabilities and applied vacuums (P_w). (From Johnson, P.C., Stanley, C.C., Kemblowski, M.W., Byers, D.L., and Colthart, J.D. 1990. A practical approach to the design, operation, and monitoring of in-situ soil-venting systems. *Groundwater Monitoring Review*, Spring 1990, 159-178. With permission.)

Air Sparging in Stratified Sands

Air Injection System — Off-gas Treatment System — Blower or Vacuum Pump — Ground Level

Vadose Zone — Vapor Extraction Wells — Mounded Water Table — Contaminated Zone — Silt Lenses — Air Flow Channels — Saturated Zone — Uniform Sands — Ground Water Sparging Wells
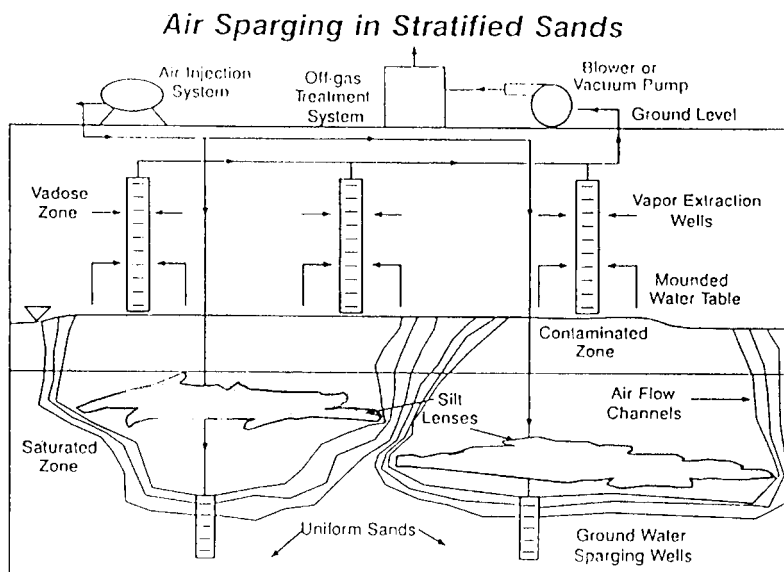
FIGURE 25.22 Typical air sparging system configuration. (From Hinchee, R. E. 1994. *Air Sparging for Site Remediation*. Lewis Publishers, Chelsea, Michigan.; Inyang, H.I. 1992. Selection and design of slurry walls as barriers to control pollutant migration. *A Seminar Presentation, Office of Solid Waste and Emergency Response*, USEPA, Washington, D.C.)

$$P_H = r_w g \left(L_s - L_{gw} \right) \tag{14}$$

where P_H is the hydrostatic pressure corresponding to the water column height that is displaced (atm), r_w is the density of water (1000 kg/m³), g is the acceleration due to gravity (9.8 m/s²), L_s is the depth below ground to the top of the sparge wellscreen (m), and L_{gw} is the depth to groundwater (m).

Two concerns exist related to the performance of air sparge systems, both of which can be addressed through monitoring. First, the introduction of air below the water table could raise the elevation of the water table, possibly changing groundwater flow direction. This first concern can be addressed by monitoring groundwater elevations near the sparge area to confirm that groundwater elevations are not changing. The second concern is that sparge gas may migrate horizontally in more permeable zones resulting in an increase in the lateral extent of contamination. This concern can be addressed by monitoring the chemical characteristics of soil vapor at locations beyond the sparge wells to verify that contaminated vapors are not migrating offsite in an uncontrolled manner and, if necessary, by using SVE to prevent offsite migration of vapors.

25.5.2.4 Chemical and Physical Processes

Chemical and physical processes can be used to promote *in situ* treatment of contaminated groundwater. The range of chemical processes that can be applied *in situ* are limited because the range of ambient conditions that affect chemical reactions (e.g., temperature, pressure, etc.) is relatively small; for example, the range of groundwater temperatures usually falls within the narrow range of about 50 and 60°F (10 and 15°C). Further, for a chemical process to be carried out to a high rate of completion, groundwater must be well mixed with reagents, which is difficult to achieve in an aquifer because of the large size of plumes of contaminated groundwater, the relatively low permeability of soil, and the heterogeneity of most aquifer systems. Still, under some conditions, *in situ* treatment of contaminated groundwater is possible, particularly in situations where contaminated groundwater can be routed to a collection point before it is treated and where contaminated groundwater can be well mixed with treatment reagents.

Chemical processes that can be considered for *in situ* application include adsorption, precipitation, dehalogenation, oxidation/reduction, fixation, and physical transformation. Such processes can be promoted by: (1) installing a permeable reaction barrier (also known as a "reaction trench," or "*in situ* reactive barrier," or "selective" barrier) in the aquifer to intercept contaminated groundwater and provide a controlled environment for the reaction; or (2) by applying reagents directly to groundwater through wells or by infiltration from the ground surface. The barrier can be constructed using the techniques described in Section 25.4.3 and can be backfilled with reagents, such as activated carbon (for treatment of volatile organic contaminants), limestone (to promote precipitation of dissolved metal ions), or iron filings (to capture organic carbon compounds (Focht, et al, 1996)), to provide the desired treatment of groundwater. A variation of the permeable reaction barrier is the "funnel and gate" reaction barrier, which is illustrated in Figure 25.23. This approach uses relatively impermeable barrier walls (i.e., the "funnels") to route groundwater to the permeable, reactive portion of the barrier (i.e., the "gate") and minimizes the length of barrier that must be maintained or reconditioned as the reaction sites in the barrier wall become exhausted.

The range of treatment technologies that can be used for permeable reaction barriers is illustrated in Table 25.17. The primary advantages of permeable reaction barriers are a result of the passive nature of the technology; for example, there is no need for a power source, which typically represents the majority of long-term operation and maintenance costs for active (e.g., pump-and-treat) remediation systems.

25.5.2.5 Innovative *In Situ* Treatment Approaches

A large number of innovative techniques for *in situ* remediation of groundwater have been developed in the past few years that could possibly provide a higher degree of permanence, a higher contaminant removal efficiency, or lower remediation cost for some groundwater contamination problems. A good summary of available innovative technologies is presented in USEPA (1994b). Descriptions of some promising innovative *in situ* remediation techniques are presented below.
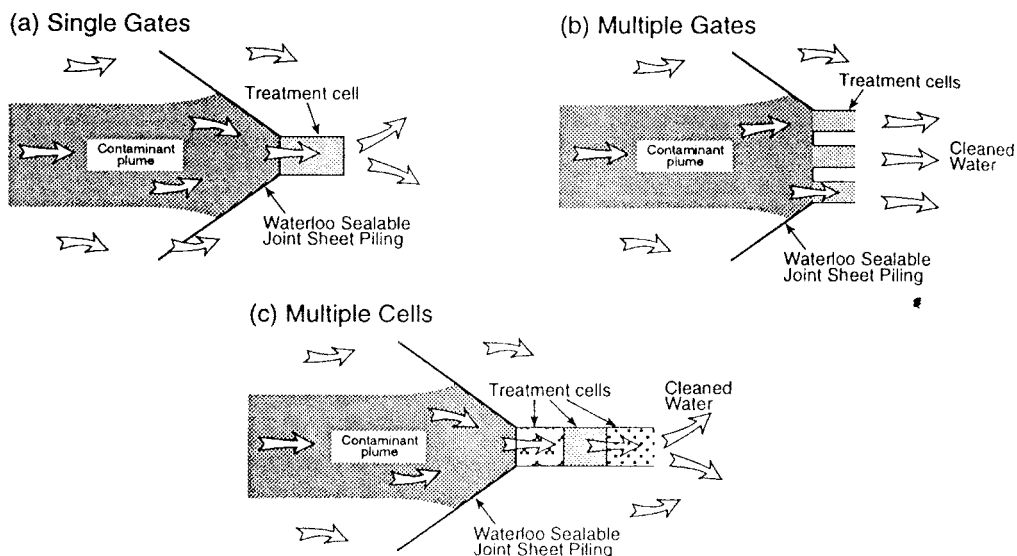
FIGURE 25.23 Funnel and gate system configuration. (From Rumer, R.R. and Mitchell, J.K. 1995. *Assessment of Barrier Containment Technologies — A Comprehensive Treatment for Environmental Remediation Applications.* National Technical Information Services, Springfield, VA.)

TABLE 25.17 Treatment Technologies Applied in Permeable Reactive Barriers

Treatment Medium	Target Contaminants	Technology Status
Zero valent iron	Halocarbons	Commercially applied
Zero valent iron	Reducible metals (Cr^{+6}, U)	Field demonstration
Limestone	Metals, acid waters	In practice (mining)
Precipitation agents (gypsum, hydroxyapatite)	Metals	Lab studies
Sorptive agents (Fe hydroxide, GAC*, zeolites, coal)	Metals and organics	Field demonstration and/or lab studies
Reducing agents (organic compost, dithionite, hydrogen sulfide)	Reducible metals	Field demonstration
Metal couples**	Halocarbons	Lab studies
Biologic electron acceptors (ORC*** oxygen source, nitrate)	BTEX	Field experiments

* GAC = granulated activated carbon
** Coupled oxidation of metal and reduction of halocarbon to produce chloride and Fe^{+2} in solution
*** ORC = oxygen release compound
From Rumer, R.R. and Mitchell, J.K. 1995. *Assessment of Barrier Containment Technologies — A Comprehensive Treatment for Environmental Remediation Applications.* National Technical Information Services, Springfield, VA.

- *Intrinsic Remediation.* Intrinsic remediation (also known as "natural attenuation") involves the evaluation and monitoring of naturally occurring processes that prevent the migration of contaminants to receptors. Currently, intensive remediation is most frequently used to describe biodegradation of contaminants by naturally occurring microbes; however, other natural physical or chemical processes (e.g., adsorption) may effectively prevent the migration of contaminants. In this approach, natural forces are allowed to act on the contaminated groundwater, without any human intervention, to remediate groundwater. Because microbes are present in nearly all aquifers and because other natural processes occur (e.g., sorption, complexation, precipitation, etc.) that mitigate groundwater contamination, these naturally occurring processes can have a beneficial impact on groundwater quality. Intrinsic remediation can be used alone or to supplement conventional remediation techniques. For example, a contaminant source could be removed prior to

intrinsic remediation to reduce contaminant loading in the aquifer and thereby improve the performance of intrinsic remediation.

- *Phytoremediation* refers to the engineered use of green plants to remove or contain contaminants in groundwater or soil (Schnoor, et al., 1995). Plants (such as poplar trees) remediate contaminants by: (1) direct uptake and accumulation in plant tissues; (2) stimulation of biodegradation; and (3) enhancement of mineralization in the root-soil interface by fungi and microbes. Phytoremediation is best suited for sites where: (1) contaminated groundwater is relatively shallow (i.e., less than 15 ft (5m) deep); (2) pollutants are moderately hydrophobic, chlorinated, or contain nitrotoluene ammonium; and (3) excess nutrients (e.g., nitrate, phosphate, etc.) are present in the subsurface soils.

- *Vacuum Vapor Extraction*. This technique (USEPA, 1995b) extracts contaminants from groundwater by circulating extracted groundwater in the extraction well, which volatizes the VOCs from the groundwater. The vapors are extracted from the well and treated; the treated groundwater is then reinjected into the aquifer formation and recirculated through the aquifer treatment zone. The technique has been shown in pilot studies to be effective at treating aquifers contaminated with both DNAPL and LNAPL constituents.

- *Wetlands Treatment for Metals-Contaminated Groundwater*. In this approach, contaminated groundwater is routed through a man-made wetland area, where the water percolates vertically through and out of the wetland. This technique was originally developed for use in treating acid mine drainage having low pH and high metals content. The process takes advantage of the fact that wetlands typically contain both an aerobic zone (i.e., at shallow depth) and an anaerobic zone (i.e., beneath the aerobic zone), which offer different opportunities for removal of contaminants from groundwater. As influent water passes vertically through the wetland, the plant matter filters and adsorbs some dissolved metals. Oxidation and precipitation reactions occur in the aerobic zone to remove metals as hydroxides, and metal sulfides are produced by reduction reactions in the anaerobic zone (McCoy and Associates, 1992).

- *Electrokinetic Decontamination*. In this approach, oppositely charged electrodes are placed in the soil and charged, causing electrolysis of groundwater. As the electrolyzed water ions (i.e., H^+ and OH^- ions) migrate toward the oppositely charged electrode, the H^+ ions create a relatively acidic front of groundwater, accelerating the desorption of contaminants that are adsorbed to soil. The process results in extraction of metals from soils, including fine-grained soils, provided that the soils have a sufficient buffering capacity (McCoy and Associates, 1992). The technology is in the advanced stages of development.

- *Density Driven Convection (DDC)*. DDC is designed to remove petroleum hydrocarbons by supplying oxygen to promote the biodegradation processes that naturally occur in soil. Air is injected into the bottom of a standard groundwater well; the aerated water rises in the well and flows out into the aquifer, where the oxygen-rich water enhances natural biodegradation. Simultaneously, groundwater flows into the bottom of the well, completing a convection current around the well. The system can be used in conjunction with a vapor extraction/treatment system if nonbiodegradable contaminants are encountered. The technology, which is currently being demonstrated, is expected to be suitable for both coarse and fine-grained soils and is also expected to have a greater radius of influence than traditional air-sparge wells (Schrauf, 1996).

25.5.3 *Ex Situ* Treatment

25.5.3.1 Biological Processes

Ex situ biological processes involve treatment of extracted groundwater in vessels until the concentration of the remaining contaminants is below a predefined level. Biological processes have been widely used for the past forty years in the United States to treat municipal wastewater. A good general reference for biological treatment is presented by Corbitt (1990). The processes used to treat municipal wastewater

are generally applicable to treatment of contaminated groundwater except in circumstances where the contaminants in groundwater are toxic to the microbes in the biological treatment system. *Ex situ* biological treatment processes include bioreactor and slurry-phase treatment, among others. A brief description of these two processes is presented below.

- *Bioreactors.* In this approach, contaminants are degraded by placing them in contact with microbes in an environment where biological growth is promoted. This can be accomplished using either suspended (i.e., activated sludge) or attached (i.e., rotating biological contactors or trickling filters) growth systems. The microbial population can be either derived from the contaminated groundwater or added to the system.
- *Slurry-Phase Treatment.* In this approach, extracted groundwater is mixed with soil and nutrients and then routinely agitated in a controlled environment to promote biodegradation of the contaminants. The slurry is typically placed in a lined area and left until the desired degree of biodegradation occurs. Finally, the slurry is dewatered, and then the water and soil (both should be uncontaminated after biodegradation is completed) are properly disposed of as nonhazardous wastes. This process offers much better control of process variables than the bioreactor approach but is slow and requires a larger area for implementation.

25.5.3.2 Volatization Processes

As described in Section 25.5.2.5 and as shown in Table 25.13, volatization processes are effective in treating groundwater that is contaminated with organic chemicals that have a high vapor pressure. Volatization processes may also be effective at removing nonorganic contaminants that have a relatively high vapor pressure (e.g., mercury). Volatization processes, or "air stripping," involve mixing of contaminated groundwater with air to allow the volatile contaminants that are dissolved in groundwater to partition from the dissolved phase to the vapor phase, allowing them to be removed from groundwater. The efficiency of the volatization processes depends on the number of air/water contacts. Air/water contacts can be maximized by using spray aeration, diffused-air aeration, packed column air stripping, tray aerators, or rotating disk aerator techniques. The primary advantage of *ex situ* over *in situ* volatization approaches is the flexibility of *ex situ* systems and the high efficiency rates of these systems.

25.5.3.3 Chemical and Physical Processes

Numerous well-developed technologies are available for treating extracted groundwater by chemical or physical means. Useful summaries of these processes are presented by the U.S. EPA and the U.S. Air Force (1993) and by Corbitt (1990). In general, chemical and physical processes are available for *ex situ* treatment of nearly all types of contaminants that occur in groundwater. *Ex situ* chemical and physical treatment techniques typically have the following advantages over *in situ* treatment techniques: (1) the ability to maintain uniformity of the influent by using equalization features (i.e., tanks, etc.) at the point of inflow to the treatment system; (2) the ability to easily evaluate the effectiveness of the treatment process by monitoring the quality of treated groundwater; and (3) ease in adjusting the treatment system components in response to changes in influent quality to optimize groundwater treatment efficiency. In Table 25.18, several chemical and physical treatment technologies are identified, and examples of their use are illustrated. Below, descriptions are provided for the *ex situ* treatment techniques that are most commonly used to treat contaminated groundwater.

- *Carbon Adsorption.* For this approach, groundwater is typically pumped through a series of canisters containing activated carbon, to which contaminants from the groundwater are adsorbed. Because activated carbon offers "attractive" locations for adsorption of organic carbons (i.e., because of the similarity between the carbon atoms on the contaminant molecule and the carbons in the activated carbon), carbon adsorption is a very effective technique for removing VOCs from contaminated groundwater. Periodically, the activated carbon must be either replaced or regenerated as the adsorption sites become saturated with VOCs.

TABLE 25.18 Summary of Selected *Ex Situ* Groundwater Treatment Technologies

Groundwater Treatment Technology	Representative Examples	Residual Streams	Status of Technology
Organic Contaminants			
Air stripping	Packed towers, surface or diffused aeration removal of volatile compounds; soil venting	Air stream with VOCs	Commercial
Liquid-phase	GAC removal of broad spectrum of VOCs	GAC for regeneration or disposal	Commercial
Steam stripping	Packed tower with steam stripping, removal of low volatile organics	Recovered solvent	Some commercial
Membranes	Ultrafiltration for removal of selected organics	Concentrated brine side stream	Commercial
Oxidation	Ozone/UV, or ozone/H_2O_2, destruction of chlorinated organics	None	Some commercial in development stages
Activated sludge	Oxygen or air biological oxidations for removal/destruction of degradable organics	Sludge	Commercial
Fixed-film biological reactors	Fixed-film fluidized bed, for oxidation of less degradable organics	Sludge	Commercial
Biophysical	Powdered carbon, with activated sludge, treatment of high strength waste waters	Powdered carbon and bacterial	Commercial, PACT process
Inorganic Contaminants			
Alkaline precipitation	Heavy metals removal	Hazardous sludge	Commercial
Coagulation	Ferric sulfate or alum for heavy metals removal	Hazardous sludge	Commercial
Ion exchange	Heavy metals; nitrate	Regeneration stream	Commercial
Adsorption	Selenium removal on activated alumina	Regeneration stream	Commercial
Filtration	Removal of clays, other particulates	Backwash wastes	Commercial
Reduction	SO_2 reduction of CR (VI)	Sludge	Commercial
Membranes	Reverse osmosis, ultrafiltration for removal of metals, other ions	Concentrated liquid waste	Commercial, new membranes are under development
Oxidation	Fe(II) and Mn(II)	Sludge	Commercial

From USEPA. 1994b. *Superfund Innovative Technology Evaluation Program*, Technology Profiles, Seventh Edition. EPA/540/R-94/526. Washington, DC.

- *Ion Exchange.* Ion exchange is similar to adsorption, although adsorption is caused by coordination bonding (i.e., bonding by van der Waals forces) of metals to specific surfaces, whereas ion exchange results from exchange of ions (typically metal ions) through electrostatic forces. Ion exchange has been used effectively for several years for cost-effective removal of dissolved metals. The technique is implemented by passing contaminated groundwater through a reaction vessel that contains resins having reaction sites that attract the contaminants. As with carbon adsorption, ion exchange resins must be replaced or rejuvenated when the reaction sites become saturated with contaminants.

- *Precipitation.* Precipitation occurs when a constituent exists at a concentration in groundwater that is greater than its solubility limit. Solubility is a function of many factors, including pH,

temperature, and the presence of other dissolved constituents. Precipitation of a target constituent is caused by manipulating these factors to change the equilibrium in the direction that promotes precipitation of that constituent. Typically, the least-soluble constituent at a given pH will precipitate from the solution. Note that precipitation can also occur *in situ* in response to changes in groundwater chemistry. Because solubility is affected by many factors, changes in groundwater quality over time may cause changes in the solubility of constituents of concern. Precipitation and dissolution may be slow reactions, depending on the solubility of the constituent; therefore, after a change occurs in groundwater chemistry that increases the constituent's solubility (e.g., drawing groundwater having lower pH through a contaminated aquifer), the previously precipitated minerals may require a long time to redissolve and be removed from the aquifer.

- *Reverse Osmosis.* In reverse osmosis (RO) (USEPA, 1996b), solids are separated from water by creating a concentration gradient across a semipermeable membrane (i.e., a membrane that is "selectively" permeable to only the solids of choice, based on size of molecules, but not to other compounds or solids). To prevent water from also passing through the membrane, a pressure gradient is applied to the opposite side of the membrane that exceeds the osmotic pressure (hence the use of the term "reverse" osmosis). The technology is extremely effective at removing dissolved solids from groundwater, but RO capital equipment is very expensive.

25.6 Performance Monitoring of Groundwater Remediation Systems

Monitoring the performance of a groundwater remediation system allows evaluation of the success of the system in meeting the remediation goal and the design criteria for the system. Different approaches are typically used to monitor the performance of groundwater extraction, barrier, and treatment remediation systems. Performance monitoring of groundwater extraction systems typically includes at least the following activities (USEPA, 1994a):

- Confirming that the system is extracting groundwater at flow rates similar to those predicted in the predesign pump test (by measurement using a flowmeter at the wellhead)
- Verifying that groundwater elevations near the extraction system are being affected as needed to control groundwater flow direction and gradient (by measuring groundwater elevations at wells or piezometers located near the extraction well(s))
- Verifying that there are no significant breaches in the groundwater extraction system that would allow contaminated groundwater to flow past the extraction system (using the groundwater elevation data collected at the wells and piezometers)
- Confirming that the system components perform consistently throughout the operational life of the facility (by routinely inspecting the components of the system and checking for signs of operational problems).

Performance monitoring of a groundwater treatment system (both *in situ* and *ex situ*) typically includes the following:

- Periodic analysis of influent groundwater quality to evaluate changes in the quality of water requiring treatment
- Periodic analysis of effluent groundwater quality, to evaluate the success of the system at meeting treatment goals
- Routine inspection of the components of the system to identify potential future problems.

Performance monitoring of subsurface barriers (see Rumer and Mitchell, 1995) typically includes the following activities:

- Verifying that the barrier is having the desired impact on groundwater flow direction and gradient (by measuring groundwater elevations at wells or piezometers near the barrier)

- Verifying that contaminated groundwater is not passing through or around the barrier (by monitoring groundwater quality beyond the barrier)
- Verifying that the barrier is not degrading over time.

These activities can be monitored in a variety of ways using a system of piezometers or wells, as described in Section 25.3.5. Note that an advantage of using groundwater extraction techniques over physical barriers is the ability to change the extraction (i.e., pumping) rates to optimize the performance of the system. For example, if it is found from operational data that the actual zone of influence of the extraction wells is less than predicted, then the pumping rate could possibly be increased to impact the required zone of influence. As another example, if it is found that contaminants are not desorbing as quickly as predicted, then the pumping system could be "pulsed" to increase the ratio of contaminant mass removed to total volume of groundwater extracted (as illustrated in Figure 25.14). When the purpose of the performance monitoring program is to verify that nearby receptors are not at risk of contaminated groundwater, external monitoring features should be located downgradient of the hydraulic containment system near receptors or near a property boundary.

For Further Information

USEPA (1989a) gives a good discussion of the basic approaches involved in performing a risk assessment. A discussion of selection of cleanup goals and current issues in groundwater remediation technology is presented by NRC (1994).

The U.S. EPA (1994b) evaluates emerging remediation technologies and routinely publishes updates of the results of the program.

References

Bear, J. 1979. *Hydraulics of Groundwater*, McGraw-Hill, New York.

Brusseau, M. 1996. Evaluation of simple methods of estimating contaminant removal by flushing. *Groundwater,* Jan./Feb. p. 19-22.

Cohen, R.M. and Miller, W.J. III, 1983. Use of analytical models for evaluating corrective actions at hazardous waste sites. Proceedings of the Third National Symposium on Aquifer Restoration and Ground-Water Monitoring. *National Water Well Association.* 86-97.

Corbitt, R.A., editor. 1990. *Standard Handbook of Environmental Engineering.* McGraw-Hill, New York.

Driscoll, F.G. 1986. *Groundwater and Wells.* Johnson Division, St. Paul, MN.

Focht, R., Vaughn, S., O'Hannesin. 1996. Field application of reactive iron walls for in-situ degradation of volatile organic compounds. *Remediation.* Summer 1996.

GeoSyntec Consultants. 1994. *Subsurface Barriers.* Report to the United States Navy, Contract SBIR Topic N93-130. Atlanta, Georgia. 80 pages.

Hinchee, R. E. 1994. *Air Sparging for Site Remediation.* Lewis Publishers, Chelsea, Michigan.

Inyang, H.I. 1992. Selection and design of slurry walls as barriers to control pollutant migration. *A Seminar Presentation, Office of Solid Waste and Emergency Response*, USEPA, Washington, D.C.

Johnson, R.I., Johnson, P.C., McWhorter, D.B., Hinchee, R.E., and Goodman, I. 1993. An overview of in situ air sparging. *Groundwater Monitoring and Remediation.* Fall 1993, 127-135.

Johnson, P.C., Stanley, C.C., Kemblowski, M.W., Byers, D.L., and Colthart, J.D. 1990. A practical approach to the design, operation, and monitoring of in-situ soil-venting systems. *Groundwater Monitoring Review*, Spring 1990, 159-178.

Keely, J.F. 1989. *Performance Evaluations of Pump-and-Treat Remediations*, EPA Superfund Groundwater Issue, EPA/540/4-89/005. Washington, DC.

McCoy and Associates. September/October 1992. *The Hazardous Waste Consultant.* 4.1-4.38.

McDonald, M.G. and Harbaugh, A.W. 1988. *A Modular Three-Dimensional Finite-Difference Groundwater Flow Model.* USGS Techniques of Water-Resources Investigations, Book 6, Chapter A1, USGS, Reston, VA.

Merritt, F.S. 1989. *Standard Handbook for Civil Engineers*. McGraw-Hill, New York.

National Research Council (NRC), Water Science and Technology Board. 1994. *Alternatives for Ground-water Cleanup*. Washington, D.C.

Rumer, R.R. and Mitchell, J.K. 1995. *Assessment of Barrier Containment Technologies — A Comprehensive Treatment for Environmental Remediation Applications*. National Technical Information Services, Springfield, VA.

Russel, M., Colglazier, E.W., and English, M.R. 1991. *Hazardous Waste Site Remediation: The Task Ahead*. Waste Management Research and Education Institute. University of Tennessee, Knoxville.

Sai, J.O. and Anderson, D.C., Barrier wall materials for containment of dense nonaqueous phase liquid (DNAPL). 1992. *Hazardous Waste and Hazardous Materials*. 9, 4, 317-330.

Schnoor, J.L., Light, L.A., McCutcheon, S.C., Wolf, N.L., and Caneira, L.H. 1995. Phytoremediation of organic and nutrient contaminants. *Environmental Science and Technology*. 29, 7.

Schrauf, T.W. 1996. A well-developed cleanup technology. *Environmental Protection*. May 1996. 24-25.

USEPA. 1985. *Remedial Action Costing Procedures Manual*. EPA/OERR/HSCD; EPA Report #600/8-87/049, Washington, D.C.

USEPA. 1989a. *Risk Assessment Guidance for Superfund: Interim Final Guidance*. Office of Emergency and Remedial Response (EPA/540/1-89/002). Washington, DC.

USEPA. September 1989b. Seminar Publication — *Transport and Fate of Contaminants in the Subsurface*. EPA/625/4-89/019. Washington, DC.

USEPA. July 1991. *Handbook — Ground Water, Volume II: Methodology*. EPA/625/6-90/016b. Washington, DC.

USEPA. 1993a. *Cleaning up the Nation's Waste Sites: Markets and Technology Trends*. EPA 542-R-92-012. Washington, DC.

USEPA. 1993b. *Guidance for Evaluating the Technical Impracticability of Ground-Water Restoration*. EPA-540-R-93-080. Washington, DC.

USEPA and U.S. Air Force, 1993. *Remediation Technologies Screening Matrix and Reference Guide*. EPA 542-B-93-005. Washington, DC.

USEPA. 1994a. *Monitoring the Performance of Groundwater Pump and Treat Systems*. EPA/600/R-94/123. Washington, DC.

USEPA. 1994b. *Superfund Innovative Technology Evaluation Program*, Technology Profiles, Seventh Edition. EPA/540/R-94/526. Washington, DC.

USEPA. 1995a. *Ground-Water and Leachate Treatment Systems*. EPA/625/R-94/005. Washington, DC.

USEPA. 1995b. *Evaluation of Technologies for In-Situ Cleanup of DNAPL Contaminated Sites*. EPA/600/R-94/0120. Washington, DC.

USEPA. 1996a. *Superfund Chemical Data Matrix*. EPA/540-R-94-009. Washington, DC.

USEPA. 1996b. *Capsule Report: Reverse Osmosis Process*. EPA/625/R-96/009. Washington, DC.

Xanthakos, P. 1979. *Slurry Walls*. McGraw-Hill, New York.

Zheng, C., Wong, H.F., Anderson, M.P., and Bradbury, K.R. 1988. Analysis of interceptor ditches for control of groundwater pollution. *Journal of Hydrology*. 98, 67-81.

Glossary

Adsorption   The adherence of ions or molecules in solution to the surface of solids.

Anisotropy   The condition under which one or more of the hydraulic properties of an aquifer vary with direction.

Biodegradation   The biologically mediated conversion of a compound to simple products.

Contaminant   An undesirable substance that is not normally present in groundwater or a substance that may naturally occur in groundwater, but that is present at an undesirably high concentration.

DNAPL   An acronym for denser-than-water nonaqueous-phase liquid; a liquid, composed of one or more contaminants, that does not mix with water and is denser than water.

Desorption   The release of sorbed molecules from the solid into solution (the reverse of adsorption).

Fraction of Organic Carbon   The organic carbon content of soil, expressed as a mass fraction of the dry soil.

Hydraulic Control   Prevention of the spread of contaminated groundwater using physical features to control groundwater flow.

Ion   A molecule or atom that has a positive or negative electric charge.

LNAPL   An acronym for less-dense-than-water (i.e., "light") nonaqueous-phase liquid.

Maximum Contaminant Level (MCL)   The maximum amount of a compound allowed in drinking water under the Safe Drinking Water Act.

Octanol-Water Partition Coefficient   A measure that indicates the extent to which a compound is attracted to an organic phase (for which octanol is a proxy) and, hence, the tendency of the compound to sorb to subsurface materials.

Partitioning   A chemical equilibrium condition in which the concentration of a chemical is apportioned between two different phases according to the partition coefficient.

Plume   A zone of contaminated groundwater.

Receptor   An organism that is exposed to a contaminant.

Retardation   The movement of a solute through a geologic medium at a velocity less than that of the groundwater as a result of phenomena that separate a fraction of the solute mass from the groundwater.

Sorption   Refers to a reversible process involving physical reaction of aquifer material and dissolved constituents.

Upgradient   In the direction of increasing hydrostatic head.

Volatization   The transfer of a chemical from the liquid to the gas phase.

26
Biodegradation

Loring F. Nies
Purdue University

Vivek Kapoor
Georgia Institute of Technology

26.1 Introduction

The objective of this chapter is to present the basic principles and concepts of groundwater microbiology and transport processes as they relate to biodegradation. Understanding of these processes is necessary to model, forecast, and manipulate subsurface contaminant transport and biodegradation. The basic principles presented may be applied to the biorestoration of contaminated groundwater, a field now commonly called "bioremediation" (Lee et al., 1988; National Research Council, 1993; Thomas and Ward, 1989).

26.1.1 Subsurface Microbiology and Geochemistry

26.1.1.1 Basic Microbiology

Microorganisms mediate the biotransformation of many different chemicals in the subsurface. Historically there has been much scientific debate over the relative contribution of abiotic and biotic chemical processes in the subsurface. Until the last 50 years, very little was known about microorganisms in groundwater, and thus abiotic transformations were thought to dominate. However, more recent research has contributed new knowledge about microbial life in the subsurface and this view has changed. While our understanding of subsurface microbiology and biogeochemistry is far from complete, biologically mediated reactions of both inorganic and organic compounds in groundwater are known to be significant, and in many cases they control groundwater chemistry (Chapelle, 1993). With respect to the biotransformation of organic pollutants, microbially mediated reactions are by far the most important. The abiotic transformation of pollutants does occur (e.g., dehalogenation, polymerization, and hydrolysis reactions), however, the environmental significance of these abiotic reactions relative to microbially mediated reactions is difficult to assess. This is partly due to the difficulty in distinguishing between

abiotic and biotic processes *in situ*. In addition, certain abiotic reactions are dependent on two environmental parameters which are often controlled by microbial processes, redox and pH. Thus these latter types of abiotic reactions are indirectly mediated by microorganisms as well.

Microorganisms can be classified in a number of different ways (e.g., phylogenetically), however for the purpose of understanding biodegradation it is most useful to begin the characterization of microorganisms according to their source of energy and their source of carbon for cell growth. Potential energy sources are organic carbon (including pollutants), inorganic compounds, and sunlight. Microbes which oxidize organic compounds to obtain energy are organotrophs, and those that oxidize inorganic compounds are lithotrophs. Potential sources of carbon for cell growth are either organic carbon or inorganic carbon (HCO_3^-, CO_2). Microbes which degrade organic compounds to obtain carbon for the synthesis of cellular constituents are heterotrophs and those that utilize inorganic carbon are autotrophs. Logically, most lithotrophs are also autotrophs and most organotrophs are also heterotrophs. Photoautotrophs and photoheterotrophs may exist in groundwater, but their utilization of photometabolism is obviously limited by access to sunlight (Brock et al., 1997; Gottschalk, 1986).

The types of microorganisms found in groundwater include protozoa, fungi, bacteria (in this chapter includes both Bacteria and Archaea), and viruses. For purposes of examining microscale phenomena (biodegradation) relative to macroscopic subsurface particle and pore sizes, it is useful to consider microorganism size. Protozoa have a size range of approximately 2 to 200 μm, the smallest single-celled fungi are approximately 2 μm, and bacteria range in size from 0.1 μm to 15 μm, although a bacterium which is 500 μm long has recently been discovered. Viruses are much smaller, ranging from approximately 0.1 μm to 0.01 μm (Brock et al., 1997).

Protozoa are relatively large heterotrophic microorganisms whose distribution in groundwater has not been thoroughly studied. Protozoa are not known to be significant biodegraders of organic pollutants, although many protozoa do feed on bacteria. This predation is thought to reduce bacterial numbers in some groundwater (Chapelle, 1993). Thus it is possible that protozoa could indirectly adversely influence biodegradation rates in groundwater by reducing populations of pollutant biodegrading bacteria. Some protozoa are also human pathogens, therefore their distribution in groundwater is a public health interest (e.g., *Entamoeba histolytica*, the causative agent of dysentery). Drinking water is the primary route of human exposure to pathogenic protozoa. Fortunately, the cysts of the recently problematic pathogenic protozoa *Cryptosporidium* and *Giardia* have thus far been detected only in surface waters (LeChevallier and Norton, 1995).

Fungi are heterotrophic biodegraders of immense capability. There are many types of single-celled microscopic fungi as well as multicelled macroscopic organisms, such as mushrooms. The ability of fungi to biodegrade organic pollutants is well known, however, in groundwater the bacteria dominate in terms of both biomass and as catalysts of biodegradation. Therefore, all microbial biodegradation discussed in this chapter can be assumed to be mediated by bacteria, unless stated otherwise.

Viruses are obligate parasites whose hosts include bacteria as well as macroscopic organisms. The transport and survival of viruses in groundwater is of interest with respect to the control of infectious diseases (Bitton et al., 1983). Viruses have no catabolic capability and therefore do not contribute directly to biodegradation. The influence of the viral infection of bacteria on biodegradation, and on microbial ecology in general, has not been systematically examined. Since viruses are essentially containers of genetic material which is inserted into the host, they can be used as gene vectors to insert desirable genetic information (e.g., biodegradation genes) into bacteria. While this type of technology is available in the laboratory, it is not currently used in the field.

Bacteria are ubiquitous in groundwater, and in addition, they have very diverse catabolic abilities, morphology, physiology, and biochemical constituents. This diversity allows bacteria to survive in some of the most extreme environments on earth. Bacteria have been found up to 2.8 kilometers below the surface (Fredrickson and Onstott, 1996). Although they are microscopic, bacteria are the dominant biodegraders and drivers of biogeochemical cycling in groundwater. Bacteria live attached to subsurface particles as individuals or in colonies and also as unattached motile organisms. Attached microbes conserve energy while removing nutrients from the surrounding water, as well as deriving some protection

from predation. The growth and survival of attached organisms is dependent on obtaining nutrients by transport through groundwater. Motile organisms are known to move in response to chemical gradients (chemotaxis), moving toward higher concentrations of nutrients and away from toxics. Bacterial transport in groundwater and environmental factors which induce bacteria to exhibit attached or motile phenotypes have been studied (Corapcioglu and Haridas, 1984, 1985; Dawson et al., 1981; Fletcher and Marshall, 1982; Harvey et al., 1989). However, our ability to control and influence these processes in the subsurface in order to promote biodegradation is still under development.

26.1.1.2 Basic Metabolism

All living organisms obtain energy by mediating oxidation/reduction reactions. Reduced organic and inorganic compounds serve as the reductant (electron donor) in the reaction. Oxidized, usually inorganic, species serve as the oxidant (electron acceptor). Aerobic respiring organisms utilize oxygen, which is the most ubiquitous electron acceptor in the environment, as an oxidant. However, the solubility of oxygen in water is relatively low (~8 mg/L at 20°C) and the oxygen diffusion rate is slow, therefore, the availability of oxygen in groundwater is significantly limited by mass transfer. A variation of typical oxidation/reduction reactions occurs when a fraction of a compound is oxidized and the remaining fraction is reduced. This type of catabolism is called fermentation, in contrast to respiration. Almost all fermenting bacteria cannot tolerate oxygen and are found in anaerobic environments living in close association with sulfate reducing or methanogenic bacteria (Brock et al., 1997; Chapelle, 1993; Gottschalk, 1986).

Organisms have evolved elegant mechanisms for converting the electron transfer which occurs during oxidation/reduction reactions into energy. Respiring organisms oxidize their energy source, stripping off the high-energy electrons. The electrons are shuttled on electron carriers (NADH) to the cell membranes where electron transport phosphorylation occurs (Figure 26.1). Respiring organisms use electron transport to pump protons across the membrane to create a pH differential (charge and proton gradient) between the inside and outside of the cell. In this energized state, protons are driven by the gradient through specialized enzymes which convert the "proton motive force" into stored chemical energy, adenosine triphosphate (ATP). Driving the proton pumps consumes the electron's energy. At the terminal end of the electron transport chain, the now low-energy electrons are transfered to a terminal electron acceptor.

Fermenting bacteria make ATP through the direct conversion of chemical bond energy in a process called "substrate level phosphorylation (SLP)." The energy-producing steps occur primarily during oxidation of carbonyl groups (Figure 26.2). Since no external electron acceptor is utilized, and biological electron carriers are in limited supply and must be recycled, the electrons generated from the energy-producing oxidations of SLP must be discarded. This is accomplished by reducing some of the substrate, producing alcohols, or by reducing protons, producing hydrogen. Thus fermenters produce a mixture of both oxidized and reduced products such as CO_2, carboxylic acids, alcohols, and H_2.

Anaerobic respiration occurs when oxygen is depleted and other suitable oxidized species are available to act as an electron acceptor. Common anaerobic electron acceptors are oxidized nitrogen compounds (nitrate, nitrite), iron(III), oxidized sulfur compounds (sulfate, sulfite, S°), and CO_2. Other oxidized inorganic species are utilized in anaerobic respiration as well, and in addition, some organic compounds, such as fumarate, may also serve as electron acceptors (Zehnder, 1988). Bacteria obtain less energy from using electron acceptors other than oxygen. The amount of energy obtained from the oxidation of a given substrate is proportional to the reduction potential ($E°'$) of the electron acceptor (Table 26.1). Electron acceptors which yield the largest amount of energy (e.g., oxygen) tend to be utilized preferentially over others that are available, probably because bacteria with access to more energy can grow faster and compete for resources more successfully. In addition, biochemical mechanisms exist which maximize energy production. For example, facultative aerobes which can use both oxygen and nitrate as electron acceptor are common. The presence of oxygen inhibits synthesis of the enzyme at the terminus of the electron transport chain which transfers electrons to nitrate (nitrate reductase), thus ensuring more energetically favorable oxygen utilization whenever possible.
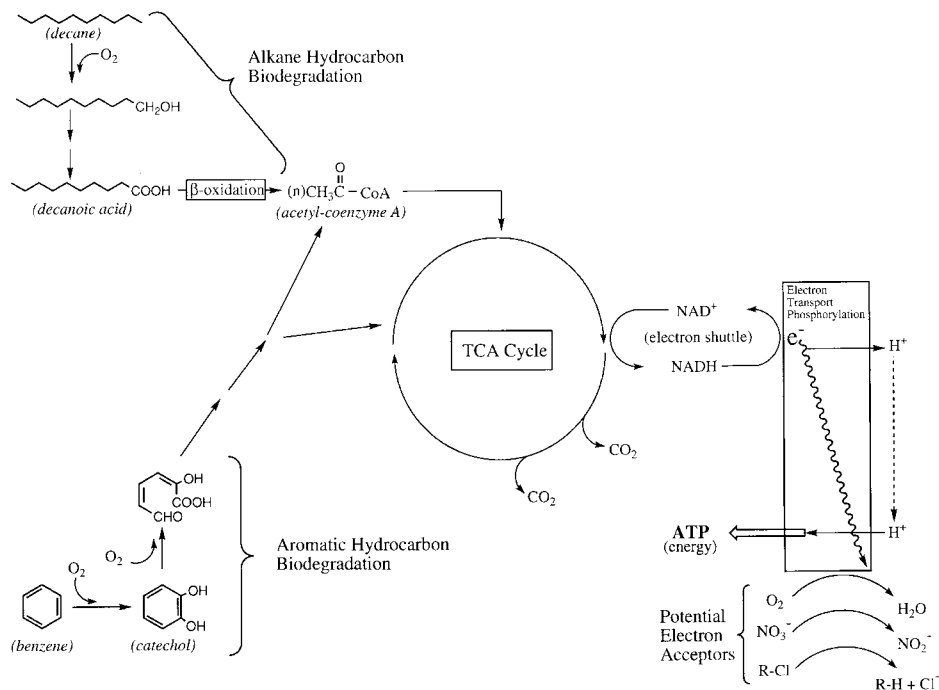
FIGURE 26.1 Idealized schematic of hydrocarbon biodegradation by a respiring organism.

During biodegradation of organic compounds, large molecules are broken down into small molecules, which are oxidized, yielding electrons for energy production. For example, aromatic rings are cleaved open to form aliphatic chains, which are then cleaved into two-carbon pieces, which are subsequently oxidized to CO_2 . Likewise, long alkanes are cleaved into two-carbon pieces prior to oxidation to CO_2. However, the initiation of biodegradation of many compounds requires "activation" by the insertion of molecular oxygen (O_2) into the hydrocarbon, forming mono- and dihydroxylated compounds. Oxygenase enzymes perform this function. Once activated by oxygenase enzymes, further biodegradation of hydrocarbons often does not require molecular oxygen. An exception is the biodegradation of aromatic rings, during which oxygen is used as a reactant for ring activation and ring cleavage. Thus,
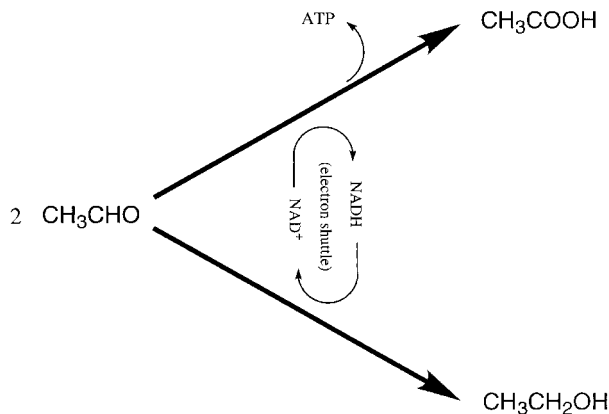
FIGURE 26.2 Schematic of fermentation.

TABLE 26.1 Reduction Potentials of Important Biological Oxidants (electron acceptors)

	$E^{\circ\prime}$ (Volts)
$\frac{1}{4}O_2(g) + H^+ + e^- = \frac{1}{2}H_2O$	0.820
$\frac{1}{2}MnO_2(s) + \frac{1}{2}HCO_3^-(10^{-3}) + \frac{3}{2}H^+ + e^- = \frac{1}{2}MnCO_3(s) + H_2O$	0.527
$\frac{1}{2}NO_3^- + H^+ + e^- = \frac{1}{2}NO_2^- + \frac{1}{2}H_2O$	0.423
$\frac{1}{6}NO_2^- + \frac{4}{3}H^+ + e^- = \frac{1}{6}NH_4^+ + \frac{1}{3}H_2O$	0.344
$FeOOH(s) + HCO_3^-(10^{-3}) + 2H^+ + e^- = FeCO_3(s) + 2H_2O$	−0.047
$\frac{1}{6}SO_4^{2-} + \frac{4}{3}H^+ + e^- = \frac{1}{6}S(s) + \frac{2}{3}H_2O$	−0.195
$\frac{1}{2}S(s) + H^+ + e^- = \frac{1}{2}H_2S(g)$	−0.243
$\frac{1}{8}CO_2(g) + H^+ + e^- = \frac{1}{8}CH_4(g) + \frac{1}{4}H_2O$	−0.244
$H^+ + e^- = \frac{1}{2}H_2$	−0.414

Values apply for unit activity (1M or 1atm) in water at pH 7.0 and 25˚C, except $HCO_3^- = 10^{-3}$M which more typically represents environmental conditions.
$\Delta G^{\circ\prime} = -nFE^{\circ\prime}$, where n is the number of electrons transferred and F is Faraday's constant.
F = 96.5 kJ/volt-mole.
Adapted from Brock, T. D., M. T. Madigan, J. M. Martinko, and J. Parker. 1997. *Biology of Microorganisms.* Prentice Hall, Englewood Cliffs, NJ.; Stumm, W. and Morgan, J. J. 1981. *Aquatic Chemistry.* John Wiley & Sons, New York.; Thauer, R. K., Jungermann, K., and Decker, K. 1977. Energy conservation in chemotrophic anaerobic bacteria. *Bacteriol. Rev.* 41:100-180.

it should be observed that oxygen can be utilized by bacteria as both electron acceptor and as a reacting co-substrate for initiating biodegradation reactions. A critical issue in biodegradation is that there are many alternatives to oxygen as an electron acceptor, while there are few alternatives to oxygen as a reactant for initiating hydrocarbon biodegradation. For this reason, when more rapid and complete biodegradation of certain hydrocarbons is desired, aerobic conditions are preferred. The bioremediation of groundwater aquifers which are contaminated with hydrocarbons is often limited by insufficient oxygen. Hydrogen peroxide, which decomposes to oxygen, is often used as a highly soluble source of oxygen for groundwater systems (Pardieck et al., 1990; Pardieck et al., 1992).

One good alternative to oxygen as an electron acceptor is nitrate (NO_3^-). Significant research has also been devoted to the feasibility of using nitrate as a supplemental electron acceptor (Dolfing et al., 1990; Evans et al., 1991; Hutchins, 1991a,b). Nitrate is more soluble in water than oxygen, is a strong oxidant, and could potentially be more readily introduced into groundwater aquifers through injection wells. Interestingly, many monoaromatic compounds (BTEX) are biodegraded under anaerobic nitrate reducing conditions, although unlike oxygen, the nitrate is used only as an electron acceptor and not as a reactant in biotransformation reactions. Nitrate has relatively low toxicity, but does have known health hazards (methemoglobinemia) and is a regulated drinking water contaminant, and therefore, is used as a supplemental electron acceptor only in carefully controlled situations.

26.1.2 Environmental Conditions which Influence Biodegradation

The activity and types of microorganisms present in groundwater are greatly influenced by subsurface physical and chemical properties (Alexander, 1994). Biodegradation kinetics are generally highly dependent on temperature because enzyme function is temperature sensitive. Bacteria can be classified by the range of temperature within which they can grow. At optimum temperatures biodegradation kinetics reach maximum rates, while slightly above optimum temperature, cell constituents usually begin to degrade. As temperatures decrease below the optimum, biodegradation rates decrease until enzymes function at rates too slow to support growth. Temperature extremes will select for climate-adapted organisms. For example, cold-loving bacteria (psychrophiles) can grow at temperatures below 0°C and generally die at temperatures greater than 20°C. Mesophilic bacteria thrive at temperatures which are comfortable to humans. Although the vast majority of biodegradation studies have been done with mesophilic bacteria at temperatures ranging between 15 and 45°C, some evidence exists that suggests that psychrophilic (or at least psychrotolerant) (Kellems et al., 1994) and thermophilic bacteria also possess pollutant biodegrading ability (Chen and Taylor, 1995).

Bacteria can also be classified according to the pH range within which they can grow. Our knowledge of biodegradation has been derived chiefly from studies conducted at "neutral" pH range, between approximately 6 and 8. Most groundwater also falls into this pH range. Acidophilic bacteria can tolerate a pH as low as 2, although most are lithoautotrophs which do not degrade organics. Alkaline conditions exist (pH 9 to 11) in certain areas where carbonate rocks predominate. Sodium concentrations in these alkaline environments are often more than 10X greater than seawater, thus indigenous bacteria are halophiles in addition to their adaptation to extreme pH. Knowledge of pollutant biodegradation in extreme pH environments is scarce.

Bacteria are sensitive to salinity. Halophilic bacteria have evolved a mechanism to counteract the tendency for cells to become desiccated by high salt concentrations. Numerous studies have demonstrated that marine bacteria are capable of hydrocarbon biodegradation, however, knowledge about pollutant biodegradation by extreme halophiles is lacking. Since a primary concern is the bioremediation of contaminated aquifers which are sources of potable water, acquiring knowledge about organisms adapted to extreme environments has not been a high priority.

All living things, including bacteria, require inorganic nutrients for growth, in addition to a source of energy and carbon. Nitrogen and phosphorus are the nutrients most likely to be a limiting factor for biodegradation and, similar to agricultural applications, are often added to the environment to remove potential nutrient limitations.

Oxidation/reduction potential is by definition the electrical potential (in volts) of the oxidation/reduction reaction occurring between the electron donor and electron acceptor. In practice when considering conditions in the environment the term *redox* potential is commonly used. Redox potential typically refers to the reduction potential of the dominant electron acceptor in the environment. Probes are available to measure environmental redox potential, however when assessing biodegradation, actual measurement of the electron acceptor of interest is more useful. For example, if aerobic biodegradation is desired, oxygen concentrations should be monitored. Likewise, the simultaneous disappearance of nitrate and organics, with the concurrent appearance of nitrite, is better circumstantial evidence of the existence of denitrifying bacteria than a redox measurement. More reducing, anaerobic conditions can be inferred from measurements of sulfate, sulfide, and the production of methane, as well as redox measurements.

26.1.3 Biotransformations

Bacteria biodegrade organic compounds by breaking large molecules apart with an array of biotransformation reactions such as hydrolysis, oxidation, reduction, dehalogenation, deamination, decarboxylation, and rearrangement reactions. As mentioned previously, molecular oxygen participates directly in a number of biotransformation reactions (Table 26.2).

TABLE 26.2 Biotransformation Reactions Mediated by Oxygenase Enzymes

Biotransformation	Compound	Reference
Deamination	Aniline	Bachofer et al., 1975
Dehalogenation	Pentachlorophenol	Xun et al., 1992
Ether cleavage	2,4,-Dichlorophenoxyacetate	Tiedje and Alexander, 1969
Aromatic ring hydroxylation	Benzene	Gibson, 1984
Aromatic ring fission	Benzene	Gibson, 1984
Alkane hydroxylation	Octane	Watkinson and Morgan, 1990

Anaerobic bacteria can perform many of the biotransformation reactions listed in Table 26.2 by different mechanisms without molecular oxygen (Schink, 1988). In addition, anaerobic bacteria can mediate reductive dehydroxylation, reductive deamination, and reductive dehalogenation reactions. Of these, reductive dehalogenations are of the greatest interest with respect to the biotransformation of hazardous pollutants. No anaerobic bacteria have been shown to hydroxylate alkanes. Thus, saturated hydrocarbons persist under anaerobic conditions. While BTEX biodegradation by denitrifying and iron-reducing bacteria has been well documented and a few reports of aromatic hydrocarbon degradation under sulfate reducing and methanogenic conditions exist, detailed knowledge about the biochemistry of anaerobic aromatic hydrocarbon transformation is limited (Evans, 1988; Lovley et al., 1989; Lovley and Lonergan, 1990).

In most cases, bacteria biodegrade organic compounds to obtain energy for growth. Due to the relaxed specificity of some enzymes, bacteria perform certain biotransformation reactions on compounds that are not growth substrates. The term *cometabolism* can be broadly taken to mean the "gratuitous biotransformation of a compound from which the organism derives no benefit." Often these gratuitous reactions produce products which can be more easily biodegraded by other organisms.

26.1.3.1 Hydrocarbons

Biodegradation of alkanes is usually initiated with terminal hydroxylation by a monooxygenase enzyme and subsequent oxidation of the alcohol to a carboxylic acid. The initial step requires oxygen, and thus far, no other anaerobic mechanism for the initiation of biodegradation of alkanes has been reported. The carboxylic acid can easily be further oxidized to CO_2 through two nearly universal biochemical pathways, ß-oxidation and the Krebs (tricarboxylic acid) cycle (Figure 26.1). Branched alkanes can be more difficult to degrade depending on the degree of branching. Branching interferes with ß-oxidation, and significant branching can result in complete inhibition of biodegradation (Watkinson and Morgan, 1990). Alkenes can be aerobically biodegraded similarly to alkanes, however the double bond can also be hydrolyzed to initiate biodegradation under anaerobic conditions (Schink, 1988). Since alkanes have relatively low water solubilities, they are less of a hazard to migrate as soluble constituents of groundwater. Biodegradation of alkanes often occurs at the hydrocarbon/water interface and is mediated by bacteria which produce biosurfactants for hydrocarbon uptake.

As described previously, aerobic biodegradation of aromatic compounds is initiated by the hydroxylation of the ring by oxygenase enzymes. Ring cleavage also requires oxygen. The aromatic compounds of most common concern are benzene, ethylbenzene, toluene, o-xylene, m-xylene, and p-xylene (BTEX). Initiation of the biodegradation of the alkylbenzenes occurs either by dioxygenase attack on the aromatic ring or by monooxygenase attack on the alkyl group followed by a dioxygenase-mediated ring cleavage (Smith, 1990). Biodegradation rates are generally observed in the following order, which may vary from site to site: toluene, ethylbenzene > benzene > m-xylene, p-xylene > o-xylene. BTEX are constituents of gasoline and therefore often appear together as a mixture in contaminated groundwater. Interactions between bacteria and BTEX mixtures may be complex and site specific (Alvarez and Vogel, 1991). Biodegradation of monoaromatic hydrocarbons (e.g., toluene, ethylbenzene) also occurs under anaerobic nitrate reducing conditions. Benzene appears to be more recalcitrant under anaerobic conditions than other monoaromatics. The biodegradation of aromatic rings has also been observed under iron reducing,

sulfate reducing, and methanogenic conditions (Edwards et al., 1992; Grbic-Galic and Vogel, 1987; Lovley et al., 1989).

Simple nonhalogenated hydrocarbon solvents such as ethanol, methanol, and acetone are easily biodegraded at dilute concentrations (e.g., 0.1%). The cyclical ethers, and ether structures in general, are typically recalcitrant. However, the biodegradation of furans and 1,4-dioxane has been reported, although biochemical mechanisms and the distribution of this ability in the environment remains unknown. Space limitations prohibit an exhaustive review of the biodegradation of all large production organics which might be found in groundwater, but several more complete references are available (Gibson, 1984; Howard, 1989; Leahy and Colwell, 1990; Young and Cerniglia, 1997).

26.1.3.2 Halogenated Compounds

Halogenated organic compounds are among the most problematic environmental pollutants encountered (Chaudhry and Chapalamadugu, 1991). Halogenation typically increases environmental stability and toxicity, and significantly alters the kinds of biochemical and chemical reactions compounds undergo (Vogel et al., 1987). Halogenation increases the oxidation state of a compound relative to analogous nonhalogenated compounds, and this significantly influences biodegradation as well. Several reviews of microbial transformation of halogenated compounds exist (Fetzner and Lingens, 1994; Mohn and Tiedje, 1992; Neilson, 1990). Biodegradation of halogenated compounds can be considered from two perspectives, bioenergetics and biochemical mechanisms. In general, very oxidized compounds are thermodynamically less favorable electron donors than reduced compounds, and therefore, as the degree of halogenation (and oxidation) increases compounds have fewer and fewer electrons to give up as electron donors and they potentially would yield correspondingly less energy when microorganisms oxidize them. Alternatively, polyhalogenated compounds are potentially good electron acceptors (Figure 26.3). Halogenated compounds acting as electron acceptors can undergo a reaction called reductive dehalogenation, in which two electrons are transferred to the compound, the halogen leaves as a halide ion, and is replaced by a hydrogen atom. Depending on the degree of halogenation, and the type of environment and microorganisms present where it is found, halogenated compounds may be used as either electron donors or electron acceptors, with the more halogenated compounds making better electron acceptors and the less halogenated compounds making better electron donors. Consideration of thermodynamics suggests that sequential anaerobic dechlorination followed by aerobic biodegradation would be successful.

Halogenated organics influence biodegradation mechanistically because of the large atomic size of halogens relative to hydrogen (which halogens usually replace), halogen electronegativity, and the strength of the carbon–halogen bond. Halogen size may prevent biochemical reactions simply due to steric hindrance. Halogen electronegativity causes charge separation in bonds and may result in dipole moments in molecules, profoundly affecting chemical reactivity. For example, halogen substitution may result in compounds more susceptible to nucleophilic substitution reactions (e.g., hydrolysis), whereas many oxygenase enzymes are strong electrophiles which are better suited to attack saturated nonhalogenated hydrocarbons. Strong bonds require large activation energies for cleavage and may prevent reactions from occurring. For example, the carbon–fluorine bond is exceptionally strong and is rarely broken during biological processes.

26.1.3.3 Halogenated Aliphatic Solvents

The halogenated aliphatic solvents which are most commonly found in groundwater are the chlorinated methanes (e.g., carbon tetrachloride [CCl_4]) and chlorinated ethenes (tetrachloroethene or "perc" [PCE] and trichloroethene [TCE]) (Vogel et al., 1987; Schwarzenbach et al., 1993). Under anaerobic conditions, PCE can undergo stepwise reductive dechlorination to TCE, then dichloroethene (DCE), chloroethene (i.e., vinyl chloride [VC]), and finally ethene. The rates of dechlorination tend to be proportional to the number of chlorines, thus PCE dechlorinates faster than dichloroethene. In groundwater systems where PCE and/or TCE are undergoing reductive dechlorination, vinyl chloride often accumulates for reasons that are currently unknown, but possibilities include thermodynamic, kinetic, or toxic limitations. Since vinyl chloride is a known human carcinogen, the presence of this metabolite is extremely undesirable.
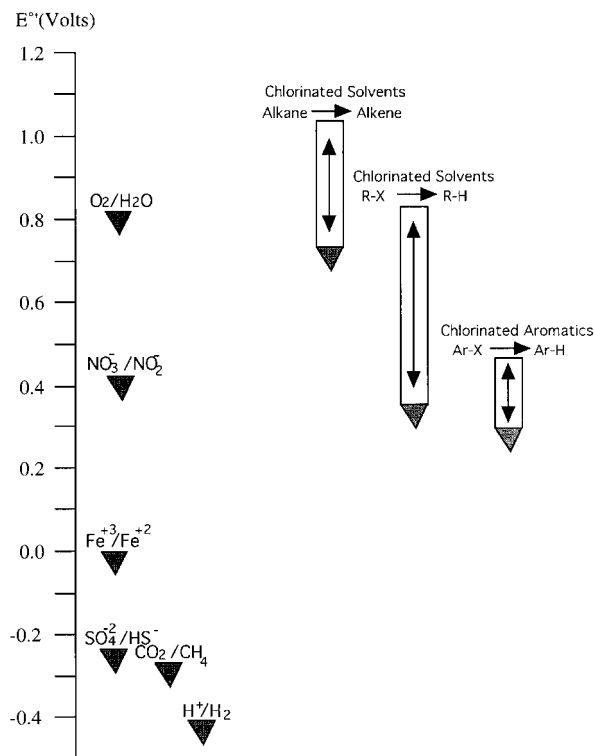
FIGURE 26.3 Redox potentials of biological electron acceptors and chlorinated compounds (Data from Brock, T. D., M. T. Madigan, J. M. Martinko, and J. Parker. 1997. *Biology of Microorganisms.* Prentice Hall, Englewood Cliffs, NJ., Dolfing, J. and Harrison, B. K. 1992. Gibbs free energy of formation of halogenated aromatic compounds and their potential role as electron acceptors in anaerobic environments. *Environ. Sci. Technol.* 26:2213-2218; Holmes, D. A., Harrison, B. K., and Dolfing, J. 1993. Estimation of Gibbs free energies of formation for polychlorinated biphenyls. *Environ. Sci. Technol.* 27:725-731; Schwarzenbach, R. P., Gschwend, P. M., and Imboden, D. M. 1993. *Environmental Organic Chemistry.* John Wiley & Sons, New York; Thauer, R. K., Jungermann, K., and Decker, K. 1977. Energy conservation in chemotrophic anaerobic bacteria. *Bacteriol. Rev.* 41:100-180; Vogel, T. M., Criddle, C. S., and McCarty, P. L. 1987. Transformations of halogenated aliphatic compounds. *Environ. Sci. Technol.* 21:722-736.)

Aerobically, TCE, DCE, and VC, but not PCE, can be cometabolized by certain bacteria with monooxygenase enzymes. This phenomena can potentially be exploited for use in the restoration of contaminated aquifers. These bacteria require specific growth substrates (e.g., methane or phenol) which induces synthesis of the monooxygenase enzymes which act on the chlorinated compounds. The metabolites of (at least TCE) cometabolism are potentially toxic and kill the cells mediating the reaction. Since the growth substrate and the cometabolized pollutant both compete for the same reactive enzyme site, a balance must be achieved between maintaining sufficient growth rates and acceptable degradation rates. Optimizing growth with excessive substrate could inhibit degradation of the pollutants, while low substrate concentrations may not induce the monooxygenase enzymes and biomass could be lost due to the production of suicide metabolites. For effective biodegradation, cell growth and pollutant biotransformation could potentially be separated in either space or time. This could be accomplished in groundwater by alternating periods of growth enhancement (adding growth substrate) with periods of starvation (Alvarez-Cohen, McCarty, 1991; Nelson et al., 1987).

26.1.3.4 Halogenated Aromatic Compounds

The biodegradation of halogenated aromatic compounds is highly dependent on the position and number of halogen substituents (Reineke and Knackmuss, 1988). It is useful to subdivide halogenated aromatics into two groups, nonionizable compounds and ionizable compounds. PCBs and chlorobenzenes are examples of nonionizable halogenated aromatic compounds found in the environment. In general, increased halogen substitution results in greater hydrophobicity. Therefore, the more soluble, less chlorinated PCB and chlorobenzene congeners are a greater threat to migrate in groundwater relative to more highly chlorinated congeners. The biodegradation of PCBs and chlorobenzenes is highly dependent on the degree of chlorination as well (Bedard and Haberl, 1990; Furukawa et al., 1978). PCBs and chlorobenzenes can be aerobically degraded similar to nonhalogenated aromatics, however, chlorine substitution often inhibits one or more enzymatic reactions. Thus, less chlorinated congeners are significantly more easily biodegraded by aerobic microorganisms than highly chlorinated compounds. Anaerobically, highly chlorinated PCBs and chlorobenzenes undergo microbially mediated reductive dechlorination to less chlorinated congeners which could potentially be biodegraded aerobically, but tend to persist under anaerobic conditions (Abramowicz, 1990). A field demonstration of PCB bioremediation was recently described by Harkness et al. (1993) during which it was found that bioavailability of PCBs severely limited biodegradation.

The behavior of ionizable aromatic compounds, such as phenols, anilines, and benzoates, is dependent on groundwater pH. For example, at a typical pH of 7, greater than 99% of the dissolved pentachlorophenol (PCP) (pKa = 4.75) will exist in the more soluble, less hydrophobic, ionized form. Unlike most PCB and chlorobenzene congeners, the complete biodegradation of PCP to CO_2 occurs under both aerobic and anaerobic conditions. In the better-understood aerobic process, a single microorganism can catabolize PCP for energy and carbon. Anaerobically, PCP is sequentially dechlorinated to phenol, which can be further degraded to methane and CO_2. Relatively little is known about the microbiology and biochemistry of anaerobic PCP biodegradation (Haggblom and Valo, 1995; McAllister et al., 1996).

26.1.4 Limits to Biodegradation

It should be mentioned that the environmental conditions which influence biodegradation and the limits to biodegradation are intrinsically related, and therefore, should not necessarily be viewed as separate topics. The most extreme limit to biodegradation is the absence of any known biochemical mechanism for the transformation of a specific compound (McCarty and Semprini, 1993). Discoveries of new transformations and microbial evolution of new enzymes will continue to challenge this limitation (Shannon and Unterman, 1993). Biodegradation potential can often be predicted from structure-activity models and a review of the biodegradation literature (Huesemann, 1995; Scow, 1990). However, since microbial distribution and environmental conditions are extremely heterogeneous, actual site-specific assays, such as laboratory treatability studies, provide the most reliable evidence that *in situ* biodegradation is possible.

26.1.4.1 Toxic Environmental Conditions

As discussed previously, microorganisms have adapted to life in many naturally occurring extreme environments. However, most of these organisms are autotrophs which will not significantly biodegrade organic pollutants. Extreme toxic conditions resulting from human activities and chemical releases to the environment are more difficult to ameliorate. The addition of oxygen and nutrients, as well as a moderate ability to influence pH, comprise the options currently available to influence environmental conditions. High concentrations of pollutants may cause toxic conditions and prevent biodegradation. Under these circumstances, removal of source material might lower groundwater concentrations to nontoxic levels at which biodegradation could occur. Moderate environmental conditions at near neutral pH, with adequate moisture, nutrients, and electron acceptor are the most likely to promote biodegradation.

26.1.4.2 Bioavailability and Mass Transfer Limitations

Most bacteria take up dissolved nutrients and substrate from the surrounding water. Phenomena which lower the dissolved concentrations or dissolution rate of compounds will limit biodegradation. Sorption of hydrophobic compounds to soils results in significant mass fractions of these compounds being unavailable to microorganisms. Biodegradation of the soluble phase results in desorption to reestablish the phase distribution equilibrium. In this situation, desorption rates may control biodegradation kinetics (Bosma et al., 1997). Long-term exposure of hydrophobic compounds to soils often results in a fraction of compound which remains unavailable for biodegradation due to mechanisms which are not yet fully understood (Hatzinger and Alexander, 1995). Therefore, it should be clear that *in situ* biodegradation kinetics may reflect processes other than microbial metabolism such as desorption of the pollutants, pollutant transport, and availability of electron acceptors.

26.1.4.3 Absence of Organisms

In some cases novel pollutant-degrading microorganisms have been isolated and cultured in laboratories, while the widespread existence of these microorganisms in the environment has not been observed. The introduction of novel nonindigenous pollutant-degrading microorganisms to resolve this situation has several potential problems. The survival and effectiveness of nonindigenous microorganisms *in situ* has rarely been carefully documented. Current research is attempting to assess the transport and survival of introduced organisms in contaminated zones. However, the ability to introduce and effectively distribute pollutant degrading microorganisms into groundwater currently does not exist. The well-known ecological problems caused by the introduction of invasive nonindigenous macroscopic organisms (e.g., zebra mussels, invasive weeds, rabbits) has provoked considerable evaluation of the practice of introducing non-native organisms. The ecological effect of introducing invasive microorganisms is a relatively unexplored topic.

26.1.4.4 Mixed Wastes/Metals

Mixtures of different chemicals may influence biodegradation in several ways. Microorganisms have biochemical mechanisms for optimizing energy production by specifically utilizing preferred substrates, while repressing catabolism of other substrates. Thus, compounds which are readily biodegraded when present individually may persist when present in a mixture. Studies of BTEX biodegradation have revealed that substrate interactions are important and are likely to be complex and diverse (Alvarez and Vogel, 1991). Compounds which require different redox conditions for biotransformation may further complicate the biodegradation of mixtures. For example, aerobic conditions are desirable for petroleum hydrocarbon biodegradation, but anaerobic conditions are necessary for reductive dechlorination of chlorinated solvents. Thus it may not be possible to have optimum biodegradation conditions for all compounds present in a mixture.

26.2 Quantitative Description of Reactive Transport and Biodegradation in Porous Media

A quantitative understanding of coupled chemical transformations, sorption, and transport processes is required to predict, control, and possibly optimize biodegradation in environmental problems in a natural or engineered setting. In this section, the basic mass conservation statements and constitutive models that provide a framework for reactive transport modeling are presented. Applications to biodegradation are discussed, and example calculations are presented to illustrate some of the interactions that can be expected among the transport and reaction mechanisms. Although such calculations are inevitably based on idealized constitutive behavior, they enable determining the sensitivities of performance measures to the constitutive parameters/models and can help in prioritizing data collection.

26.2.1 Conservation Statement

Multiple species undergoing chemical reactions and transport in fluid flowing through porous media are considered here. The subscript m denotes the m*th* species. Some further notation:

c_m (M/L^3):   Mass per unit volume of fluid.

s_m (M/M):   Adsorbed mass per unit mass of porous media solids.

n (L^3/L^3):   Pore volume per unit bulk volume of porous media. For saturated media this denotes the volume of fluid per unit bulk volume.

ρ_b (M/L^3):   The mass of solids per unit bulk volume of porous media.

F_{mi} (M/L^2/T): Time rate of transport of mass per unit area of porous medium in the i*th* direction.

r_m (*M/L^3/T*):   The rate of introduction of mass per unit bulk volume due to chemical transformations other than sorption, and/or externally imposed source/sink terms.

The mass conservation statement for the m*th* species in saturated porous media is given by

$$\frac{\partial\left(nc_m + \rho_b s_m\right)}{\partial t} + \frac{\partial F_{mi}}{\partial x_i} = r_m \tag{1}$$

Assuming n and ρ_b to be time invariant, the mass conservation statement becomes

$$n\frac{\partial c_m}{\partial t} + \rho_b\frac{\partial s_m}{\partial t} + \frac{\partial F_{mi}}{\partial x_i} = r_m \tag{2}$$

A sum over the repeated index i and j is implied in this chapter.

26.2.2 Constitutive Models

26.2.2.1 Flux

For the dissolved or suspended species undergoing transport, the flux, accounting for advective transport associated with the seepage velocity v_i, and the dispersion associated with the pore-scale velocity variability and molecular diffusion is given by

$$F_{mi} = n\left(v_i c_m - d_{mij}\frac{\partial c_m}{\partial x_j}\right) \tag{3}$$

A sum over the repeated index j is implied in this representation of mass flux (Chapter 14 provides a discussion of transport mechanisms). The constitutive assumption is that of a Fickian local dispersive flux, with the associated local-dispersion coefficients for the m*th* species being d_{mij}. The local dispersion coefficients are related to the seepage velocity field v_i, molecular diffusion coefficients for the species in porous medium d_m (taking into account the tortuosity), and the local dispersivities α_{mL}, α_{mT}:

$$d_{mij} = \left(\alpha_{mT}\sqrt{v_i v_i} + d_m\right)\delta_{ij} + \frac{\left(\alpha_{mL} - \alpha_{mT}\right)}{\sqrt{v_i v_i}}v_i v_j \tag{4}$$

The local dispersivities are typically taken to be species independent, however, that assumption has not been directly tested extensively for biodegradable species.

Not all species may undergo transport. For example, if the biomass is attached to the soil particles and does not undergo advection or local dispersion, then F_{mi} may be set to zero in the conservation statement for the biomass.

26.2.2.2 Sorption

A model for $\partial s_m/\partial t$ needs to be specified to complete the representation of the effects of sorption in the mass conservation equation (2).

26.2.2.3 Equilibrium Model

In batch tests an equilibrium relationship between the dissolved/suspended and adsorbed phase concentrations, of the general form, $s_m = N_m(c_m)$ may be established. This equilibrium relationship between s_m and c_m may also be influenced by the concentration of other species (c_n, $n \neq m$), depending on the specific chemical nature of the contaminant and soil being considered. In that case $\partial s_m/\partial t$ will be related to the time derivatives of all those species. For the case in which the other concentrations that control N_m are not time variant, we may write $\partial s_m/\partial t = (dN_m(c_m)/dc_m)\partial c_m/\partial t$, and therefore the first two terms on the left-hand side of the mass balance (2) may be written as

$$n\frac{\partial c_m}{\partial t} + \rho_b\frac{\partial s_m}{\partial t} = nR_m\frac{\partial c_m}{\partial t}$$

$$R_m = 1 + \frac{\rho_b}{n}\frac{dN_m(c_m)}{dc_m}$$

(5)

Compared to the concentration of a transported species that is not adsorbed, the propagation of the aqueous (dissolved/suspended) concentration of an adsorbed species, in the equilibrium sorption model, is therefore slowed down by a factor called the retardation factor R_m.

For a linear equilibrium isotherm, $s_m = K_d c_m$, where $K_d = dN_m(c_m)/dc_m$ is the distribution or partition coefficient. For some organic compounds, the partition coefficient is expressed as a product of the partition coefficient with respect to organic carbon and the fraction of organic carbon in the soil. In general, the function N_m needs to be experimentally established over the range of concentrations of interest, and a concentration-dependent retardation factor can be introduced in numerical solutions of the transport equation.

26.2.2.4 Nonequilibrium Model

That the rate of sorption is controlled by differences between the sorbed and aqueous concentrations or their nonlinear functions can be accounted for in a general nonequilibrium model, $\partial s_m/\partial t = A_m(s_m) - B_m(c_m)$, where A_m and B_m are nonlinear functions consistent with the equilibrium isotherm $s_m = N_m(c_m)$, i.e., $A_m^{-1}(B_m(c_m)) = N_m(c_m)$. Nonlinear and nonequilibrium sorption effects have been illustrated in the experimental study of Streck et al. (1995). A linear nonequilibrium model has been used for TCE (Roberts et al., 1989)

$$\frac{\partial s_m}{\partial t} = \beta\left(\frac{dN_m(c_m)}{dc_m}c_m - s_m\right)$$

(6)

where β (T^{-1}) is a rate constant. Karickhoff (1980), Valocchi (1985), Wu and Gschwend (1986), Brusseau and Rao (1989), and Ball and Roberts (1991) present inferences and implications of the time-scales in the nonequilibrium sorption process.

26.2.2.5 Transformation Kinetics

In the mass balance equation r_m is the rate at which mass of species m is being introduced per unit bulk volume of the porous media. Therefore if the species m is being depleted r_m is negative and vice versa. Without any external source or sink of a species, r_m represents the effects of chemical transformations (other than sorption). In specifying a model for the chemical kinetics, through r_m, it needs to be remembered that the adsorbed and aqueous phases of the species may undergo reactions at different

rates. For example, if the adsorbed and aqueous phases undergo first-order decay with different decay constants, κ_1 and κ_2, then $r_m = -n\kappa_1 c_m - \rho_b\kappa_2 s_m$. While such a linear decay model is simple and requires few constitutive parameters, it simply does not apply to the situation when a reaction requires the interaction of more than one species. An elementary example of such an irreversible reaction is a bimolecular second-order reaction, which may be represented as $r_m = -n\kappa c_1 c_2$, assuming that the aqueous phase of the two-species system is involved in the reaction. In many situations the consideration of multiple species is necessary in specifying reaction kinetics. Extensive laboratory experiments are required to define the biodegradation kinetics, which will be contaminant and soil specific. Borden and Bedient (1986) and Baveye and Vallochi (1989) provide a comparison of different approaches for representing biodegradation kinetics in porous media, and Dykaar and Kitanidis (1996) have quantified the implications of diffusive transport limitations at the pore-scale. Examples of experimentally inferring constitutive parameters in biodegradation kinetics of BTX contaminants are provided in MacQuarrie et al. (1990) and Alvarez et al. (1991), and by Walton et al. (1996) and references therein.

Here is a biodegradation kinetics model that has been previously applied to simulate the biodegradation by Roberts et al. (1989), Semprini and McCarty (1991 and 1992), and MacDonald (1995) for the biodegradation of TCE:

m = 1: TCE

$$\frac{\partial s_1}{\partial t} = \beta\left(K_d c_1 - s_1\right)$$

$$r_1 = -n\kappa_1 c_4\left(\frac{c_1}{\chi_1 + \left(\chi_1/\chi_3\right)c_3 + c_1}\right)\left(\frac{c_2}{\chi_2 + c_2}\right)$$

m = 2: Electron acceptor

$$s_2 = 0$$

$$r_2 = -n c_4\left(\frac{c_2}{\chi_2 + c_2}\right)\left[\kappa\Phi\left(\frac{c_3}{\chi_3 + c_3}\right) + \omega\right]$$

m = 3: Electron donor

$$s_3 = 0$$

$$r_3 = -n\kappa c_4\left(\frac{c_2}{\chi_2 + c_2}\right)\left(\frac{c_3}{\chi_3 + c_3}\right)$$

m = 4: Biomass

$$F_4 = 0$$

$$s_4 = 0$$

$$r_4 = n c_4\left(\frac{c_2}{\chi_2 + c_2}\right)\left[\kappa Y\left(\frac{c_3}{\chi_3 + c_3}\right) - b\right]$$

In specifying this four-species model for the biodegradation of TCE, the symbols other than the dependent variables c_m and s_m represent constitutive parameters that need to be experimentally determined in batch and column tests. In the presence of significant amounts of electron donor and acceptor, this model has been shown to result in large biomass concentrations (MacDonald, 1995), which can significantly decrease the permeability of the porous medium (biofouling, discussed in Taylor and Jaffe, 1990, and Vandevivere and Baveye 1992a,b), and may also require considering the influence of shear stress on the biomass (Rittmann, 1982) and transport mechanisms for the biomass.

26.2.3 Illustration of Reactive Transport in Heterogeneous Formations

In any field application the effects of small-scale heterogeneity are potentially important, as has long been established for the spreading of solutes in the field (Gelhar et al., 1992, discussed in Chapter 14). In addition to the enhanced spreading (macrodispersion) of contaminants caused by the field heterogeneity, in the context of reactive transport and biodegradation, two additional effects of heterogeneity are catalogued here: (1) macrosegregation and transport limitations and (2) influence of a correlation between sorption and hydraulic conductivity. These effects are illustrated for flow in a saturated heterogeneous hydraulic conductivity field (Figure 26.4), described previously in Chapter 14.
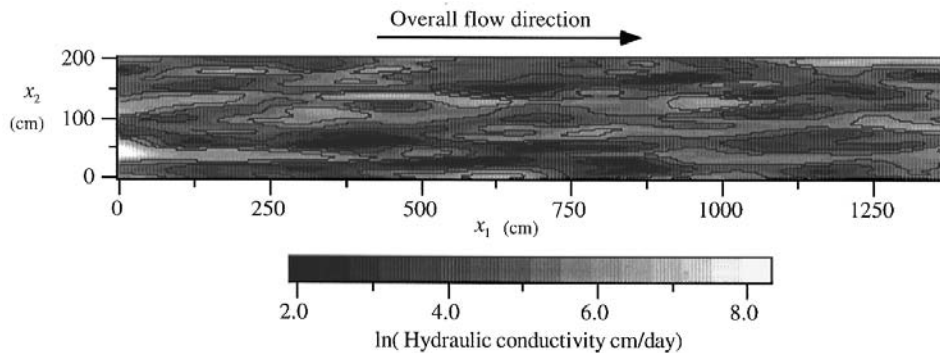
FIGURE 26.4 Heterogeneous hydraulic conductivity field.

The specific example employed will be a simplified one in which the biomass concentration (M) and electron donor concentration are assumed to be constant, and the biodegradable contaminant is a benzene solution, which undergoes linear equilibrium sorption and aerobic biodegradation, at a rate controlled by possibly both the dissolved benzene concentration (c_1) and oxygen concentration (c_2):

m = 1: Benzene

A linear equilibrium sorption model is assumed, i.e.,

$$R_1 \equiv \left[1 + \frac{\rho_b k_d}{n} \right]$$

The biodegradation kinetics are assumed to be described by a dual Monod model:

$$r_1 = -n\kappa M \left(\frac{c_1}{\chi_1 + c_1} \right) \left(\frac{c_2}{\chi_2 + c_2} \right)$$

An initially Gaussian pulse of benzene (with an x_1 radius of gyration of 10 cm) is introduced, uniformly over x_2 (Figures 26.4 and 26.5).
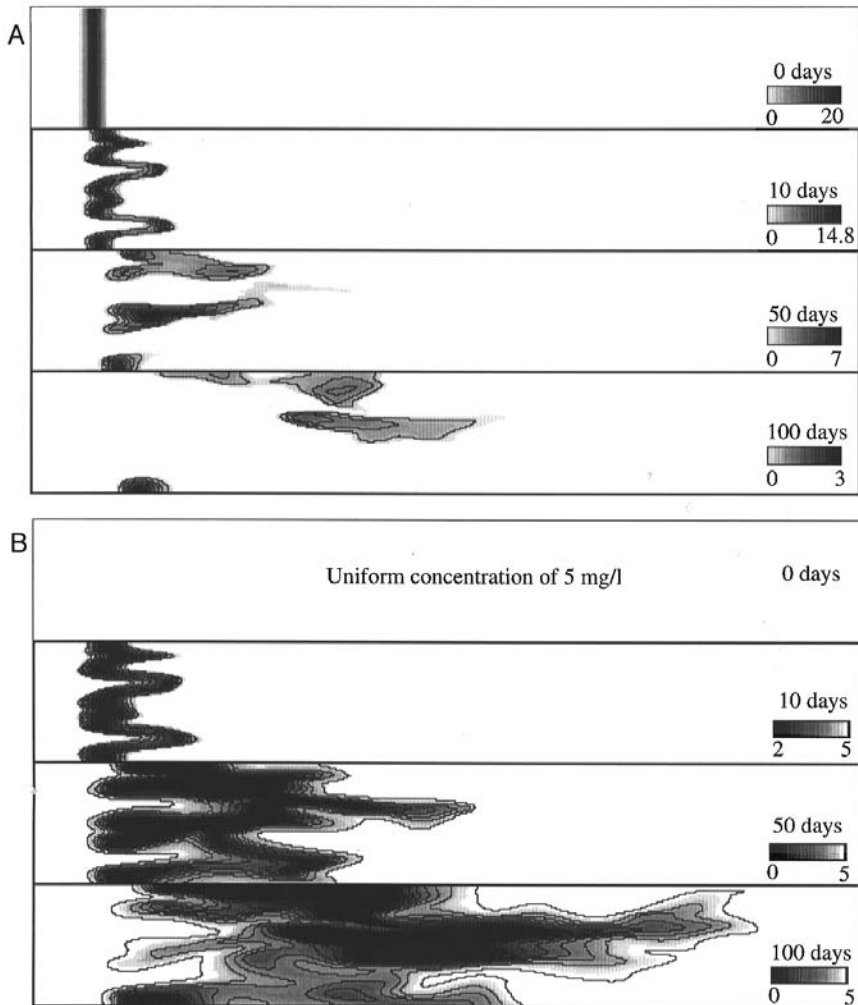
FIGURE 26.5 Concentration fields, A) benzene concentration (mg/l), B) dissolved oxygen concentration (mg/l).

m = 2: Oxygen

Oxygen does not undergo sorption ($s_2 = 0$) . The oxygen consumption rate per unit bulk volume of porous media is assumed to be

$$r_2 = -Wr_1$$

Both oxygen and benzene undergo advection and dispersive flux and the dispersivities are taken to be $\alpha_L = 1.0$ cm and $\alpha_T = 0.5$ cm, for both the species, and the molecular diffusion coefficient, taking into account the influence of tortuosity, is taken to be 0.5×10^{-5} cm^2/sec.

The hydraulic conductivity and flow field is the same as that presented in Chapter 14, to illustrate the influence of heterogeneity on transport (Figure 26.4). The hydraulic conductivity, K, is modeled as a random field-$\ln K = \ln K_G + f$, where f is a spatially correlated random field. The additional features here are the presence of two species, the nonlinear nature of the reaction rate, and the effects of sorption on benzene.

Parameters pertinent to transport and reactions:

Porosity n = 0.3
Benzene half saturation constant χ_1 = 0.5 mg/L
Dissolved oxygen half-saturation constant χ_2 = 0.1 mg/L
Microbial mass concentration M = 0.2 mg/L
Reaction rate parameter κ = 0.5 day^{-1}
Stoichiometric factor W = 3.13
Initial maximum benzene concentration = 20 mg/L
Initial background dissolved oxygen concentration = 5 mg/L

Two different descriptions of sorption are examined.

26.2.3.1 Case 1: Constant Partition Coefficient

The retardation factor for benzene is taken to be a constant, R_1 = 1.4.

26.2.3.2 Case 2: Partition Coefficient Negatively Correlated with Hydraulic Conductivity

Paired data of partition coefficients and hydraulic conductivity have not been widely collected. However, existing data for the sorption of inorganic solutes suggest a modest negative correlation between these parameters (Robin et al., 1991; Foster-Reid, 1994), which can be attributed to the relationship of both the partition coefficient and hydraulic conductivity to the porous material grain-size, which may be a surrogate for the capacity of a material to adsorb certain contaminants. For organic contaminants it may also be argued that the fraction of organic carbon is greater in the finer material (with a greater clay content), which could manifest in a negative correlation between the partition coefficients with hydraulic conductivity. To illustrate the impact of a negative correlation between the retardation factor and the hydraulic conductivity, the retardation factor is assumed to be perfectly negatively correlated with the log-hydraulic conductivity: R_1 = a lnK_G – b f. With a = 0.304 and b = 0.28 the mean retardation factor is 1.4 (the same as Case 1), and its coefficient of variation is 0.2.

The evolution of the benzene concentration and dissolved oxygen concentration (calculated by a standard explicit in time, second order centered in space, finite-difference scheme) are shown in Figure 26.5. As the benzene is retarded the dissolved oxygen moves faster than the benzene. This difference in effective velocities of dissolved oxygen and benzene controls the bulk advective mixing of the two species, and therefore importantly controls the overall biodegradation rate (Miralles-Wilhelm et al., 1994). As a consequence of this difference, the dissolved oxygen plot shows a large area over which oxygen is depleted (Figure 26.5) relative to its original background value.

26.2.3.3 Macrosegregation and Transport Limitations

Figure 26.6 shows x_2 transects of benzene and oxygen concentration at the center of mass of the benzene solute body at 10 days. Note that the zones over which the benzene concentration is greater than average, correspond to the zones over which the dissolved oxygen is smaller than average. The benzene concentration and dissolved oxygen concentration are negatively correlated, i.e., the two species are segregated. The physical explanation for this negative correlation is simple — as the benzene channels through the high conductivity zones, it displaces the oxygen in those zones. This negative correlation has an important consequence for overall transformation rates in heterogeneous field settings. Due to segregation, the actual biodegradation rate can be much smaller than what would be inferred by substituting x_2 averages of c_1 and c_2 in the expression for the biodegradation kinetics. Therefore in assessing biodegradation rates in natural flowing waters in the subsurface, it needs to be recognized that the transformation rates pertinent to the usual simplified depth averaged descriptions of concentrations (neglecting heterogeneity) can be much smaller than those implied by the chemical kinetic parameter inferred in well-mixed batch tests (Kapoor et al., 1997). Molz and Widdowson (1988) have previously shown sharp gradients in the dissolved oxygen concentration, in the field and pointed out the inability of depth-averaged macroscopic descriptions of biodegradation to take into account the influence of these variations on the biodegradation process.
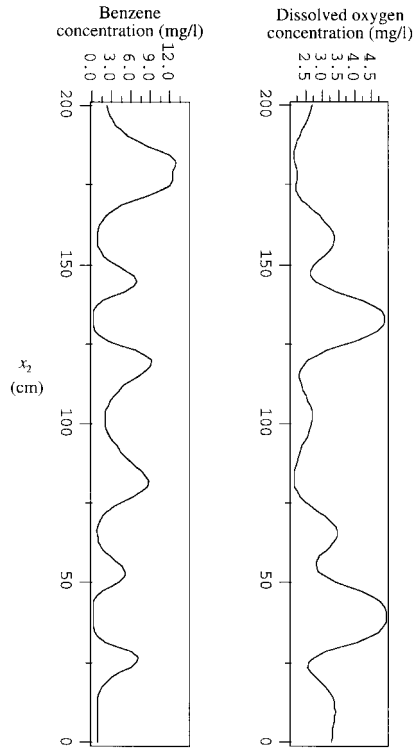
FIGURE 26.6 x_2 transect of concentration. The x_1 location corresponds to the center of mass of benzene. Higher than average values of benzene concentration occur together with lower than average values of the dissolved oxygen oxygen concentration. This macrosegregation occurs due to the channeling of benzene in the high conductivity zones and causes the reaction rate to be smaller than what would be inferred based on x_2 averaged concentrations.

26.2.3.4 Consequences of a Negative Correlation Between Retardation Factor and Hydraulic Conductivity

It has been theoretically predicted by Garabedian et al. (1988), and Miralles-Wilhelm and Gelhar (1996) that a negative correlation between hydraulic conductivity and retardation factor can result in enhanced spreading of the solute body, and field experiments at Cape Cod (Garabedian et al., 1988; Leblanc et al., 1991; Gelhar, 1993) document an enhanced spreading for the sorbing tracer (without any biodegradation effects).

The numerical simulation (Case 2) effected an idealized perfect negative correlation between the retardation factor and hydraulic conductivity to illustrate the theoretical prediction of enhanced spreading. Figure 26.7 shows the evolution of the x_1 second spatial moment of the benzene solute body for both the cases. More spreading occurs in Case 2 than in Case 1, with a constant retardation factor. A simple physical explanation for this is as follows. The contaminant partitions between aqueous and sorbed phases. If the spatial locations of larger-than-average partition coefficients also correspond to the lower-than-average velocities, then the effective range of velocities experienced by the aqueous phase is amplified by the negative correlation. Therefore, sorption can also profoundly alter the spreading characteristics of solute bodies, and in multispecies transport, the field-scale dispersion of sorbed and nonadsorbed tracers can be significantly different (Garabedian et al., 1988; Miralles-Wilhelm et al., 1994; Miralles-Wilhelm and Gelhar, 1996). Similarly, if the biodegradation kinetic parameters are spatially variable, their correlations with physical parameters can have important macroscopic consequences. Therefore, to understand biodegradation phenomena in the field, it is important to concurrently infer both biochemical constitutive parameters and physical constitutive parameters of porous media in which biodegradation is of interest.
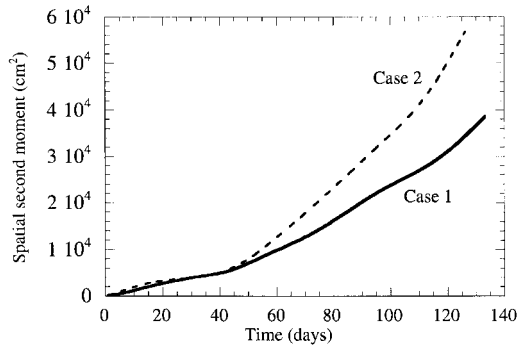
FIGURE 26.7 Spatial second moment evolution. Spatial heterogeneity of the partition coefficient and its negative correlation with hydraulic conductivity (case 2) has been previously predicted to cause an enhanced spreading (Garabedian et al., 1988), compared to the case of a constant partititon coefficient (case 1) as illustrated in this plot of numerical results.

26.3 Field Applications of Bioremediation

Bioremediation is the utilization of naturally occurring microbial biodegradation processes to restore a site to a nonhazardous condition. A detailed discussion of bioremediation application and design is presented elsewhere (Baker and Herson, 1994; Cookson, 1995; Flathman et al., 1994; King et al., 1992; Riser-Roberts, 1992). There are many advantages to using bioremediation. Since pollutant destruction occurs *in situ*, the potential liability and environmental risk associated with the removal, handling, transport, and storage of hazardous contaminated materials is eliminated. An additional important advantage of bioremediation is that it is often the most economical solution available. Successful bioremediation requires a thorough site investigation and evaluation of treatment options. A site history is compiled which should include property uses, chemicals stored, and location of utilities and buildings, as well as the location of nearby wells. Available information about local hydrogeology, geology, and topography should also be included. Soil and groundwater samples are obtained to identify the contaminants present and estimate their concentration and distribution. Soil borings are made to determine the local hydraulic gradient and sometimes in conjunction with pumping tests, the hydraulic conductivity is estimated. Whenever possible, an assessment of the local groundwater geochemistry should be made by measuring alternative electron acceptor concentrations (nitrate, sulfate), pH, and aquifer buffering capacity. Data from samples taken outside the zone of contamination are useful for assessing background microbiological and geochemical conditions. Laboratory studies are currently the most reliable method for assessing the biodegradation potential of the indigenous organisms. From the assembled information, predictions of contaminant migration and biodegradation can be used to evaluate the current hazard posed and potential remediation options.

26.3.1 Intrinsic Bioremediation

Intrinsic bioremediation is the natural *in situ* biodegradation of pollutants without the engineered manipulation of environmental conditions. Intrinsic bioremediation is appropriate when a site investigation shows that natural biodegradation processes are sufficiently attenuating the migration of contaminants and there is little imminent risk or liability associated with the site. An ongoing monitoring program is essential to regularly evaluate the progress of intrinsic bioremediation until site closure.

26.3.2 Engineered Bioremediation

Engineered systems are installed when it is necessary to overcome some limitation to biodegradation, or when health and liability concerns make it desirable to accelerate naturally occurring processes. In current

practice the addition of some form of oxygen to the aquifer is usually the primary objective. Nutrients (N&P) are often added as well, usually without site-specific evidence of a nutrient deficiency, but because it is an easy and inexpensive option. Pumping wells, injection wells, infiltration galleries, vacuum pumps, and compressors may be installed in various combinations and configurations depending on the design objective. In addition to adding an electron acceptor or nutrients to an aquifer, it may be desirable to attempt to contain the contaminant plume, install a groundwater recirculation system, or combine bioremediation with other remediation processes.

26.3.3 Case Histories

The objective of listing a few case histories is to provide the reader with references to further study well-documented examples of applied bioremediation. Therefore, only brief summaries of several types of bioremediation field studies are presented, and the interested reader is encouraged to obtain the cited references for detailed descriptions.

26.3.3.1 Aerobic Petroleum Hydrocarbon Bioremediation

The aerobic bioremediation of motor oil, diesel fuel, and gasoline which had leaked from a used oil sump into groundwater was described by Nelson et al., (1994). Hydrogen peroxide and nutrients were infiltrated into the contaminated zone through an installed groundwater recirculation system. A vapor extraction system was combined with bioremediation to enhance hydrocarbon removal in the unsaturated zone.

26.3.3.2 BTEX Bioremediation Under Denitrifying Conditions

The use of nitrate as an electron acceptor for the bioremediation of JP-4 jet fuel was described by Hutchins et al., (1991). The system consisted of an infiltration gallery, interdiction wells, pumping wells, and monitoring wells. The bioremediation was effective in removing BTEX constituents, however, a significant amount of relatively insoluble alkanes remained as NAPL.

26.3.3.3 Aerobic Chlorinated Aliphatic Hydrocarbon Bioremediation

The feasibility of *in situ* methanotrophic co-oxidation of chlorinated ethenes was tested at Moffett Naval Air Station, CA (Roberts et al., 1990; Semprini et al., 1990). With sufficient oxygen and methane the methanotrophic bacteria cometabolized TCE, DCE, and VC in a two-meter-long biostimulated zone. Approximately 90% of the VC and 80 to 90% of the trans-DCE was biodegraded in this short distance. TCE and *cis*-DCE were more recalcitrant. However, longer residence times in biostimulated zones would improve the extent of removal.

26.3.3.4 Intrinsic Bioremediation

Intrinsic biodegradation processes have been responsible for the containment of petroleum hydrocarbons from a pipeline leak in Minnesota (Baedecker et al., 1993). Careful monitoring has confirmed that the soluble BTEX constituents are not migrating with the groundwater flow. This is due to the biodegradation rate being equal to the release rate of BTEX from the residual hydrocarbon mixture.

26.3.4 Emerging Technologies

Increased utilization of bioremediation will depend on technological advances which will remove limitations to when bioremediation can be reliably and effectively applied (Atlas, 1995). Advances in understanding novel bioremediation technologies have been made in several areas, for example, bioavailability, molecular biology, microbiology, and anaerobic biodegradation.

Microbially produced biosurfactants and synthetic surfactants increase the apparent solubility of hydrophobic compounds and could potentially reduce bioavailability limitations. However, much remains to be learned about the biodegradation of surfactant solubilized compounds and potential surfactant toxicity, as well as cost and effectiveness. A combined treatment process consisting of surfactant/solvent soil washing followed by bioremediation to remove residual contamination has potential for

the remediation of NAPL (nonaqueous phase liquids) contaminated sites. However, the microecological effect of the soil washing is unknown.

Advances in molecular biology have contributed much to our understanding of the biochemistry and microbiology of biodegradation, and will continue to do so. Use of molecular techniques is now commonplace in biodegradation research. The importance of understanding the biochemistry, genetics, and microbiology of biodegradation cannot be overestimated. In the coming years, molecular techniques for detecting and enumerating bacteria *in situ* will likely become routine. Genetic techniques are being used to construct novel biodegradation pathways, essentially creating organisms with new capabilities.

Additional examples of emerging technologies include water recirculation systems, oxygen-releasing peroxides, the use of zero valent metals to enhance reductive dechlorination, as well as advances in *in situ* measurement of physical, chemical, and biological parameters. Ongoing development of pollutant degrading thermophiles, alkaline-tolerant bacteria, and bacteria with membranes resistant to high solvent concentrations will ensure that biodegradation will continue to grow as a primary groundwater remediation option.

For Further Information

Brock (1997) provides an excellent comprehensive work on microbiology, covering introductory cell chemistry and cell biology as well as more advanced material.

Gottschalk (1986) is an outstanding reference that describes basic bacterial biochemistry.

National Research Council (1993), *In Situ Bioremediation. When does it work?* describes the fundamental principles and limitations of bioremediation.

References

Abramowicz, D. A. 1990. Aerobic and anaerobic biodegradation of PCBs: a review. *CRC Crit. Rev. Biotechnol.* 10:241-251.

Alexander, M. 1994. *Biodegradation and Bioremediation.* Academic Press, San Diego.

Alvarez, P. J. J. and Vogel, T. M. 1991. Substrate interactions of benzene, toluene, and para-xylene during microbial degradation by pure cultures and mixed culture aquifer slurries. *Appl. Environ. Microbiol.* 57:2981-2985.

Alvarez, P. J. J., Anid, P. J., and Vogel, T. M. 1991. Kinetics of aerobic biodegradation of benzene and toluene in sandy aquifer material. *Biodegradation.* 2:43-51.

Alvarez-Cohen, L. and McCarty, P. L. 1991. Product toxicity and cometabolic competitive inhibition modeling of chloroform and trichloroethylene transformation by methanotrophic resting cells. *Appl. Environ. Microbiol.* 57:1031-1037.

Atlas, R. M. 1995. Bioremediation. *Chem. Eng. News.* 73:32-42.

Bachofer, R., Lingens, F., and Schafer, W. 1975. Conversion of aniline into pyrocatechol by a Nocardia sp.: incorporation of oxygen-18. *FEBS Lett.* 50:288-290.

Baedecker, M. J., Cozzarelli, I. M., Eganhouse, R. P., Siegei, D. I., and Bennett, P. C. 1993. Crude oil in a shallow sand gravel aquifer III. Biogeochemical reactions and mass balance modeling in anaoxic groundwater. *Appl. Geochem.* 8:569-586.

Baker, K. H. and Herson, D. S. 1994. *Bioremediation.* McGraw-Hill, New York.

Ball, W. P. and Roberts, P. V. 1991. Long-term sorption of halogenated organic chemicals by aquifer material, 1, Equilibrium, *Environ. Sci. Technol.* 25, 1223-1236.

Baveye, P. and Valocchi, A. 1989. An evaluation of mathematical models of the transport of biologically reacting solutes in saturated soils and aquifers, *Water Resour. Res.* 25(6):1413-1421.

Bedard, D. L. and Haberl, M. L. 1990. Influence of chlorine substitution pattern on the degradation of polychlorinated biphenyls by eight bacterial strains. *Microb. Ecol.* 20:87-102.

Bitton, G., Farrah, S. R., Ruskin, R. H., Butner, J., and Chou, Y. J. 1983. Survival of pathogenic and indicator organisms in groundwater. *Groundwater.* 21:405-410.

Borden, R. C. and Bedient, P. B. 1986. Transport of dissolved hydrocarbons influenced by oxygen-limited biodegradation. 1. Theoretical development, *Water. Resour. Res.* 22(13):1973-1982.

Bosma, T. N. P., Middeldorp, P. J. M., Schraa, G., and Zehnder, A. J. B. 1997. Mass transfer limitation of biotransformation: quantifying bioavailability. *Environ. Sci. Technol.* 31:248-252.

Brock, T. D., Madigan, M. T., Martinko, J. M., and Parker, J. 1997. *Biology of Microorganisms.* Prentice Hall, Upper Saddle River, NJ.

Brusseau, M. L. and Rao, P. S. C. 1989. Sorption nonideality during organic contaminant transport in porous media, *Crit. Rev. Environ. Control.* 19:33-99.

Chapelle, F. H. 1993. *Groundwater Microbiology & Geochemistry.* John Wiley & Sons, New York.

Chaudhry, G. R. and Chapalamadugu, S. 1991. Biodegradation of halogenated organic compounds. *Microbiological Reviews.* 55:59-79.

Chen, C. I. and Taylor, R. T. 1995. Thermophilic biodegradation of BTEX by two *Thermus* species. *Biotech. Bioeng.* 48:614-624.

Cookson, J. T. Jr. 1995. *Bioremediation Engineering: Design and Application.* McGraw-Hill, New York.

Corapcioglu, M. Y. and Haridas, A. 1984. Transport and fate of microorganisms in porous media: a theoretical investigation. *J. Hydrol.* 72:149-169.

Corapcioglu, M. Y. and Haridas, A. 1985. Microbial transport in soils and groundwater: a numerical model. *Adv. Water Resour.* 8:188-200.

Dawson, M. P., Humphrey, B. A., and Marshall, K. C. 1981. Adhesion: a tactic in the survival strategy of a marine vibrio during starvation. *Current Microbiol.* 6:195-199.

Dolfing, J. and Harrison, B. K. 1992. Gibbs free energy of formation of halogenated aromatic compounds and their potential role as electron acceptors in anaerobic environments. *Environ. Sci. Technol.* 26:2213-2218.

Dolfing, J., Zeyer, J., Binder-Eicher, P., and Schwarzenbach, R. P. 1990. Isolation and characterization of a bacterium that mineralizes toluene in the absence of molecular oxygen. *Arch. Microbiol.* 154:336-341.

Dykaar, B. B. and Kitanidis, P. K. 1996. Macrotransport of a biologically reacting solute through porous media. *Water Resour. Res.* (32)2, 307-320.

Edwards, E. A., Wills, L. E., Reinhard, M., and Grbic-Galic, D. 1992. Anaerobic degradation of Toluene and Xylene by aquifer microorganisms under sulfate-reducing conditions. *Appl. Environ. Microbiol.* 58:794-800.

Evans, P. J., Mang, D. T., Kim, K. S., and Young, L. Y. 1991. Anaerobic degradation of toluene by a denitrifying bacterium. *Appl. Environ. Microbiol.* 57:1139-1145.

Evans, W. C. 1988. Anaerobic degradation of aromatic compounds. *Annu. Rev. Microbiol.* 42:289-317.

Fetzner, S. and Lingens, F. 1994. Bacterial dehalogenases: biochemistry, genetics, and biotechnical applications. *Microbiol. Rev.* 58:641-685.

Flathman, P. E., Jerger, D. E., and Exner, J. H. 1994. *Bioremediation — Field Experience.* Lewis Publishers, Boca Raton, FL.

Fletcher, M. and Marshall, K. C. 1982. Are solid surfaces of ecological significance to aquatic bacteria? *Adv. Microbial. Ecol.* 6:199-236.

Foster-Reid, G. H. 1994. *Variability of Hydraulic Conductivity and Sorption in a Heterogeneous Aquifer.* M.S. thesis, Mass. Inst. of Technol., Cambridge.

Fredrickson, J. K. and Onstott, T. C. 1996. Microbes deep inside the earth. *Scientific American.* 275:68-73.

Furukawa, K., Tonomura, K., and Kamibayashi, A. 1978. Effect of chlorine substitution on the biodegradability of polychlorinated biphenyls. *Appl. Environ. Microbiol.* 35:223-227.

Garabedian, S. P., Gelhar, L. W., and Celia, M. A. 1988. Large-scale dispersive transport in aquifers: Field experiments and reactive transport theory, Parsons Laboratory Report 315, Massachusetts Institute of Technology, Cambridge.

Gelhar, L. W., Welty, C., and Rehfeldt, K. R. 1992. A critical review of data on field-scale dispersion in aquifers. *Water Resour. Res.* 28(7):1955-1974.

Gelhar, L. W. 1993. *Stochastic Subsurface Hydrology.* Prentice Hall, Englewoods Cliffs, NJ.

Gibson, D. T. 1984. *Microbial Degradation of Organic Compounds*. Marcel Dekker, New York.

Grbic-Galic, D. and Vogel, T. M. 1987. Transformation of toluene and benzene by mixed methanogenic cultures. *Appl. Environ. Microbiol.* 53:254-260.

Gottschalk, G. 1986. *Bacterial Metabolism*. Springer-Verlag, New York.

Haggblom, M. M. and Valo, R. J. 1995. Bioremediation of chlorophenol wastes, in L.Y. Young and C.E. Cerniglia (eds.), *Microbial Transformation and Degradation of Toxic Organic Chemicals*. Wiley-Liss, New York, 389.

Harkness, M. R., McDermott, J. B., Abramowicz, D. A., Salvo, J. J., Flanagan, W. P., Stephens, M. L., Mondello, F. J., May, R. J., Lobos, J. H., Carroll, K. M., Brennan, M. J., Bracco, A. A., Fish, K. M., Warner, G. L., Wilson, P. R., Dietrich, D. K., Lin, D. T., Morgan, C. B., and Gately, W. L. 1993. In situ stimulation of aerobic PCB biodegradation in Hudson River sediments. *Science*. 259:503-507.

Harvey, R. W., George, L. H., Smith, R. L., and LeBlanc, D. R. 1989. Transport of microspheres and indigenous bacteria through a sandy aquifer: results of natural and forced gradient tracer experiments. *Environ. Sci. Technol.* 23:51-56.

Hatzinger, P. B. and Alexander, M. 1995. Effect of aging of chemicals in soil on their biodegradability and extractability. *Environ. Sci. Technol.* 29:537-545.

Holmes, D. A., Harrison, B. K., and Dolfing, J. 1993. Estimation of Gibbs free energies of formation for polychlorinated biphenyls. *Environ. Sci. Technol.* 27:725-731.

Howard, P. H. 1989. *Handbook of Environmental Fate and Exposure Data for Organic Chemicals*, Lewis Publishers, Chelsea, MI.

Huesemann, M. H. 1995. Predictive model for estimating the extent of petroleum hydrocarbon biodegradation in contaminated soils. *Environ. Sci. Technol.* 29:7-18.

Hutchins, S. R. 1991a. Biodegradation of monoaromatic hydrocarbons by aquifer microorganisms using oxygen, nitrate, or nitrous oxide as the terminal electron acceptor. *Appl. Environ. Microbiol.* 57:2403-2407.

Hutchins, S. R. 1991b. Optimizing BTEX biodegradation under denitrifying conditions. *Environ. Toxicol. Chem.* 10:1437-1448.

Hutchins, S. R., Downs, W. C., Wilson, J. T., Smith, G. B., Kovacs, D. A., Fine, D. D., Douglas, R. H., and Hendrix, D. J. 1991. Effect of nitrate addition on biorestoration of fuel-contaminated aquifer: field demonstration. *Ground Water*. 29:571-580.

Kapoor, V., Gellhar, L. W., and Wilhelm, F. M. 1997. Bimolecular second order reactions in spacially varying flows: Segregation induced scale dependent transformation rates. *Water Resour. Res.* 33(4): 527-536.

Karickhoff, S. W. 1980. Sorption kinetics of hydrophobic pollutants in natural sediments, in *Contaminants and Sediments*, Vol. 2, Butterworth, Stoneham, MA, 193.

Kellems, B. L., Leeson, A., and Hinchee, R. E. 1994. Review of bioremediation experience in Alaska, in R.E. Hinchee, B.C. Alleman, R.E. Hoeppel, and R.N. Miller (eds.), *Hydrocarbon Bioremediation*. Lewis Publishers, Boca Raton, FL, 438.

King, R. B., Long, G. M., and Sheldon, J. K. 1992. *Practical Environmental Bioremediation*. Lewis Publishers, Boca Raton, FL.

Leahy, J. G. and Colwell, R. R. 1990. Microbial degradation of hydrocarbons in the environment. *Microbiol. Rev.* 54:305-315.

LeBlanc, D. R., Garabedian, S.P., Hess, K.M., Gelhar, L.W., Quadri, R.D., Stollenwerk, K.G., and Wood, W.W. 1991. Large-scale natural gradient tracer test in sand and gravel, Cape Cod, Massachusetts: 1. Experimental design and observed tracer movement. *Water Resour. Res.* 27(5):895-910.

LeChevallier, M. W. and Norton, W. D. 1995. Giardia and Cryptosporidium in raw and finished water. *J. Am. Water Works Assoc.* 87:54-68.

Lee, M. D., Thomas, J. M., Borden, R. C., Bedient, P. B., Ward, C. H., and Wilson, J. T. 1988. Biorestoration of aquifers contaminated with organic compounds. *CRC Crit. Rev. Environ. Control.* 18:29-88.

Lovley, D. R., Baedecker, M. J., Lonergan, D. J., Cozzarelli, I. M., Phillips, E. J. P., and Siegei, D. I. 1989. Oxidation of aromatic contaminants coupled to microbial iron reduction. *Nature*. 339:297-299.

Lovley, D. R. and Lonergan, D. J. 1990. Anaerobic oxidation of toluene, phenol and p-cresol by the dissimilatory iron-reducing organism, GS-15. *Appl. Environ. Microbiol.* 56:1858-1864.

MacDonald, T. R. 1995. Flow from a Recirculation Well for Enhanced In-Situ Bioremediation, PhD dissertation, Department of Civil Engineering, Stanford University, Stanford, CA.

Macquarrie, K. T. B., Sudicky, E. A., and Frind, O. E. 1990. Simulation of biodegradable organic contaminants in groundwater, 1. Numerical formulations in principal directions, *Water Resour. Res.* 26(2):207-222.

McAllister, K. A., Lee, H., and Trevors, J. T. 1996. Microbial degradation of pentachlorophenol. *Biodegradation.* 7:1-40.

McCarty, P. L. and Semprini, L. 1993. Engineering and hydrogeological problems associated with in situ treatment. *Hydrol. Sci.* 38:261-272.

Miralles-Wilhelm, F. and Gelhar, L. W. 1996. Stochastic analysis of sorption macrokinetics in heterogeneous aquifers. *Water Resour. Res.* 32(6), 1541-1549.

Miralles-Wilhelm, F., Kapoor, V., and Gelhar, L. W. 1994. Modeling oxygen limited biodegradation in three dimensionally heterogeneous aquifers. Prepared Under Contract API GW-14-360-8, Department of Environmental Health and Sciences, American Petroleum Institute.

Mohn, W. W. and Tiedje, J. M. 1992. Microbial reductive dehalogenation. *Microbiological Reviews.* 56:482-507.

Molz, F. J. and Widdowson, M. A. 1988. Internal inconsistencies in dispersion-dominated models that incorporate chemical and microbial kinetics. *Water Resour. Res.* (24)4:615-619.

National Research Council. 1993. In Situ Bioremediation. When Does it Work? National Academy Press, Washington, D.C.

Neilson, A. H. 1990. A review — The biodegradation of halogenated organic compounds. *J. Appl. Bacteriol.* 69:445-470.

Nelson, C. H., Hicks, R. J., and Andrews, S. D. 1994. In situ bioremediation: an integrated systems approach, in R.E. Hinchee, B.C. Alleman, R.E. Hoeppel, and R.N. Miller (eds.), *Hydrocarbon Bioremediation*, Lewis Publishers, Boca Raton, FL, 125.

Nelson, M. J. K., Montgomery, S. O., Mahaffey, W. R., and Pritchard, P. H. 1987. Biodegradation. of trichloroethylene and involvement of an aromatic biodegradative pathway. *Appl. Environ. Microbiol.* 53:949-954.

Pardieck, D. L., Bouwer, E. J., and Stone, A. T. 1990. Hydrogen peroxide as a source of oxidant capacity for the biotransformation of benzene, toluene and xylene in biofilms. *J. Environ. Engin.* 374.

Pardieck, D. L., Bouwer, E. J., and Stone, A. T. 1992. Hydrogen peroxide use to increase oxidant capacity for in situ bioremediation of contaminated soils and aquifers: a review. *J. Contam. Hydrol.* 9:221-242.

Reineke, W. and Knackmuss, H. J. 1988. Microbial degradation of haloaromatics. *Annu. Rev. Microbiol.* 42:263-287.

Riser-Roberts, E. 1992. *Bioremediation of Petroleum Contaminated Sites.* C.K. Smoley, Boca Raton, FL.

Rittman, B. E. 1982. The effect of shear stress on biofilm loss rate, *Biotechnology and Bioengineering* 24: 501-506.

Roberts, P. V., Hopkins, G. D., Mackay, D. M., and Semprini, L. 1990. A field evaluation of in-situ biodegradation of chlorinated ethenes: part I, methodology and field site characterization. *Ground Water.* 28:591-604.

Roberts, P. V., Semprini, L., Hopkins, G. D., Grbic-Galic, D., and McCarty, P. L. 1989. In situ aquifer restoration of chlorinated aliphatics by methanotrophic bacteria. Environmental Protection Agency, EPA/600/2-89/033. Washington, DC.

Robin, M. J. L., Sudicky, E. A., Gillham, R. W., and Kachonaski, R. G. 1991. Spatial variability of strontium distribution coefficients and their correlation with hydraulic conductivity in the Canadian Forces Base Borden aquifer. *Water Resour. Res.* 27(10):2619-2632.

Schink, B. 1988. Principles and limits of anaerobic degradation: environmental implications and technological aspects, in A.J.B. Zehnder (ed.), *Biology of Anaerobic Microorganisms.* John Wiley & Sons, New York, 771.

Schwarzenbach, R. P., Gschwend, P. M., and Imboden, D. M. 1993. *Environmental Organic Chemistry.* John Wiley & Sons, New York.

Scow, K. M. 1990. Rate of biodegradation, W.J. Lyman, W.F. Reehl, and D.H. Rosenblatt (eds.), *Handbook of Chemical Property Estimation Methods*, American Chemical Society, Washington, DC..

Semprini, L. and McCarty, P. L. 1991. Comparison between model simulations and field results for in-situ biorestoration of chlorinated aliphatics: Part 1. Biostimulation of methanotrophic bacteria. *Groundwater.* 29(3):365-374.

Semprini, L. and McCarty, P. L. 1992. Comparison between model simulations and Field Results for in-situ biorestoration of chlorinated Aliphatics: Part 2. Cometabolic transformations. *Groundwater.* 30(1):37-44.

Semprini, L., Roberts, P. V., Hopkins, G. D., and McCarty, P. L. 1990. A field evaluation of in-situ biodegradation of chlorinated ethenes: part 2, results of biostimulation and biotransformation experiments. *Ground Water.* 28:715-727.

Shannon, M. J. R. and Unterman, R. 1993. Evaluating bioremediation: distinguishing fact from fiction. *Annu. Rev. Microbiol.* 47:715-738.

Smith, M. R. 1990. The biodegradation of aromatic hydrocarbons by bacteria. *Biodegradation.* 1:191-206.

Streck, T., Poletika, N. N., Jury, W. A., and Farmer, W. J. 1995. Description of simazine transport with rate-limited, two-stage, linear and nonlinear sorption. *Water Resour. Res.* 31(4):811-822.

Stumm, W., Morgan, J. J. 1981. *Aquatic Chemistry.* John Wiley & Sons, New York.

Taylor, S. W. and Jaffe, P. R. 1990. Biofilm growth and the related changes in the physical properties of a porous medium. 1. Experimental investigation, *Water Resour. Res.* 26(9):2153-2159.

Thauer, R. K., Jungermann, K., and Decker, K. 1977. Energy conservation in chemotrophic anaerobic bacteria. *Bacteriol. Rev.* 41:100-180.

Thomas, J. M. and Ward, C. H. 1989. In situ biorestoration of organic contaminants in the subsurface. *Environ. Sci. Technol.* 23:760-766.

Tiedje, J. M. and Alexander, M. 1969. Enzymatic cleavage of the ether bond of 2,4-dichlorophenoxyacetate. *J. Agric. Food Chem.* 17:1080-1084.

Valocchi, A. J. 1985. Validity of the local equilibrium assumption for modeling sorbing solute transport through homogeneous soils. *Water Resour. Res.* 21(6):808-820.

Vandevivere, P. and Baveye, P. 1992a. Saturated hydraulic conductivity reduction caused by aerobic bacteria in sand columns. *Soil Sci. Soc. Amer. J.* 56(1):1-13.

Vandevivere, P. and Baveye, P. 1992b. Effect of bacterial extracellular polymers on the saturated hydraulic conductivity of sand columns. *Applied Environmental Microbiology.* 58(5):1690-1698.

Vogel, T. M., Criddle, C. S., and McCarty, P. L. 1987. Transformations of halogenated aliphatic compounds. *Environ. Sci. Technol.* 21:722-736.

Walton, R. K., Hornberger, G. M., Herman, J. S., and Mills, A. L. 1996. Kinetics of BTX biodegradation and mineralization in batch and column systems. *J. Contaminant Hydrology.* 23:113-132.

Watkinson, R. J. and Morgan, P. 1990. Physiology of aliphatic hydrocarbon degrading microorganisms. *Biodegradation.* 1:79-92.

Wu, S. and Gschwend, P. M. 1986. Sorption kinetics of organic hydrophobic compounds to natural sediments and soils. *Environ. Sci. Technol.* 20, 717-725.

Xun, L., Topp, E., and Orser, C. S. 1992. Confirmation of oxidative dehalogenation of pentachlorophenol by a Flavobacterium pentachlorophenol hydroxylase. *J. Bacteriol.* 174:5745-5747.

Young, L. Y. and Cerniglia, C. E. 1997. *Microbial Transformation and Degradation of Toxic Organic Chemicals.* Wiley-Liss, New York.

Zehnder, A. J. B. 1988. *Biology of Anaerobic Microorganisms.* John Wiley & Sons, New York.

Glossary

Aerobic   Microorganisms which use molecular oxygen as a terminal electron acceptor or an environment which contains molecular oxygen.

Anaerobic   Microorganisms which do not use molecular oxygen as a terminal electron acceptor or an environment which contains no molecular oxygen.

Autotroph   An organism which uses inorganic carbon as a source for biosynthesis of cellular constituents.

Bacteria   In this chapter bacteria includes all prokaryotic organisms. In fact, Achaea and Bacteria are separate phylogenetic domains of prokaryotic organisms.

Bioavailability   Refers to the availability of a substance for uptake and transformation by living organisms.

Bioremediation   The use of naturally occurring biodegradation processes for the remediation of sites contaminated with pollutants.

Cell   An assembly of highly structured macromolecules which is the smallest living unit. Bacteria are single-celled organisms.

Citric Acid Cycle   A nearly universal biochemical cycle which oxidizes organic carbon to CO_2 and NADH.

Cometabolism   A gratuitous biochemical reaction from which the microorganism derives no benefit.

Electron Acceptor   The terminal oxidant in respiratory processes. Oxygen is the electron acceptor of aerobic organisms.

Electron Donor   The compound which is oxidized as a source of energy by organisms.

Eukaryote   Organisms which have a membrane-enclosed compartment containing their DNA (a nucleus).

Fermentation   Substrate transformation without the use of an external electron acceptor, therefore the substrate serves as both electron donor and electron acceptor.

Halogens   Elements in group VII of the Periodic Table. F, Cl, Br, I, At are halogens.

Heterotroph   An organism which uses organic carbon as a source for biosynthesis of cellular constituents.

Lithotroph   An organism which oxidizes inorganic carbon for energy.

Organotroph   An organism which oxidizes organic carbon for energy.

pH   A logarithmic unit of H^+ ion concentration. $pH = -\log(H^+)$

Prokaryote   Single-celled organisms without a membrane-enclosed compartment containing DNA.

Redox Potential   The electrical potential of a given oxidation–reduction reaction.

Reductive Dechlorination   Removal of halogens as halide ions by the reduction of the organic compound with the replacement of the halogen by a hydrogen atom.

Respiration   Substrate transformation in which an external oxidant is used as a terminal electron acceptor.

27
Geosynthetics

Jorge G. Zornberg
GeoSyntec Consultants

Barry R. Christopher
Independent Consultant

27.1 Introduction

Geosynthetics can be defined as planar products manufactured from polymeric material, which are used with soil, rock, or other geotechnical engineering-related material as an integral part of a man-made project, structure, or system (ASTM, 1995). Geosynthetics are widely used in many geotechnical and environmental applications related to groundwater quality and control. This is the case, for example, of base and cover liner systems for modern landfills, which are designed making extensive use of geosynthetics. The main purpose of geosynthetic liner systems is to minimize potential groundwater contamination. Moreover, the use of geosynthetics is rapidly increasing in applications related directly to groundwater control. This is the case of high-density polyethylene (HDPE) vertical barrier systems, which are used instead of traditional soil-bentonite cutoff walls in projects involving groundwater remediation and control.

The geosynthetics market is strong and rapidly increasing because of the continued use of geosynthetics in well-established applications and, particularly, because of the increasing number of new applications which make use of these products. The strength of the geosynthetics market can be appreciated by evaluating the growth in the estimated amount of geosynthetics in North America over the years. While the total amount of geosynthetics produced in North America was slightly more than 83 million m^2 in 1980, the production of geosynthetics exceeded 500 million m^2 in 1995. Table 27.1 shows the estimated North American shipments of geosynthetics for 1995 and predictions until 2001. Collectively, 150 manufacturers of geosynthetic products shipped 525 million m^2 of material in 1995 and grew at an annual rate of 7% (Industrial Fabrics Association International, 1996). In the 2001 forecast period, manufacturers are expected to ship approximately 734 million m^2 of materials.

TABLE 27.1 North American Shipments of Geosynthetic Materials, 1995-2001 (In Million m^2)

	1995	**1996**	**1998**	**2001**
Geotextiles	346.2	356.2	419.7	477.4
Geomembranes	62.4	64.4	74.6	86.8
Geogrids	22.9	24.3	29.1	36.8
Geosynthetic clay liners	5.0	5.4	6.1	8.2
Erosion-control products	72.7	77.8	82.8	93.6
Specialty geosynthetics	16.7	20.1	25.9	31.8

From Industrial Fabrics Association International. 1996. *North American Market for Geosynthetics — 1996.*

Geosynthetics applications are very diverse. In order to fulfill different functions in the design of geotechnical-, environmental-, and groundwater-related systems, the geosynthetic industry has developed a number of products. In addition to the already mentioned examples regarding use of geomembranes in landfill liner systems and the use of HDPE vertical panels in groundwater control projects, other examples of geosynthetics applications include the use of geotextiles as filtration elements in dams and waste containment systems, the use of geocomposites as erosion control elements in channels and slopes, and the use of geogrids as reinforcement elements in soil embankments, to mention a few more.

Geosynthetics have numerous material properties. Many of the reported properties are important in the manufacture and quality control of geosynthetics; however, many others are also important in design. The material properties related to the manufacture and quality control of geosynthetics are generally referred to as index properties and those related to the design as design or performance properties. Considering their different properties, the several geosynthetic products can perform different functions and, consequently, they should be designed to satisfy minimum criteria to adequately perform these functions. The different functions performed by geosynthetics are discussed in Section 27.2. The geosynthetic functions are as follows:

- Separation
- Reinforcement
- Filtration
- Drainage
- Infiltration barrier
- Protection (or stress relief)

Geosynthetics are manufactured in a factory-controlled environment. They are packaged in sheets, placed in a roll or carton, and finally transported to the site. At the project site the geosynthetic sheets are unrolled on the prepared subgrade surface, overlapped to each other to form a continuous geosynthetic blanket, and often physically joined to each other. The individual types of products within the geosynthetics family are discussed in Section 27.3. The geosynthetic types are as follows:

- Geotextiles
- Geomembranes
- Geogrids
- Geosynthetic clay liners (GCLs)
- Geocomposite sheet drains
- Geocomposite strip (wick) drains
- Geocells
- Erosion control products
- HDPE vertical barrier systems

Note in the list above that different types of geocomposite drains are treated separately, and that HDPE vertical barriers are not lumped together with the rest of the geomembrane products. These geosynthetics are described separately in this chapter because of their particular relevance in groundwater-related applications.

Geotechnical, environmental, and groundwater systems frequently incorporate several types of geosynthetics, which are designed to perform more than one function in the system. The bottom and cover liners of waste containment facilities are good examples of applications that make use of geosynthetics for multiple purposes. In these facilities, the different geosynthetic products are combined to fulfill the functions of infiltration barrier, filtration, separation, drainage, protection, and reinforcement. The multiple use of geosynthetics in the design of modern landfills is described in Section 27.4. Finally, a case history illustrating the use of HDPE panels as a vertical barrier in a groundwater control project is presented in Section 27.5. A glossary of relevant terms and a list of sources are included for further information.

27.2 Geosynthetic Functions

27.2.1 Design by Function of Geosynthetics

As with other engineering materials, there are several design approaches that could be used during the selection process of geosynthetic products. The most common geosynthetic design methods are by experience, by specification, or by function (Koerner, 1994).

Design-by-experience is generally based on the use of the manufacturer's literature and of the designer's experience and familiarity with geosynthetic products. Design-by-specification is practiced, for example, by government agencies (e.g., state departments of transportation) for routine applications. It often consists in selecting geosynthetic products for common application areas, taking as a basis minimum- or maximum-specified property values.

Design-by-function can be used in addition to the aforementioned methods and is required for those applications not covered by specifications or of such a nature that large property or personal damage would result in the event of a failure. A generic design process that applies to the different geosynthetic functions is summarized as follows (Koerner, 1994):

1. Evaluate the criticality and severity of the application
2. Determine the function(s) of the geosynthetic
3. Calculate, estimate, or otherwise determine the required property value for the function(s)
4. Test or otherwise obtain the allowable property of the candidate geosynthetic material
5. Calculate the factor of safety (*FS*) ratio as follows:

$$FS = \frac{\text{allowable (test) value}}{\text{required (design) value}}$$

6. Determine if the resulting factor of safety is significantly high for the site-specific situation under consideration
7. Prepare specifications and construction documents
8. Observe construction and post-construction performance.

If the factor of safety is sufficiently high for the specific application, the candidate geosynthetic is acceptable. The same process can be repeated for a number of available geosynthetics, and the final selection among acceptable products is based on availability and cost.

The design-by-function approach is the general approach to be followed in the majority of the projects. As mentioned, the primary function of geosynthetics is either separation, reinforcement, filtration, drainage, infiltration barrier, or protection. However, a certain geosynthetic product can perform different

TABLE 27.2 Function of Different Geosynthetic Products

	Geo-textile	Geo-membrane	Geo-grid	GCL	Geo-composite Sheet Drain	Geo-composite Strip (Wick) Drain	Geocell	Erosion Control Product	HDPE Vertical Barrier
Separation	X						X	X	
Reinforcement	X		X				X		
Filtration	X								
Drainage	X				X	X			
Infiltration barrier	X[1]	X		X					X
Protection	X			X					

[1] Asphalt-saturated geotextiles

functions and, similarly, the same function can often be performed by different types of geosynthetics. The specific function(s) of the different geosynthetic(s) are presented in Table 27.2. Each of these functions is described in Sections 27.2.2 to 27.2.7.

27.2.2 Separation Function

Separation is the introduction of a flexible, porous geosynthetic product between dissimilar materials so that the integrity and functioning of both materials can remain intact or be improved. For example, a major cause of failure of roadways constructed over soft foundations is contamination of the aggregate base courses with the underlying soft subgrade soils (Figure 27.1A). Contamination occurs due to: (1) penetration of the aggregate into the weak subgrade due to localized bearing capacity failure under stresses induced by wheel loads, and (2) inclusion of fine-grained soils into the aggregate because of pumping or subgrade weakening due to excess pore water pressures. Subgrade contamination results in inadequate structural support, which often leads to premature failure of the system. A geotextile can be placed between the aggregate and the subgrade to act as a separator and prevent the subgrade and aggregate base course from mixing (Figure 27.1B).

Among the different geosynthetics, geotextiles have been the products generally used in the function of separation. Examples of separation applications are the use of geotextiles between subgrade and stone base in roads and airfields, and between geomembranes and drainage layers in landfills. In addition to these applications, in which separation is the primary function of the geotextile, it could be said that most geosynthetics generally include separation as a secondary function.

Geosynthetics used as erosion control systems can also be considered as performing a separation function. In this case, the geosynthetic separates the ground surface from the prevailing atmospheric
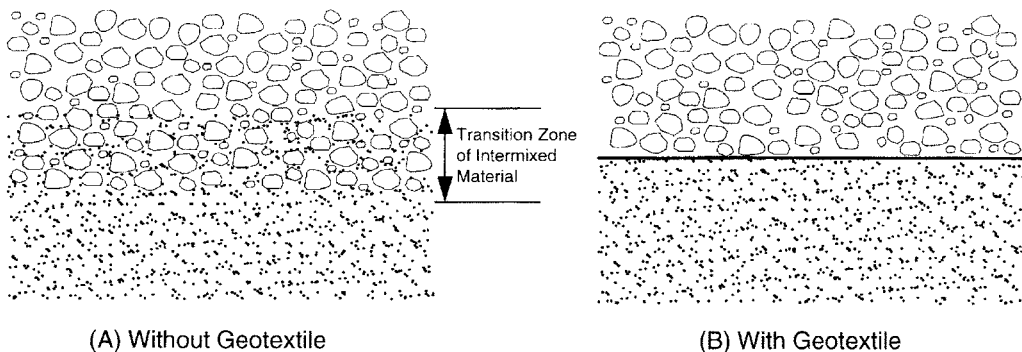
FIGURE 27.1 Separation function of a geotextile placed between road aggregate and soft subgrade.

conditions (i.e., wind, rain, snow, etc.). Specialty geocomposites have been developed for the specific purpose of erosion control. The general goal of these products is to protect soil slopes from both sheet and gully erosion, either permanently or until vegetation is established. General references on the design of geosynthetics for separation applications can be found in Christopher and Holtz (1985) and in Koerner (1994).

27.2.3 Reinforcement Function

Geosynthetic inclusions within a soil mass can provide a reinforcement function by developing tensile forces which contribute to the stability of the geosynthetic–soil composite (a reinforced soil structure). Design and construction of stable slopes and retaining structures within space constraints are aspects of major economical significance in geotechnical engineering projects. For example, when geometry requirements dictate changes of elevation in a highway project, the engineer faces a variety of distinct alternatives for designing the required earth structures. Traditional solutions have been either a concrete retaining wall or a conventional, relatively flat, unreinforced slope (Figure 27.2). Although simple to design, concrete wall alternatives have generally led to elevated construction and material costs. On the other hand, the construction of unreinforced embankments with flat slope angles dictated by stability considerations is an alternative often precluded in projects where design is controlled by space constraints.
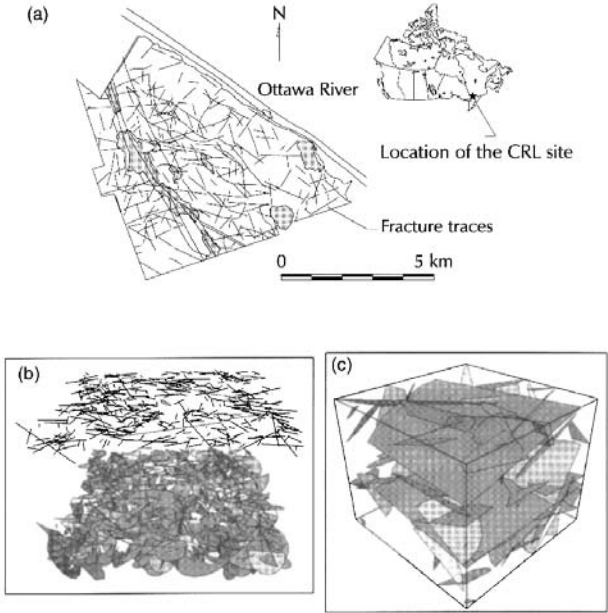
FIGURE 27.2 Reinforcement function of geosynthetics used to optimize the design of earth retaining structures.

Geosynthetics are particularly suitable for soil reinforcement. Geosynthetic products typically used as reinforcement elements are nonwoven geotextiles, woven geotextiles, geogrids, and geocells. Reinforced soil vertical walls generally provide vertical grade separations at a lower cost than traditional concrete walls. Reinforced wall systems involve the use of shotcrete facing protection or of facing elements such as precast or cast-in-place concrete panels. Alternatively, steepened reinforced slopes may eliminate the use of facing elements, thus saving material costs and construction time in relation to vertical reinforced walls. As indicated in Figure 27.2, a reinforced soil system generally provides an optimized alternative for the design of earth-retaining structures.

The effect of geosynthetic reinforcements on the stability of sand slopes is illustrated in Figure 27.3, which shows a reduced scale geotextile-reinforced slope model built using dry sand as backfill material.
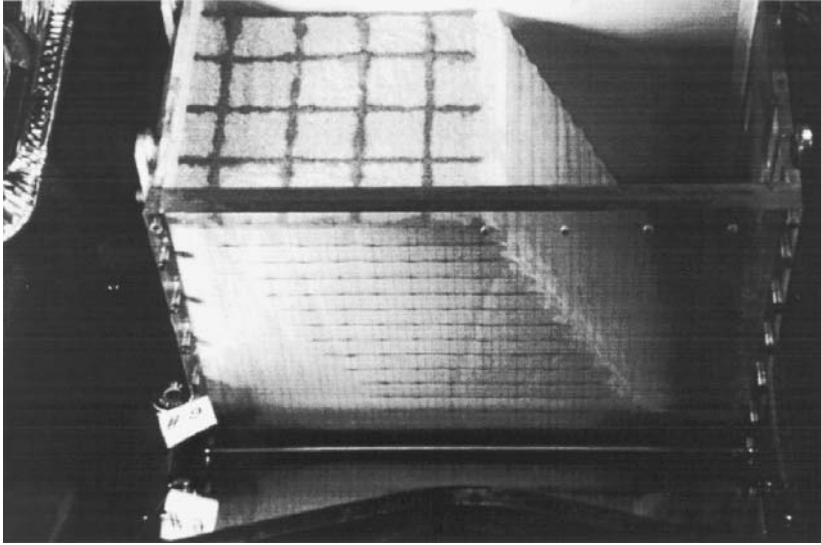
FIGURE 27.3 Model of a sand slope reinforced with geosynthetics.
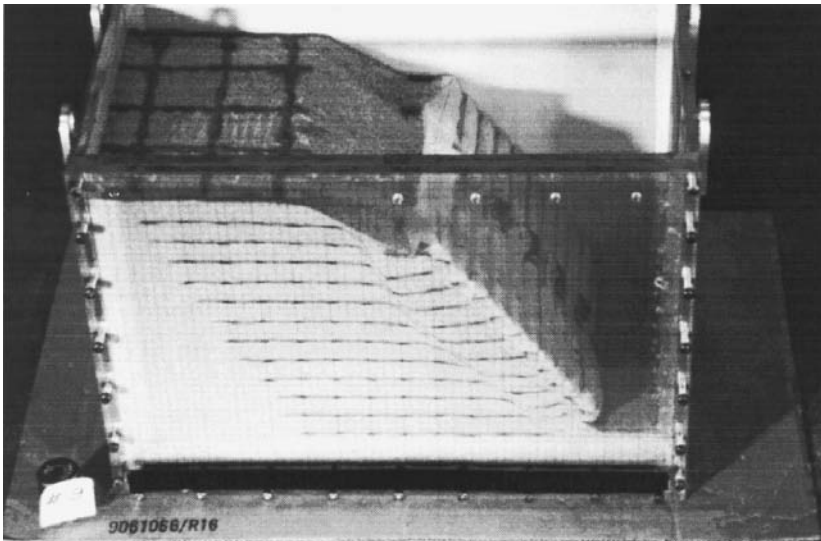
FIGURE 27.4 Reinforced slope model brought to failure by increasing the unit weight of the backfill.

The maximum slope inclination of an unreinforced sand under its own weight is the angle of repose of the sand, which is well below the inclination of the slope face of the model. Horizontal geotextile reinforcements placed within the backfill provided stability to the steep sand slope. In fact, not only did the reinforced slope model not fail under its own weight, but its failure only occurred after the unit weight of the backfill was increased 67 times by placing the model in a geotechnical centrifuge (Zornberg et al., 1997). Figure 27.4 shows the reinforced slope model after centrifuge testing.

The use of inclusions to improve the mechanical properties of soils dates to ancient times. However, it is only within the last quarter of century or so (Vidal, 1969) that analytical and experimental studies have led to contemporary soil reinforcement techniques. Soil reinforcement is now a highly attractive alternative for embankment and retaining wall projects because of the economic benefits it offers in relation to conventional retaining structures. Moreover, its acceptance has also been triggered by a number

of technical factors, that include aesthetics, reliability, simple construction techniques, good seismic performance, and the ability to tolerate large deformations without structural distress. The design of reinforced soil slopes is based on the use of limit equilibrium methods to evaluate both external (global) and internal stability of the structure. The required tensile strength of the reinforcements is selected during design so that the margins of safety, considering an internal failure similar to the one shown in Figure 27.4, are adequate. Guidance in soil reinforcement design procedures is provided by Mitchell and Villet (1987), Christopher et al. (1989), and Elias and Christopher (1997).

27.2.4 Filtration Function

The filtration function involves movement of liquid through the geosynthetic and, at the same time, retention of soil on its upstream side. As indicated in Table 27.2, geotextiles are the geosynthetic product generally used in filtration. Both adequate hydraulic conductivity (provided by a geotextile with a relatively open structure) and adequate soil retention (provided by a geotextile with a relatively tight structure) should be offered by the selected product. In addition, considerations should be made regarding the long-term soil-to-geotextile flow compatibility such that the flow through the geotextile will not reduce excessively by clogging during the lifetime of the system. The geosynthetic-to-soil system should then achieve an equilibrium that allows for adequate liquid flow with limited soil loss across the plane of the geotextile over a service lifetime compatible with the application under consideration. Filtration concepts are well established in the design of soil filters, and similar concepts can be used in the design of geotextile filters.

The flow of liquid is perpendicular to the plane of the geosynthetic and, consequently, filtration refers to the cross-plane hydraulic conductivity. Some of the geosynthetics used for this purpose are relatively thick and compressible. For this reason, geosynthetics are generally characterized by their permittivity, which is defined as:

$$\psi = k_n/t$$

where ψ is the permittivity, k_n is the cross-plane hydraulic conductivity, and t is the geosynthetic thickness at a specified normal pressure.

Testing procedures for geotextile permittivity follow similar guidelines used for testing soil hydraulic conductivity. Some designers prefer to work directly with hydraulic conductivity and require the geotextile hydraulic conductivity to be some multiple of the adjacent soil's hydraulic conductivity (Christopher and Fischer, 1992).

As the flow of liquid through the geotextile increases, the geotextile voids should be larger. However, large geotextile voids can lead to an unacceptable situation called soil piping, in which the soil particles are continuously carried through the geotextile, leaving large soil voids behind. The liquid velocity then increases, which accelerates the process and may lead to the collapse of the soil structure. This process can be prevented by selecting a geotextile with voids small enough to retain the soil on the upstream side of the fabric. It is the coarser soil fraction that must be initially retained. The coarser-sized particles eventually filter the finer-sized particles and build up a stable upstream soil structure (Figure 27.5).

Several approaches have been developed for soil retention design using geotextiles; most of them compare the soil particle size characteristics to the 95% opening size of the geotextile (defined as O_{95} of the geotextile). The test method used in the United States to determine the geotextile opening size is called the apparent opening size (AOS) test.

Some of the soil particles will embed themselves on or within the geotextile structure and will cause a reduction in the hydraulic conductivity or permittivity of the geotextile. Although some partial clogging should be expected, the designer should ensure that the geotextile will not excessively clog. That is, the flow of liquid will not be decreased to a point at which the system will not adequately perform its function. Thus, the geotextile voids should be large enough to allow the finer soil particles to pass. Guidelines are available for clogging evaluation of noncritical, nonsevere cases, but laboratory testing is necessary in
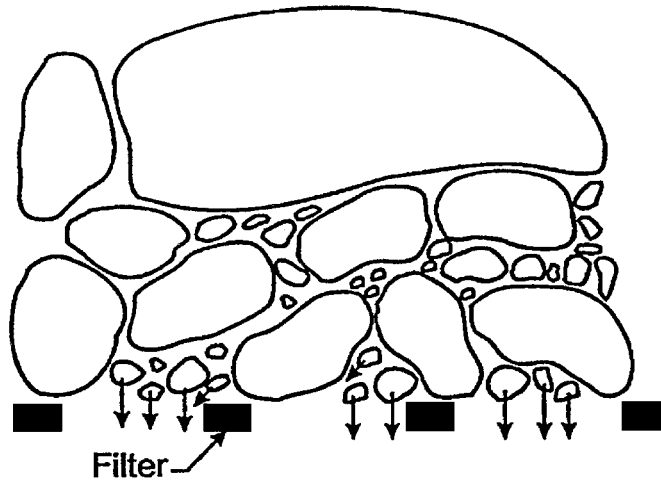
FIGURE 27.5 Geotextile providing adequate filtration through selection of adequate opening size.

important applications. Either the gradient ratio test (Haliburton and Wood, 1982), the long-term flow test (Halse et al., 1987), or the hydraulic conductivity ratio test (Williams and Abouzakhm, 1989) should be performed. An evaluation of the filtration function of geotextiles is provided by Christopher and Fischer (1992), Giroud (1996), Bhatia and Smith (1996a,b), Bhatia et al. (1996), and Holtz et al. (1997).

27.2.5 Drainage Function

Geosynthetics provide a drainage function by transmitting liquid within the plane of their structure. As shown in Table 27.2, the geosynthetics generally used for drainage purposes are geotextiles and geocomposites. The drainage function of geosynthetics allows for adequate liquid flow with limited soil loss within the plane of the geotextile over a service lifetime compatible with the application under consideration.

Thick, needle-punched nonwoven geotextiles have considerable void space in their structure and can convey large amounts of liquid. Geocomposite drains can transmit one to two orders of magnitude more liquid than geotextiles. Proper design should dictate what type of geosynthetic drainage material is necessary.

Except for the consideration of flow direction, the soil retention and the long-term compatibility considerations regarding the drainage function of geosynthetics are the same as those discussed in Section 27.2.4 regarding the filtration function of geosynthetics. Since the geosynthetic thickness decreases with increasing normal stress, the in-plane drainage of a geosynthetic is generally quantified by its transmissivity, which is defined as:

$$\theta = k_p \cdot t$$

where θ is the transmissivity, k_p is the in-plane hydraulic conductivity, and t is the geosynthetic thickness at a specified normal pressure.

The geotextile, either when used as a drain itself or when placed onto a core to form a geocomposite must fulfill the filtration function. The compatibility of the soil with the geotextile filter must be ensured over the lifetime of the system being built. General references on design methods for the use of geosynthetics for drainage applications can be found in Holtz et al. (1997) and in Koerner (1994).

27.2.6 Infiltration Barrier Function

The infiltration barrier function can be performed by geosynthetic products that have hydraulic conductivity low enough to provide containment to liquid or vapor. As shown in Table 27.2, the infiltration barrier function may be provided by several types of geosynthetics, namely, geomembranes and geosynthetic clay liners (GCLs). Other geosynthetic products also used as infiltration barriers include membrane-encapsulated soil layers (MESLs) used with paved or unpaved road construction, asphalt-saturated geotextiles used in the prevention of bituminous pavement crack reflection problems, and geofoam used for insulation against moisture and/or temperature.

Geosynthetic barriers are commonly used as liners for surface impoundments storing hazardous and nonhazardous liquids, as covers above the liquid surface of storage reservoirs, and as liners for canals used to convey water or chemicals. Geosynthetic barriers are also used as secondary containment for underground storage tanks and in applications related to dams and tunnels. Of particular relevance for groundwater applications is the use of geosynthetic barriers for seepage control (HDPE vertical barrier systems). A common application of geosynthetics as infiltration barriers is for base and cover liner systems of landfills. In landfill applications, infiltration barriers are typically used instead of or in addition to low-hydraulic conductivity soils. Base liners are placed below the waste to prevent liquids from the landfill (leachate) from contaminating the underlying ground and the groundwater. Geosynthetic cover liner systems are placed above the final waste configuration to keep precipitation water from entering the waste and generate leachate. If a building or other structure is constructed on a landfill, a geosynthetic barrier may be placed under the building foundation to provide a barrier for vapors such as landfill gas. The use of geosynthetics in infiltration barriers is further described in Koerner (1994).

27.2.7 Protection Function

Geosynthetics (mainly geotextiles) can be used to protect other geosynthetics (mainly geomembranes) against damage. A common example is the use of geotextiles to provide protection against puncture of geomembranes in waste and liquid containment systems. Adequate mechanical protection must be provided to resist both short-term equipment loads and long-term loads imparted by the waste. Experience has shown that geotextiles can play an important role in the successful installation and longer-term performance of geomembranes by acting as a cushion to prevent puncture damage of the geomembrane. In the case of landfill base liners, geotextiles can be placed (1) below the geomembrane to resist puncture and wear due to abrasion caused by sharp-edged rocks in the subgrade, and (2) above the geomembrane to resist puncture caused either by the drainage aggregate or direct contact with waste materials. Likewise, in the case of landfill cover liners, geotextiles can be placed below the geomembrane to reduce risk of damage by sharp objects in the landfill and above the geomembrane to prevent damage during placement of drainage aggregate or cover soil. Key characteristics for the geotextile cushions are polymer type, mass density, method of manufacture, and construction survivability. The selection process of a geotextile that fulfills a protective function of a geomembrane involves the following three steps: (1) selection of polymer type and method of manufacture; (2) evaluation of the geotextile's capacity to provide puncture protection for the geomembrane; and (3) evaluation of construction survivability. Detailed procedures and methods for conducting these evaluations are described by Holtz et al. (1997), Koerner et al. (1996), Narejo et al. (1996), and Wilson-Fahmy et al. (1996).

27.3 Geosynthetic Types

27.3.1 Geotextiles

Among the different geosynthetic products, geotextiles are the ones that present the widest range of properties. They can be used to fulfill all the different functions listed in Table 27.2 for many different geotechnical, environmental, and groundwater applications. For example, Figure 27.6 shows the con-

FIGURE 27.6 Placement of a high-strength nonwoven geotextile to perform a dual function of reinforcement and in-plane drainage in a reinforced slope.

struction of a reinforced slope in which geotextiles were selected as multipurpose inclusions within the fill, because they can provide not only the required tensile strength (reinforcement function), but also the required transmissivity (drainage function) needed for that particular project (Zornberg et al., 1996).

Geotextiles are manufactured from polymer fibers or filaments which are later formed to develop the final product. Approximately 75% of the geotextiles used today are based on polypropylene resin. An additional 20% are polyester, and the remaining 5% is a range of polymers including polyethylene, nylon, and other resins used for specialty purposes. As with all geosynthetics, however, the base resin has various additives, such as for ultraviolet light protection.

The most common types of fibers used in the manufacture of geotextiles are monofilament, staple, and slit-film. If fibers are twisted or spun together, they are known as a yarn. Monofilament fibers are created by extruding the molten polymer through an apparatus containing small-diameter holes. The extruded polymer strings are then cooled and stretched to give the fiber increased strength. Staple fibers are also manufactured by extruding the molten polymer; however, the extruded strings are cut into 25- to 100-mm portions. The staple fibers may then be spun into longer fibers known as staple yarns. Slit-film fibers are manufactured by either extruding or blowing a film of a continuous sheet of polymer and cutting it into fibers by knives or lanced air jets. Slit-film fibers have a flat, rectangular cross-section instead of the circular cross-section shown by the monofilament and staple fibers.

The fibers or yarns are formed into geotextiles using either woven or nonwoven methods. Figure 27.7 shows a number of typical woven and nonwoven geotextiles. Woven geotextiles are manufactured using traditional weaving methods and a variety of weave types. Nonwoven geotextiles are manufactured by placing and orienting the fabrics on a conveyor belt and subsequently bonding them by needle punching or melt bonding. The needle-punching process consists of pushing numerous barbed needles through the fiber web. The fibers are thus mechanically interlocked into a stable configuration. As the name implies, the heat (or melt) bonding process consists of melting and pressurizing the fibers together.

Common terminology associated with geotextiles includes machine direction, cross machine direction, and selvage. Machine direction refers to the direction in the plane of fabric in line with the direction of manufacture. Conversely, cross machine direction refers to the direction in the plane of fabric perpendicular to the direction of manufacture. The selvage is the finished area on the sides of the geotextile
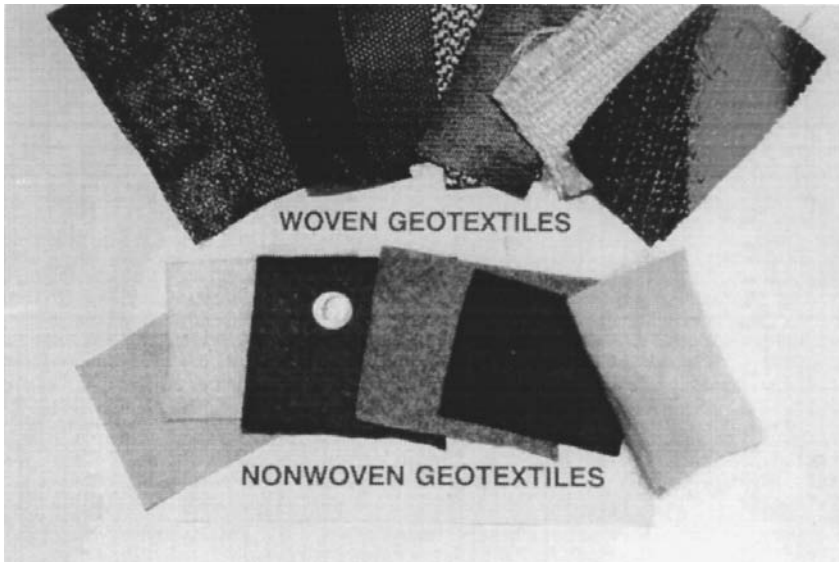
FIGURE 27.7 Typical woven and nonwoven geotextiles.

width that prevents the yarns from unraveling. Adjacent rolls of geotextiles are seamed in the field by either overlapping or sewing. Sewing is generally the case for geotextiles used as filters in landfill applications but may be waived for geotextiles used in separation. Heat bonding may also be used for joining geotextiles in filtration and separation applications.

Numerous tests have been developed to evaluate the properties of geotextiles. In developing geotextile specifications, it is important that the designer understands the material tests and that he or she specifies material properties important for the geotextiles' intended use. Table 27.3 describes the tests commonly performed in geotextile products (ASTM, 1995). Several of the reported material properties and test methods were borrowed from the textile industry. Consequently, several geotextile properties reported by manufacturers are index or quality control tests and are not intended for engineering design.

27.3.2 Geomembranes

Geomembranes are flexible, polymeric sheets that have very low hydraulic conductivity (typically less than 10^{-11} cm/sec) and, consequently, are used as liquid or vapor barriers. The most common types of geomembranes are high-density polyethylene (HDPE), very flexible polyethylene (VFPE), polyvinyl chloride (PVC), and reinforced chlorosulfonated polyethylene (CSPE). Figure 27.8 shows a number of geomembranes currently available in the geosynthetics market.

Polyethylene is the type of geomembrane most commonly used in landfill applications for base and cover liner systems. This is primarily because of its high chemical resistance and durability. Specifically, high-density polyethylene (HDPE) is typically used in base liner systems. This material is somewhat rigid but generally has good physical properties and can withstand the large stresses often imposed on the geomembrane during construction.

VFPE and PVC are the most commonly used geomembrane materials besides HDPE. The term VFPE encompasses various polyethylene grades such as very-low-density polyethylene (VLDPE) and certain types of linear low-density polyethylene (LLDPE). The linear structure and lack of long-chain branching in both LLDPE and VLDPE arise from their similar polymerization mechanisms. Due to the large settlements that may occur, cover liner systems commonly require a flexible geomembrane. VFPE is often used in this application since it provides chemical resistance similar to that of HDPE but is more flexible and can more readily conform to underlying refuse settlements without puncturing.

TABLE 27.3 Standard Tests for Geotextiles

Property	Test Standard	Test Name
Thickness	ASTM D 5199	Standard Test Method for Measuring Nominal Thickness of Geotextiles and Geomembranes
Mass per unit area	ASTM D 5261	Standard Test Method for Measuring Mass per Unit area of Geotextiles
Grab rupture	ASTM D 4632	Standard Test Method for Breaking Load and Elongation of Geotextiles (Grab Method)
Uniaxial tensile strength	ASTM D 4595	Standard Test Method for Tensile Properties by the Wide-Width Strip Method
Multiaxial tensile or burst tests	ASTM D 3786	Standard Test Method for Hydraulic Bursting Strength of Knitted Goods and Nonwoven Fabrics — Diaphragm Bursting Strength Tester Method
Puncture resistance	ASTM D 4833	Standard Test Method for Index Puncture Resistance of Geotextiles, Geomembranes, and Related Products.
Trapezoid tear strength	ASTM D 4533	Standard Test Method for Trapezoid Tearing Strength of Geotextiles
Apparent opening size	ASTM D 4751	Standard Test Method for Determining Apparent Opening Size of a Geotextile
Permittivity	ASTM D 4491	Standard Test Methods for Water Permeability of Geotextiles by Permittivity
Gradient ratio	ASTM D 5101	Standard Test Method for Measuring the Soil-Geotextile System Clogging Potential by the Gradient Ratio
Transmissivity	ASTM D 4716	Standard Test Method for Constant Head Hydraulic Transmissivity (In-Plane Flow) of Geotextiles and Geotextile Related Products
Ultraviolet resistance	ASTM D 4355	Standard Test Method for Deterioration of Geotextiles from Exposure to Ultraviolet Light and Water (Xenon-Arc Type Apparatus)
Seam strength	ASTM D 1683	Failure in Sewn Seams of Woven Fabrics
Seam strength	ASTM D 4884	Standard Test Method for Seam Strength of Sewn Geotextiles
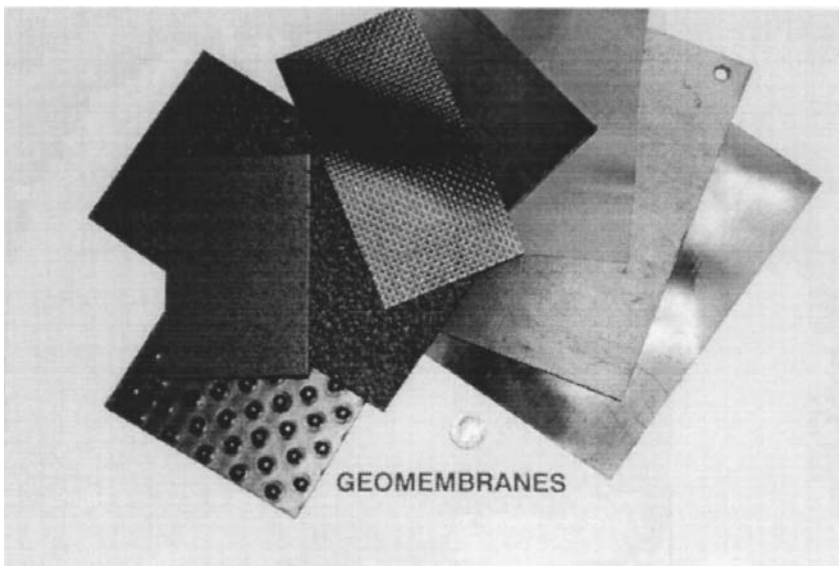
FIGURE 27.8 Typical geomembranes.

PVC geomembranes are used in liners for many waste containment applications, such as contaminated soils containment and liquid storage ponds. In the United States, PVC geomembranes have generally been recommended for relatively short-term applications, approximately 1 to 5 years, due to long-term durability concerns. The merits of PVC geomembranes are that they are generally less expensive than

polyethylene geomembrane and can be factory manufactured in relatively large panels. The large panel sizes allow easier installation since there are fewer field fabricated seams.

In landfill applications, geomembranes are typically used as a base or a cover liner in place of or in addition to low-hydraulic conductivity soils. The key performance factors related to the selection of geomembrane polymer types for landfill applications are summarized in Table 27.4. Geomembrane thickness ranges from 0.75 to 2.5 mm (30 to 100 mils). Table 27.5 summarizes the key performance factors related to the selection of the thickness of HDPE geomembranes for landfill applications. Geomembranes are placed after subgrade preparation, and placement is followed by seaming, inspection, and backfilling. A properly designed geomembrane has the potential of hundreds of years of service lifetime, but installation must follow high-quality management principles. In the early uses of geomembranes for waste containment applications, the main concerns were related to the chemical compatibility between geomembranes and waste, and to the service life of geomembranes. Now, construction quality issues are viewed as the principal limitations to the performance of geomembranes.

TABLE 27.4 Criteria for Selection of HDPE, PVC, or CSPE Geomembranes

Criteria	Considerations for Selection
Liquid barrier	All three polymers have acceptable characteristics as liquid barriers, although HDPE geomembranes have the best. All three have extremely low hydraulic conductivity and are impermeable for practical purposes.
Mechanical properties	Although the mechanical properties vary somewhat with geomembrane thickness, HDPE is relatively stiff and has relatively small yield strain. PVC, in contrast, is relatively extensible and does not exhibit yield. The tensile properties of CSPE often fall between those of HDPE and PVC but are difficult to generalize because CSPE is often made with embedded reinforcing fabrics which affect tensile response.
Construction survivability	All three polymers have acceptable ability to maintain integrity when subjected to concentrated stresses. However, the best performance is obtained with more extensible geomembranes. Therefore, based on the relative extensibility, PVC offers the most favorable performance.
Installation	Key considerations include ease of placement and seaming. PVC and CSPE are easier to place than HDPE because their greater flexibility makes them conform more easily to the foundation and makes them less prone to thermal expansion wrinkles. Acceptable placement and wrinkle control, however, can be achieved with all three polymers if appropriate installation procedures are used. All three polymers are easily seamed, with HDPE usually achieving the highest seam strength and quality.
Chemical resistance	HDPE has the highest degree of compatibility with a wide variety of chemicals encountered in wastes. CSPE has good resistance to many chemicals but is attacked by some which are relatively common, namely chlorinated solvents and hydrocarbons. PVC typically is the least chemically resistant of the three polymers.
Long-term durability	HDPE offers the best performance. HDPE is a highly inert and durable material that is not susceptible to chemical degradation under conditions generally encountered in landfills. In addition, HDPE is not susceptible to physical degradation (extraction). The durability of PVC geomembranes is significantly less favorable than that of HDPE. This is because PVC geomembranes are composed of approximately two-thirds PVC resin and one-third plasticizers. Over time, physical degradation (extraction) may cause plasticizer loss which results in reduced geomembrane flexibility. The durability of CSPE geomembranes is typically between that of HDPE and PVC.

For continuity of the impermeable barrier, geomembranes should be seamed in the field. The fundamental mechanism of seaming polymeric geomembrane sheets together is to temporarily reorganize (melt) the polymer structure of the two surfaces to be joined in a controlled manner. This reorganization can be done either through thermal or chemical processes. These processes may involve the addition of extra polymer in the bonded area. There are four general categories of seaming methods: extrusion welding, thermal fusion or melt bonding, chemical fusion, and adhesive seaming. Extrusion welding and thermal fusion are the methods most commonly used, and are described next.

TABLE 27.5 Criteria for Selection of HDPE Geomembrane Thickness

Criteria	Considerations for Selection of Thickness
Abrasion resistance	The abrasion resistance of HDPE geomembranes increases with geomembrane thickness. Experience indicates that geomembranes with thickness less than [1 mm (40 mils)] may not have acceptable abrasion resistance.
Response to differential settlements	The thicker HDPE geomembrane have higher stiffness. This issue is more significant for geomembrane cover systems than for geomembrane liner systems because the cover system must be flexible enough to accommodate differential settlements. From this viewpoint, a thickness of not more than 2 mm (80 mils) is desirable.
Effective welding	The thinner the HDPE geomembrane, the more difficult is the welding of adjacent panels. For most effective welding, a thickness of at least 1 mm (40 mils) is desirable, and 1.5 mm to 2 mm (60 to 80 mils) is preferred.

Extrusion welding is presently used exclusively on geomembranes made from polyethylene. A ribbon of molten polymer is extruded over the edge of, or in between, the two surfaces to be joined. The molten extrudate causes the surface of the sheets to become hot and melt, after which the entire mass cools and bonds together. The technique is called extrusion fillet seaming when the extrudate is placed over the leading edge of the seam, and is called extrusion flat seaming when the extrudate is placed between the two sheets to be joined. Fillet extrusion seaming is essentially the only practical method for seaming polyethylene geomembrane patches, for seaming in poorly accessible areas such as sump bottoms and around pipes, and for seaming of extremely short seam lengths.

In thermal fusion or melt bonding (the most common seaming method), portions of the opposing surfaces are truly melted. Temperature, pressure, and seaming rate play important roles since excessive melting weakens the geomembrane and inadequate melting results in low seam strength. The hot wedge, or hot shoe, method consists of an electrically heated resistance element in the shape of a wedge that travels between the two sheets to be seamed. A standard hot wedge creates a single uniform width seam, while a dual hot wedge (or "split" wedge) forms two parallel seams with a uniform unbonded space between them. This space can then be conveniently used to evaluate seam quality and continuity by pressurizing the unbonded space with air and monitoring any drop in pressure that may signify a leak in the seam (Figure 27.9).

The material properties of geomembranes are divided into the properties of the raw polymer or resin used in manufacture of the geomembrane sheet and the manufactured geomembrane properties. Table 27.6 lists the tests commonly performed for evaluation of the raw polymer properties. Table 27.7 summarizes the tests commonly performed to evaluate the manufactured geomembrane sheet properties (ASTM, 1995). As with the geotextiles, many of these tests provide index or quality control properties.

27.3.3 Geogrids

Geogrids constitute a category of geosynthetics designed preliminarily to fulfill a reinforcement function. Geogrids have a uniformly distributed array of apertures between their longitudinal and transverse elements. The apertures allow direct contact between soil particles on either side of the installed sheet, thereby increasing the interaction between the geogrid and the backfill soil.

Geogrids are composed of polypropylene, polyethylene, polyester, or coated polyester. They are formed by several different methods. The polyester and coated polyester geogrids are typically woven or knitted. Coating is generally performed using PVC or acrylics to protect the filaments from construction damage. The polypropylene geogrids are either extruded or punched sheet drawn, and polyethylene geogrids are exclusively punched sheet drawn. Figure 27.10 shows a number of typical geogrid products.

Although geogrids are used primarily for reinforcement, some products are used for asphalt overlay and some are combined with other geosynthetics to be used in waterproofing or in separation and stabilization applications. In waste containment systems, geogrids may be used to support a lining system over a weak subgrade or to support final landfill cover soils on steep refuse slopes. A relatively new

FIGURE 27.9 Monitoring seaming of a geomembrane liner.

TABLE 27.6 Tests for Raw Geomembrane Polymers

Property	Test Standard	Test Name
Density	ASTM D 792	Standard Test Method for Specific Gravity and Density of Plastics by the Density-Gradient Technique
Density	ASTM D 1505	Standard Test Method for Density of Plastics by the Density-Gradient Technique
Melt index	ASTM D 1238	Standard Test Method for Flow Rates of Thermoplastics by Extrusion Plastometer
Chemical identification methods (fingerprinting)	ASTM D 4595	Thermogravimetric analysis (TGA)
		Differential scanning calorimetry (DSC)
		Thermomechanical analysis (TMA)
		Infrared spectroscopy (IR)
		Chromatography (GC)
		Gel permeation chromatography (GPC)

application for geogrids is in the design of "piggyback" landfills, which are landfills built vertically over older, usually unlined landfills. Regulatory agencies often require that a liner system be installed between the old and new landfill. Since the old refuse is highly compressible, it provides a poor base for the new lining system. A geogrid may be used to support the lining system and bridge over voids that may occur beneath the liner as the underlying refuse components decompose.

As with other geosynthetics, geogrids have several physical, mechanical, and durability properties. Many of the test methods used for geotextiles and geomembranes also apply to geogrids. In particular, a key design parameter for reinforcement is tensile strength, which is typically reported from wide-width tensile tests (ASTM D 4595). The wide-width tests are performed with the specimen width incorporating typically a few ribs of the geogrid. The allowable tensile strength of geogrids (and of other geosynthetics used for soil reinforcement applications) is typically significantly less than its ultimate tensile strength. The allowable tensile strength is determined by dividing the ultimate tensile strength by partial safety factors for installation damage, creep deformation, chemical degradation, and biological degradation. The partial factors of safety for installation damage, chemical degradation, and biological degradation

TABLE 27.7 Standard Tests for Geomembranes

Property	Test Standard	Test Name
Thickness	ASTM D 5199	Standard Test Method for Measuring Nominal Thickness of Geotextiles and Geomembranes
Tensile behavior	ASTM D 412	Test Method for Rubber Properties in Tension
Tensile behavior	ASTM D 638	Standard Test Method for Tensile Properties of Plastics
Tensile behavior	ASTM D 882	Test Methods for Tensile Properties of Thin Plastic Sheeting
Tensile behavior	ASTM D 4885	Standard Test Method for Determining Performance Strength of Geomembranes by the Wide Strip Tensile Method
Tear resistance	ASTM D 1004	Test Method for Initial Tear Resistance of Plastic Film and Sheeting
Puncture resistance	ASTM D 2065	Puncture Resistance and Elongation Test
Environmental stress crack	ASTM D 1693	Standard Test Method for Environmental Stress-Cracking of Ethylene Plastics
Environmental stress crack	ASTM D 2552	Environmental Stress Rupture of Type III Polyethylene Under Constant Tensile Load
Carbon black	ASTM D 1603	Standard Test Method for Carbon Black in Olefin Plastics
Carbon black	ASTM D 3015	Standard Practice for Microscopical Examination of Pigment Dispersion in Plastic Compounds
Seam strength	ASTM D 4437	Standard Practice for Determining the Integrity of Field Seams Used in Joining Flexible Polymeric Sheet Geomembranes
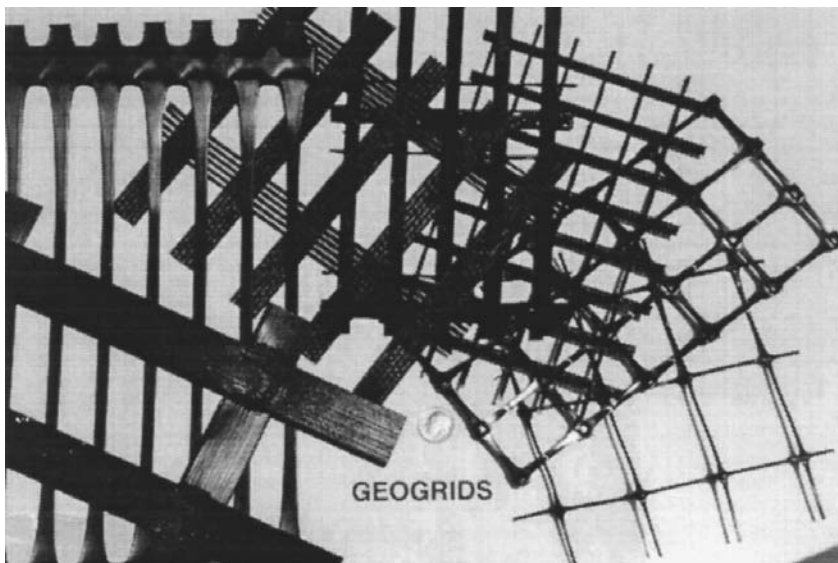
FIGURE 27.10 Typical geogrids.

range from 1.0 to 1.6, with the partial factor of safety for creep ranging from 1.5 to 3.5 (Koerner, 1994). Manufacturers typically provide recommendations for the allowable tensile strengths of their products.

27.3.4 Geosynthetic Clay Liners (GCLs)

Geosynthetic clay liners (GCLs) are rapidly expanding products in the geosynthetics market. GCLs are infiltration barriers consisting of a layer of unhydrated, loose granular or powdered bentonite placed between two or on top of one geosynthetic layer (geotextile or geomembrane). GCLs are produced in panels which are joined in the field by overlapping. They are generally used as an alternative to compacted clay liners.
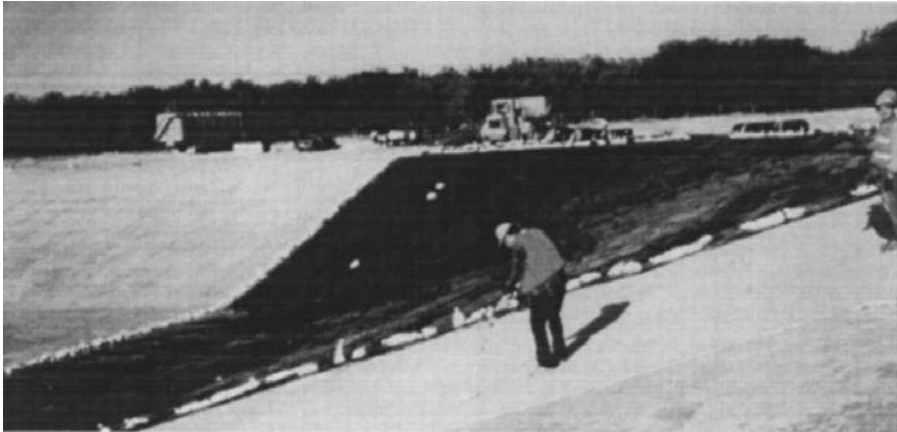
FIGURE 27.11 Installation of a GCL during construction of a landfill base liner.

Due to the inherent low shear strength of hydrated bentonite, GCL usage had initially been limited to applications where stability of the overlying materials was not a concern. In the late 1980s, however, methods were developed to reinforce the GCLs, producing a composite material with higher shear strength properties. This allowed the use of GCLs in landfill applications (Figure 27.11).

Some advantages of GCLs over compacted clay liners are that they occupy significantly less space to achieve equivalent performance, plus they are flexible, self-healing, and easy to install. In locations where low hydraulic conductivity clays are not readily available, they may offer significant construction cost savings. Moreover, since they are factory manufactured with good quality control, field construction quality assurance costs are typically less than with compacted clay liners.

Bentonite is a clay formed primarily from the mineral montmorillonite. While several types of montmorillonite exist, including calcium and sodium montmorillonite, the term bentonite typically refers to a sodium montmorillonite. Water is strongly attracted to the surface of the negatively charged montmorillonite crystal and is readily absorbed by it. In its unhydrated state, the montmorillonite crystals are densely packed. Once hydrated, the structure becomes very open and swells. The high water absorption and swell characteristics of bentonite lead to its low hydraulic conductivity and low hydrated shear strength.

Geosynthetic clay liners are manufactured by laying down a layer of dry bentonite, approximately 5-mm thick, on a geosynthetic material and attaching the bentonite to the geosynthetic. Two general configurations are currently employed in commercial processes (Figure 27.12): bentonite sandwiched between two geotextiles or bentonite glued to a geomembrane. The primary purpose of the geosynthetic component is to hold the bentonite together in a uniform layer and to permit transportation and installation of the GCL without loss of bentonite.

The outer geosynthetic layer of GCLs can be mechanically bonded using stitching or needle punching (resulting in reinforced GCLs). A different process consists in the use of an adhesive bond to glue the bentonite to the geosynthetic (resulting in unreinforced GCLs). The mechanical bonding of reinforced GCLs increases their internal shear strength. Geosynthetic clay liners contain approximately 5 kg/m^2 of bentonite that has a hydraulic conductivity of approximately 1×10^{-9} cm/s. Infiltration under unit hydraulic gradient through a material with hydraulic conductivity of 1×10^{-9} cm/s would result in an infiltration rate of 0.3 mm per year.

Since a GCL is a composite material, its relevant properties are those of the geotextile alone, of the bentonite alone, and of the composite. Geotextile properties were discussed in Section 27.3.1. The geotextile properties relevant to GCLs include mass per unit area, grab tensile, wide-width tensile, and puncture resistance. Relevant properties of the bentonite are obtained from free swell tests, which measure the absorption of water into a bentonite based on its volume change, and plate water absorption tests,
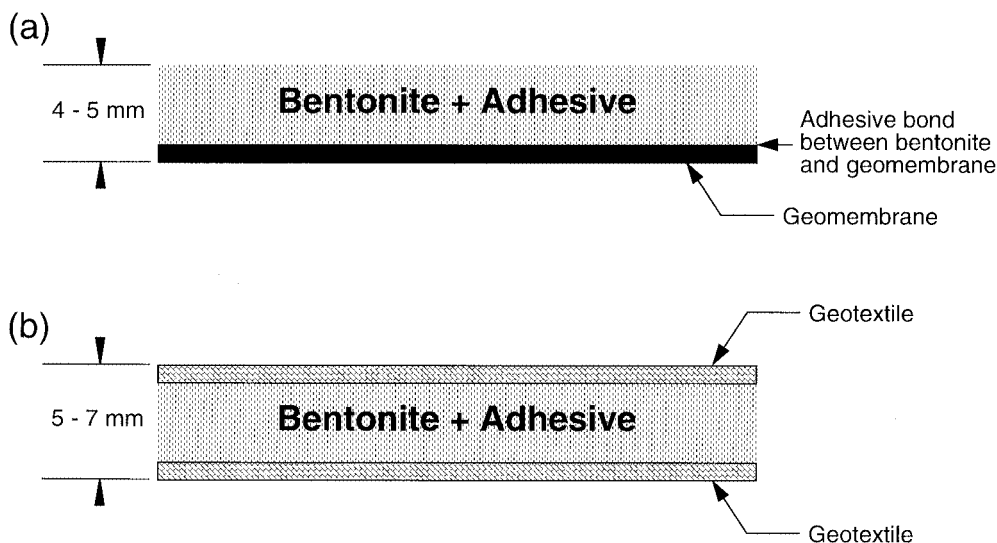
FIGURE 27.12 Typical GCL configurations: (a) bentonite glued to a geomembrane; (b) bentonite sandwiched between two geotextiles.

which measure the ability of powdered bentonite to absorb water. The relevant properties of the composite GCL material include bentonite content, which is simply a measure of the mass of bentonite per unit area of GCL, permeameter testing (ASTM D 5084) used to estimate the GCL hydraulic conductivity, tensile strength characterized either by grab tensile or wide-width tensile tests, and puncture resistance tests performed to assess the relative puncture resistance between GCLs and geomembranes or other geosynthetics.

27.3.5 Geocomposite Sheet Drains

A geocomposite consists of a combination of different types of geosynthetics. In particular, the geosynthetics industry has developed a number of geocomposite drains, which are polymeric drainage cores with continuously open flow channels sandwiched between geotextile filters. Geocomposite sheet drains are discussed in this section, while geocomposite strip (wick) drains are discussed in Section 27.3.6.

Geocomposite sheet drainage systems have been engineered to replace costly aggregate and/or perforated pipe subsurface drainage systems. They have reached rapid acceptance because they provide adequate drainage and reduce the material cost, installation time, and design complexity of conventional aggregate systems.

The core of geocomposite sheet drains are extruded sheets of plastic formed into a configuration that promotes drainage. The core of the geocomposite sheet drains are most commonly composed of polyethylene but may also be composed of polypropylene, polystyrene, high-impact polystyrene, or other materials. The structures of the core drainage products range from a dimpled core to a geonet. Geonets, a commonly used drainage product, generally consist of two sets of parallel solid or foamed extruded ribs that intersect at a constant angle to form an open net configuration. Channels are formed between the ribs to convey either liquids or gases. Figure 27.13 shows a number of geonets currently available in the market.

The geotextile serves as both a separator and a filter, and the geonet or built-up core serves as a drain. There may be geotextiles on both the top and bottom of the drainage core, and they may be different from one another. For example, the lower geotextile may be a thick needle-punched nonwoven geotextile used as a protective material for the underlying geomembrane, while the top geotextile may be a thinner
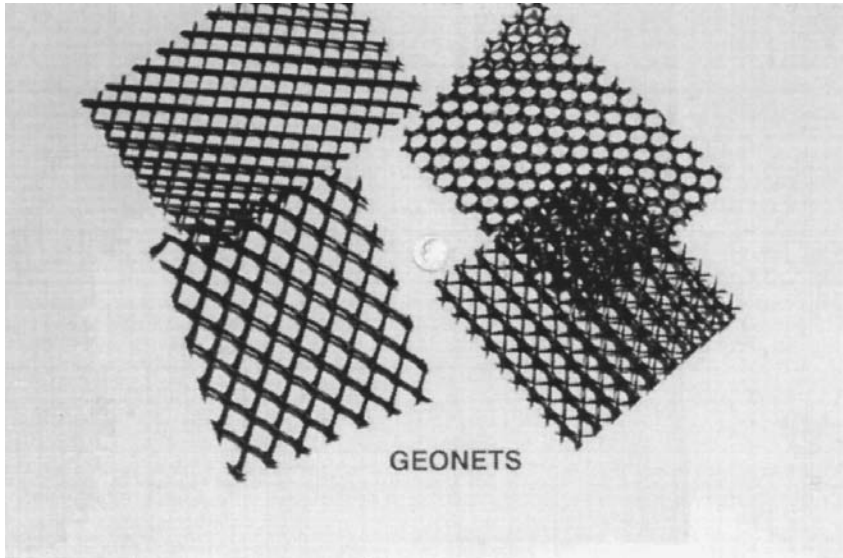
FIGURE 27.13 Typical geonets used as the core of geocomposite sheet drains.

FIGURE 27.14 Typical geocomposite sheet drains.

nonwoven or woven product. Composite drainage nets are typically formed by thermally bonding the geotextile and geonet. Glueing and solvent welding can also be used to bond the geosynthetic core to the geotextile. In producing geocomposite drainage nets, the melt temperatures of the geotextile and geonet must be compatible so that the properties of each material are retained. Figure 27.14 shows a number of available geocomposite sheet drainage materials.

Since the purpose of the core is drainage, the most important properties to include in specifications are thickness, crush strength, and transmissivity under load. Table 27.8 summarizes the tests commonly performed to evaluate the properties of geocomposite sheet drains. It is also important to evaluate filtration requirements for the geotextile. Design of geocomposite drains is covered by Holtz et al. (1997).

TABLE 27.8 Standard Tests for Geocomposite Drainage Nets

Property	Test Standard	Test Name
Thickness	ASTM D 5199	Standard Test Method for Measuring Nominal Thickness of Geotextiles and Geomembranes
Crush strength	ASTM D 1621	Standard Test Method for Compressive Properties of Rigid Cellular Plastics
Transmissivity	ASTM D 4716	Standard Test Method for Constant Head Hydraulic Transmissivity (In-Plane Flow) of Geotextiles and Geotextile Related Products

27.3.6 Geocomposite Strip (Wick) Drains

Geocomposite strip drains, also called "wick drains," have been developed to replace the use of sand drains in applications involving the increase in consolidation rate of soft, saturated fine-grained soils. Geocomposite strip drains actually do not wick moisture, but simply provide a conduit for pore water pressure-induced flow. They are placed vertically through high water content silts and clays to produce short drainage paths and thus increase the rate of consolidation. Other names commonly used for these products are "band shaped drains" and "prefabricated vertical drains."

Sand drains were originally introduced in the 1930s as a method for improvement of soft soil foundations. The method of rapid consolidation of saturated fine-grained soils using sand drains involves placement of vertical columns of sand (usually 200 to 450 mm in diameter) at spacings of 1.5 to 6.0 m centers throughout the subsurface to be dewatered. Now, the use of geocomposite strip drains dominates over the use of sand drains in projects involving dewatering of saturated fine-grained soils. Their lengths are site-specific but usually extend to the bottom of the soft layer(s). Once installed, a surcharge load is placed on the ground surface to mobilize excess pore water pressures. This surcharge load is placed in incremental lifts, which induce pore water pressures in the underlying soil. The pore water pressures are then dissipated through the vertical drains. Water takes the shortest drainage path (i.e., horizontally radial) to the vertical drain, at which point it flows vertically since the drain has a much higher hydraulic conductivity than the fine-grained soil being consolidated. The rate at which surcharge fill is added is critical in this process.

Most commercially available geocomposite strip drains have adequate capacity to drain the water expelled during consolidation of the fine-grained soils. Since their flow capacity is usually adequate, selection of the spacing of the vertical drains is governed by the consolidation rate required in the project. Hansbo's equation (Hansbo, 1979) is generally used to estimate the time required to achieve a desired percentage of consolidation as a function of the horizontal coefficient of consolidation of the foundation soil, the equivalent diameter of the geocomposite strip drain, and the spacing of the drains. As with geocomposite sheet drains, the geotextile covering or wrapping serves primarily a filtration function. Determining the filtration requirements for the geotextile is an essential element of the design.

Installation of geocomposite strip drains is very rapid and uses lightweight construction equipment fitted with hollow leads (called "lances" or mandrels) for insertion to the desired depth. The bottom of the lance should be covered by an expendable shoe which keeps soil out of the lance so as not to bind the strip drain within it. The allowable flow rate of geocomposite strip drains is determined by ASTM D4716 test method. Typical values of ultimate flow rate at a hydraulic gradient of 1.0 under 207 kPa normal stress vary from 1.5 to 3.0 m^3/sec.-m. This value must then be reduced on the basis of site-specific partial factors of safety. Specifications for geocomposite strip drains are covered in Holtz et al. (1997).

27.3.7 Geocells

Geocells (or cellular confinement systems) are three-dimensional, expandable panels made from HDPE or polyester strips. When expanded during installation, the interconnected strips form the walls of a flexible, three-dimensional cellular structure into which specified infill materials are placed and compacted (Figure 27.15). This creates a system that holds the infill material in place and prevents mass
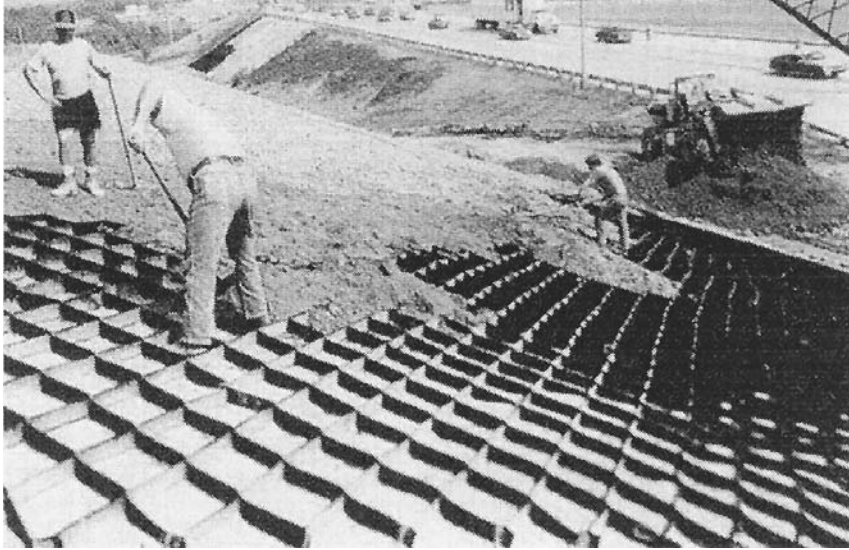
FIGURE 27.15 View of expanded geocell. (Photo courtesy of Presto Products Company.)

movements by providing tensile reinforcement. Cellular confinement systems improve the structural and functional behavior of soil infill materials.

Geocells were developed in the late 1970s and early 1980s for support of military vehicles on weak subgrade soils. The original type of geocell consists of HDPE strips 200 mm wide and approximately 1.2 mm thick. They are ultrasonically welded along their 200-mm width at approximately 330-mm intervals and are shipped to the job site in a collapsed configuration. At the job site they are placed directly on the subgrade surface and propped open in an accordion-like fashion with an external stretcher assembly. They are then generally filled with sand (although other infill materials can be selected) and compacted using a vibratory hand-operated plate compactor. Geocell applications include protection and stabilization of steep slope surfaces, protective linings of channels and hydraulic structures, static and dynamic load support on weak subgrade soils, and multilayered earth-retaining and water-retaining gravity structures.

Geocells have proven very effective in providing a stable foundation over soft soils. The cellular confinement system improves the load-deformation performance of infill materials because cohesionless materials gain considerable shear strength and stiffness under a confined condition. Confining stresses are effectively induced in a geocell by means of the hoop strength developed by the HDPE cell walls. The overall increase in the load-carrying performance of the system is provided through a combination of the cell wall strength, the passive resistance of the infill material in adjacent cells, and the frictional interaction between the infill soil and the cell walls. The cellular structure distributes concentrated loads to surrounding cells, thus reducing the stress on the subgrade directly beneath the loads.

Infill selection is primarily governed by the nature and intensity of anticipated working stresses, availability and cost of candidate materials, and aesthetic requirements for a fully vegetated appearance. Aggregates, vegetated topsoil, and concrete constitute typical geocell infill types. A complete cellular confinement system may also include geotextiles, geomembranes, geonets, geogrids, integral polymeric tendons, erosion-control blankets, and a variety of earth anchors.

27.3.8 Erosion Control Products

Erosion-control products represent one of the fastest-growing application areas in the geosynthetics industry. Erosion-control products provide protection against sheet and gully erosion on soil slopes either until vegetation is established or for long-term applications. These products can be classified as temporary

FIGURE 27.16 Erosion control mat placed to help establish the vegetation on the face of a 1H:1V reinforced soil slope.

degradable erosion control blankets, long-term nondegradable erosion control mats, and permanent hard armored systems.

Temporary degradable erosion control blankets are used to enhance the establishment of vegetation. These products are used where vegetation alone would provide sufficient site protection after the erosion control product has degraded. Some of these products are completely biodegradable (e.g., straw, hay, jute, and hydraulic mulches), while others are only partially biodegradable (e.g., erosion control meshes and nets). Long-term nondegradable erosion control mats provide permanent reinforcement of vegetation root structure. They are used in critical erosion-control applications where immediate high-performance erosion protection, followed by the permanent reinforcement of established vegetation is required. These soft armor-related products provide erosion control, aid in vegetative growth, and eventually become entangled with the vegetation to provide reinforcement to the root system. Finally, the permanent hard armored systems include geocell products with concrete infill, vegetated concrete block systems, and fabric-formed revetments.

Figure 27.16 shows an erosion control mat installed to help vegetation establishment on a steep reinforced soil slope. Installation of flexible erosion control products is straightforward. The products are usually placed on a prepared soil surface (e.g., facing of the reinforced embankment in Figure 27.16) by stapling or pinning them to the soil surface. Intimate contact between the blanket or mat and the soil is very important since water flow beneath the material has usually been the cause of poor functioning.

27.3.9 HDPE Vertical Barrier Systems

The use of geomembranes (Section 27.3.2) as horizontal barrier layers has been extended for the case of seepage control in remediation projects, in which vertically deployed geomembranes are used in vertical cutoff trenches. The construction process involves excavation of a trench and placement of a seamed geomembrane in the open trench. This procedure is usually not possible for deep trenches because of the potential collapse of the sidewalls, so the use of slurry to stabilize the trench becomes necessary. The mixture of water and bentonite clay balances the pressures exerted by the *in situ* soils. The geomembrane is placed in the slurry after trench excavation to the intended depth. Once the geomembrane is in place, the backfill can be introduced, displacing the slurry and forcing the geomembrane to the side of the trench.
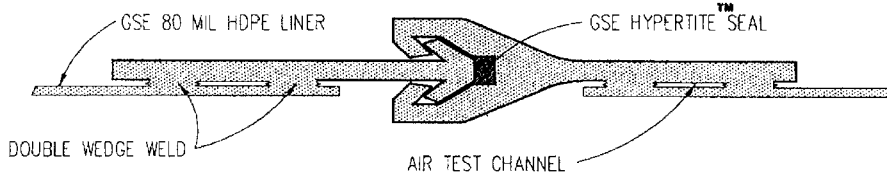
FIGURE 27.17 Interlocking system for HDPE vertical barriers. (Figure courtesy of GSE Lining Technology, Inc.)

Since installation of vertically deployed conventional geomembranes is difficult, other systems have become available. These systems involve the use of thick HDPE or nonplasticized PVC geomembranes in the form of tongue-and-groove sheeting. Sealing of the interlocks is often achieved using chloroprene-based, hydrophilic seals. Figure 27.17 shows one type of interlocking system with a hydrophilic seal. The seal is an extruded profile, typically 8 mm in diameter, which can expand up to 8 times its original volume when exposed to water. These interlocking HDPE vertical barrier systems have become increasingly used as an alternative to soil-bentonite slurry walls, especially in projects involving areas of limited access, high disposal costs, depths where performance of a slurry wall is questionable, and high concentrations of saline and/or chemicals.

An additional advantage of the HDPE vertical barrier system is that both a containment and a collection system (e.g., a geocomposite sheet drain) can be constructed in one trench. A recently developed method utilizes a biopolymer, or biodegradable, slurry. These slurries allow the HDPE panels and collection system to be installed in the same trench. Unlike bentonite, these slurries will either biodegrade or can be reversed to allow the collection system to drain clear and free of fines.

Another method of achieving construction of a containment and collection system in the same trench has been developed which utilizes a trenchless, vibratory method for installation of the HDPE panels. First, a collection trench is constructed to the required depth. This is followed with the installation of the geomembrane panel using modified pile driving techniques. Panel widths ranging typically between 0.91 m and 1.83 m are driven to depths up to 12 m. This construction method is most often reserved for sites on which excavation and disposal costs are high, access is limited, or the barrier is too close to a body of water. A case history in which this installation method was used for placement of an HDPE vertical barrier system is described in Section 27.5 .

A recent development in the installation of these systems is the use of a "one-pass" deep trencher. Installation of HDPE vertical barrier systems using this technology has proven to be fast and safe. Special trencher equipment can install a vertical geomembrane wall with a collection system consisting of HDPE pipe and a gravel fill in one trench, in one pass. Figure 27.18 shows the placement of an HDPE panel using this one-pass deep trencher.

27.4 Geosynthetic Applications in Landfill Design

The multiple use of geosynthetics in the design of modern municipal solid waste landfills is a good illustration of an application in which the different geosynthetics can be and have been used to perform all the functions discussed in Section 27.3. Virtually all the different types of geosynthetics discussed in Section 27.4 have been used in the design of both base and cover liner systems of landfill facilities. The extensive use of geosynthetics in modern landfills has been triggered by the economical and technical advantages that geosynthetics offer in relation to traditional liner systems. A geomembrane infiltration barrier of a couple of millimeters in thickness can provide performance equivalent to a soil infiltration barrier of sometimes several meters in thickness.

Landfill base liners are placed below the waste in order to minimize the release of liquids from the waste (i.e., leachate). Leachate is the main source of contamination of the soil underlying the landfill and, most importantly, of the groundwater. Landfill cover liners are placed above the final waste configuration to prevent water, usually from rain or snow, from percolating into the waste and producing

FIGURE 27.18 One pass trencher for installation of an HDPE vertical barrier system. (Photo courtesy of Ground-water Control, Inc.)

leachate. Waste containment systems employ geosynthetics to varying degrees. Figure 27.19 illustrates the extensive multiple uses of geosynthetics in both the cover and the base liner systems of a modern landfill facility.

The base liner system illustrated in Figure 27.19 is a double composite liner system. Double composite liner systems are used in some instances for containment of municipal solid waste and are frequently used for landfills designed to contain hazardous waste. The base liner system shown in the figure includes a geomembrane/GCL composite as the primary liner system and a geomembrane/compacted clay liner composite as the secondary system. The leak detection system, located between the primary and secondary liners, is a geotextile/geonet composite. The leachate collection system overlying the primary liner on the bottom of the liner system consists of gravel with a network of perforated pipes. A geotextile protection layer beneath the gravel provides a cushion to protect the primary geomembrane from puncture by stones in the overlying gravel. The leachate collection system overlying the primary liner on the side slopes of the liner system is a geocomposite sheet drain (geotextile/geonet composite) merging into the gravel on the base. A geotextile filter covers the entire footprint of the landfill and prevents clogging of the leachate collection and removal system. The groundwater level may be controlled at the bottom of the landfill by
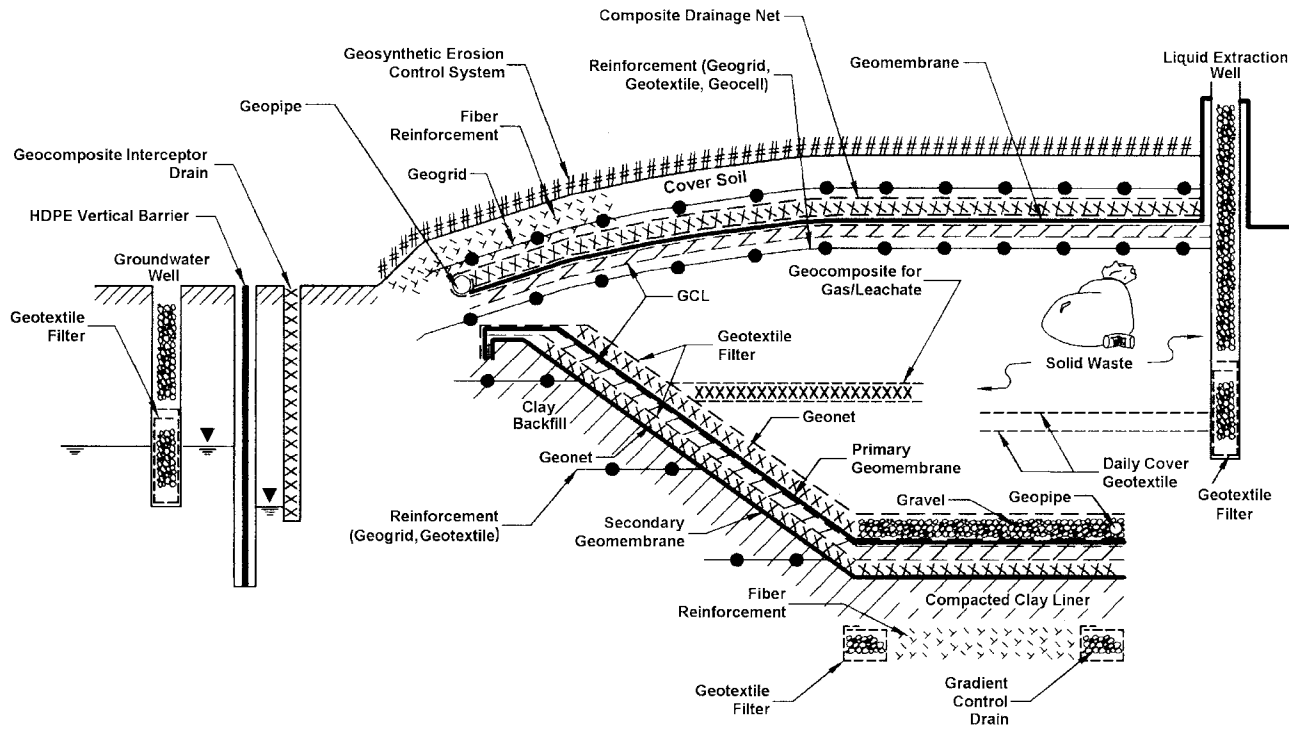
FIGURE 27.19 Multiple use of geosynthetics in landfill design.

gradient control drains built using geotextile filters. Moreover, the foundation soil below the bottom of the landfill may be stabilized as shown in the figure using randomly distributed fiber reinforcements, while the steep side soil slopes beneath the liner could also be reinforced using geogrids. Different types of geosynthetics (e.g. geogrids, geotextiles, fibers) could have been selected for stabilization of the foundation soils.

The cover system of the landfill illustrated in Figure 27.19 contains a composite geomembrane/GCL barrier layer. The drainage layer overlying the geomembrane is a geocomposite sheet drain (composite geotextile/geonet). In addition, the soil cover system may include geogrid, geotextile, or geocell reinforcements below the infiltration barrier system. This layer of reinforcements may be used to minimize the strains that could be induced in the barrier layers by differential settlements of the refuse or by a future vertical expansion of the landfill. In addition, the cover system could include a geogrid or geotextile reinforcement above the infiltration barrier to provide stability to the vegetative cover soil. Fiber reinforcement may also be used for stabilization of the steep portion of the vegetative cover soil. A geocomposite erosion control system above the vegetative cover soil is indicated in the figure and provides protection against sheet and gully erosion.

Figure 27.19 also illustrates the use of geosynthetics within the waste mass, which are used to facilitate waste placement during landfilling. Specifically, the figure illustrates the use of geotextiles as daily cover layers and of geocomposites within the waste mass for collection of gas and leachate. Geosynthetics can also be used as part of the groundwater and leachate collection well system. The use of geotextiles as filters in groundwater and leachate extraction wells is illustrated in the figure. Finally, the figure shows the use of an HDPE vertical barrier system and a geocomposite interceptor drain along the perimeter of the landfill facility. Although not all of the components shown in Figure 27.19 would normally be needed at any one landfill facility, the figure illustrates the many geosynthetic applications that can be considered in landfill design.

27.5 Case History of Vertical Barrier System

Although the use of geosynthetics in many geotechnical and environmental projects is related indirectly to groundwater applications (e.g., landfill liners, which prevent groundwater contamination), a geosynthetic application directly related to groundwater remediation and control is the use of HDPE panels as vertical barrier systems. A case history is presented herein to illustrate the use of HDPE panels as part of a remediation plan for a site contaminated with coal tar (Burson et al., 1997).

The site was a defunct manufactured gas plant in York, Pennsylvania. The site is surrounded by commercial and residential areas, and a creek (Codorus Creek) borders the site for a distance of approximately 305 meters. During years of operation and the subsequent closing of the manufactured gas plant, some process residuals migrated to subsurface soils and groundwater. Over time, the presence of coal tar-like material in the form of dense nonaqueous phase liquid (DNAPL), was observed seeping from the bank of the Codorus Creek. DNAPL was also noted in some monitoring wells on site.

Several remediation scenarios were evaluated with the purpose of intercepting the tar-like material migrating through the soil and into groundwater, encountered approximately 5.0 m below ground surface. A system consisting of a combination of soil improvement by jet grouting, a vertical barrier using HDPE panels, and a network of recovery wells was finally selected.

The use of vertical HDPE panels and trenchless technology allowed placement of the barrier as close as 3 m from the bank of Codorus Creek, which was considered not to be feasible with conventional slurry wall technology. The HDPE barrier system selected for this project was a 2-mm-thick geomembrane, which allowed for the vibratory, trenchless installation. Sealing of the interlocks was achieved with a chloroprene-based, hydrophilic seal (see Figure 27.17). HDPE panels were keyed into soil improved by jet grouting, as discussed below. The panels were installed using conventional vibratory pile driving equipment, without a trench, thus reducing the amount of contaminated spoils to be disposed of (Figure 27.20).

FIGURE 27.20 Installation of HDPE barrier wall utilizing conventional pile driving equipment. (Photo courtesy of Groundwater Control, Inc.)

In order to complete closure of the contaminated material, jet grouting was used to provide a seal to control DNAPL migration between the bottom of the HDPE panels and the irregular bedrock contact. Jet grouting consists of the high pressure injection of a cement and bentonite slurry horizontally into the soil strata in order to improve its mechanical and hydraulic properties. The containment wall was approximately 290 m in length. The soils along the alignment of the barrier system consisted of granular fills, with large amounts of cinder material. Also mixed into the fill were varying amounts of rubble and debris. These highly permeable soils were underlain by the competent bedrock. Holes were predrilled down to bedrock, and the jet grouting improvement was done by injecting the grout horizontally from the competent rock up to an elevation approximately 6 m below ground surface.

A groundwater recovery system was implemented once the barrier was completed. Since its installation in the fall of 1995, the HDPE panel jet grout barrier system has performed as intended.

For Further Information

Koerner (1994) provides an excellent, well-illustrated overview of the different types of geosynthetics and their applications.

Holtz, Christopher, and Berg (1997) provide well-documented practical design and construction information on the different uses of geosynthetic products.

Giroud et al. (1993, 1994) provide a two-volume comprehensive database on technical literature relative to geosynthetics, including technical papers from conferences, journals, books, theses, and research reports.

Technical advances on geosynthetics are also published in the two official journals of the IGS: *Geosynthetics International* and *Geotextiles and Geomembranes*. Similarly, the Geotechnical Fabrics Report (GFR), published by the Industrial Fabrics Association International (IFAI), provides updated informa-

tion, including the annual *Specifier's Guide*, which offers a summary of the properties of products available in the geosynthetics market.

The ASTM Standards on Geosynthetics, sponsored by ASTM Committee D-35 on Geosynthetics (ASTM, 1995), provides information on the standard test procedures for the different types of geosynthetics.

The Proceedings of the International Conferences on Geosynthetics, organized by the International Geosynthetic Society (IGS), offer a relevant source of information on the different topics related to geosynthetics. These international conferences are organized every four years. Equally relevant are the proceedings of conferences organized by the regional chapters of IGS. Particularly, the proceedings of conferences organized by the North American Geosynthetics Society (NAGS) every two years are an important source. Finally, the proceedings of the series of conferences organized by the Geosynthetics Research Institute (GRI) provide information on specific topics relevant to geosynthetic design.

Geosynthetic manufacturers' literature is also a valuable source of information, providing product-specific properties, suggested design methods, and recommended safety factors.

References

ASTM. 1995. *ASTM Standards on Geosynthetics.* Sponsored by ASTM Committee D-35 on Geosynthetics, Fourth Edition, 178p.

Bhatia, S. K. and Smith, J. L. 1996a. Geotextile characterization and pore-size distribution: Part I. A review of manufacturing processes, *Geosynthetics International.* 3, 1, 85-105.

Bhatia, S. K. and Smith, J. L. 1996b. Geotextile characterization and pore-size distribution: Part II. A review of test methods and results. *Geosynthetics International.* 3, 2, 155-180.

Bhatia, S. K., Smith, J. L., and Christopher, B. R. 1996. Geotextile characterization and pore-size distribution: Part III. Comparison of methods and application to design. *Geosynthetics International.* 3, 3, 301-328.

Burson, B., Baker, A. C., Jones, B., and Shailer, J. 1997. Development and installation of an innovative vertical containment system. Proceedings of the *Geosynthetics '97* Conference, Long Beach, California, March 1997, 1, 467-480.

Christopher, B. R. and Fischer, G.R. 1992. Geotextile filtration principles, practices and problems. *Journal of Geotextiles and Geomembranes.* 11, 4-6, 337-354.

Christopher, B. R. and Holtz, R. D. 1985. *Geotextiles Engineering Manual*, National Highway Institute, FHWA, Washington, D.C.

Christopher, B. R., Gill, S. A., Giroud, J. P., Juran, I., Mitchell, J., Schlosser, F., and Dunnicliff, J. 1989. *Design and Construction Guidelines for Reinforced Soil Structures, Volume I*, Report FHWA-RD-89-043. Federal Highway Administration, U.S. Department of Transportation, Washington, D.C.

Elias, V. and Christopher, B. R. 1997. Mechanically Stabilized Earth Walls and Reinforced Soil Slopes Design and Construction Guidelines, Publication No. FHWA-SA-96-071 Federal Highway Administration, U.S. Department of Transportation, Washington, D.C., 371 p.

Gerry, G. S. and Raymond, G. P. 1983. The in-plane permeability of geotextiles, *Geotechnical Testing Journal*, ASTM, 6, 4, December 1983, 181-189.

Giroud, J. P. (with cooperation of Beech, J.F. and Khatami, A.). 1993. *Geosynthetics Bibliography. Volume 1.* IGS, Industrial Fabrics Association International (IFAI) Publishers, St. Paul, MN.

Giroud, J. P. (with cooperation of Beech, J.F., Khatami, A., and Badu-Tweneboah, K.). 1994. *Geosynthetics Bibliography. Volume 2.* IGS, Industrial Fabrics Association International (IFAI) Publishers, St. Paul, MN.

Giroud, J. P. 1996. Granular filters and geotextile filters. Proceedings of *Geofilters '96*, Montreal, 565-680.

Haliburton, T. A. and Wood, P. D. 1982. Evaluation of U.S. Army Corps of Engineers gradient ratio test for geotextile performance. Proceedings of 2nd International Conference on Geotextiles, Las Vegas, NV, 1982, IFAI, 97-101.

Halse, Y., Koerner, R. M., and Lord, A. E., Jr. 1987. Filtration properties of geotextiles under long term testing, Proceedings of ASCE/Penn DOT Conf. on Advances in Geotechnical Engineering, Hershey, PA. 1-13.

Hansbo, S. 1979. Consolidation of clay by band shaped perforated drains. *Ground Engineering.* July 1979, 16-25.

Holtz, R. D., Christopher, B. R., and Berg, R. R. 1997. *Geosynthetic Engineering.* Bitech Publishers Ltd., Richmond, British Columbia, Canada.

Industrial Fabrics Association International. 1996. *North American Market for Geosynthetics — 1996.*

Koerner, R. M. 1994. *Designing with Geosynthetics.* Prentice Hall (3rd edition).

Koerner, R. M., Wilson-Fahmy, R. F., and Narejo, D. 1996. Puncture protection of geomembranes, Part III: Examples. *Geosynthetics International,* 3, 5, 655-675.

Mitchell, J. K. and Villet, W. C. B. 1987. Reinforcement of earth slopes and embankments. National Cooperative Highway Research Program, Report 290, Transportation Research Board, Washington, D.C., June.

Narejo, D., Koerner, R. M., and Wilson-Fahmy, R. F. 1996. Puncture protection of geomembranes, Part II: Experimental. *Geosynthetics International.* 3, 5, 629-653.

Vidal, H. 1969. La Terre Armée. *Annales de l'Institut Technique du Bâtiment et des Travaux Publics,* Series: Materials (38), 259-260, 1-59.

Williams, N. D. and Abouzakhm, M. A. 1989. Evaluation of geotextile/soil filtration characteristics using the hydraulic conductivity ratio analysis. *Journal of Geotextiles and Geomembranes.* 8, 1, 1-26.

Wilson-Fahmy, R. F., Narejo, D., and Koerner, R. M. 1996. Puncture protection of geomembranes, Part I: Theory. *Geosynthetics International.* 3, 5, 605-628.

Zornberg, J. G., Barrows, R. J., Christopher, B. R., and Wayne, M. H. 1995. Construction and instrumentation of a highway slope reinforced with high strength geotextiles. Proceedings of the *Geosynthetics '95* Conference, Nashville, Tennessee, February 1995, 1, 13-27.

Zornberg, J. G., Sitar, N., and Mitchell, J. K. 1997. Failure Mechanisms in Geosynthetically Reinforced Soil Slopes. Proceedings of the Fourteenth Int. Conf. on Soil Mechanics and Foundation Engineering, 3, Hamburg, Germany, September 1997, 1661-1664.

Glossary

Alloys, Polymeric   A blend of two or more polymers (e.g., a rubber and plastic) to improve a given property (e.g., impact strength).

Apparent Opening Size (AOS), O_{95}   For geotextile, a property which indicates the diameter of the approximate largest particle that would effectively pass through the geotextile. At least 95% of the openings apparently have that diameter or are smaller as measured by the dry sieve test.

Chemical Stability   Stability of a geosynthetic; ability to resist degradation from chemicals, such as acids, bases, solvents, oils, and oxidation agents, and chemical reactions, including those catalyzed by light.

Chlorosulfonated Polyethylene (CSPE)   Family of polymers that is produced by polyethylene reacting with chlorine and sulfur dioxide. Present CSPEs contain 25 to 43% chlorine and 1.0 to 1.4% sulfur.

Clogging   Movement by mechanical action or hydraulic flow of soil particles into the voids of fabric and retention therein, thereby reducing the hydraulic conductivity of the geotextile.

Cross-Machine Direction   The axis within the plane of a fabric perpendicular to the predominant axis of the direction of production.

Cross-Plane   The direction of a geosynthetic which is perpendicular to the plane of its manufactured direction. Referred to in hydraulic situations.

Fiber   Basic element of fabrics and other textile structures, characterized by having a length at least 100 times its diameter or width which can be spun into a yarn or otherwise made into a fabric.

Filament Yarn   The yarn made from continuous filament fibers.

Filtration In geotextiles, the process of retaining soil in place while allowing water to pass from soil. Removal of particle from a fluid stream.

Geocell A three-dimensional structure filled with soil, thereby forming a mattress for increased stability when used with loose or compressible subsoils.

Geocomposite A manufactured material using geotextiles, geogrids, and/or geomembranes in laminated or composite form. May or may not include natural materials.

Geogrid Open grid structure of orthogonal filaments and strands of polymeric material used primarily for tensile reinforcement.

Geomembrane Very low hydraulic conductivity synthetic membrane liners or barriers used with any geotechnical engineering-related material to control fluid migration in a man-made project, structure, or system.

Geonet A geosynthetic consisting of integrally connected parallel sets of ribs overlying similar sets at various angles for planar drainage of liquids or gases.

Geopipe Any plastic pipe used with foundation, soil, rock, earth, or any other subsurface material as an integral part of a man-made project, structure, or system.

Geosynthetic A planar product manufactured from polymeric material used with soil, rock, earth, or other geotechnical engineering-related material as an integral part of a man-made project, structure, or system.

Geosynthetic Clay Liner (GCL) Factory-manufactured hydraulic barriers consisting of a layer of bentonite clay or other very low permeability material supported by geotextiles and/or geomembranes, and mechanically held together by needling, stitching, or chemical adhesives.

Geotextile Any permeable textile used with foundation, soil, rock, earth, or any other geotechnical engineering-related material as an integral part of a man-made project, structure, or system.

Grab Test In fabric testing, a tension test in which only a part of the width of the specimen is gripped in the clamps.

Gradient Ratio The ratio of the average hydraulic gradient across the fabric and the 25 mm of soil immediately next to the fabric to the average hydraulic gradient across the 50 mm of soil between 25 and 75 mm above the fabric, as measured in a constant head permeability test.

Heat Bonded Thermally bonded by melting the fibers to form weld points.

Hot Wedge Common method of heat seaming of thermoplastic geomembranes by a fusing process wherein heat is delivered by a hot wedge passing between the opposing surfaces to be bonded.

Hydraulic Transmissivity For a geotextile or related product, the volumetric flow rate of water per unit width of specimen per unit gradient in a direction parallel to the plane of the specimen.

Index Test A test procedure which may contain a known bias but which may be used to establish an order for a set of specimens with respect to the property of interest.

In-Plane The direction of a geosynthetic that is parallel to its longitudinal, manufactured, or machine direction. Referred to in hydraulic situations.

Leachate Liquid that has percolated through or drained from solid waste or other human-emplaced materials and contains soluble, partially soluble, or miscible components removed from such waste.

Liner A layer of emplaced materials beneath a surface impoundment or landfill which serves to restrict the escape of waste or its constituents from the impoundment or landfill.

Machine Direction The direction in the plane of the fabric parallel to the direction of manufacture.

Mass Per Unit Area The proper term to represent and compare to amount of material per unit area (units are oz/yd^2 or g/m^2) of a geosynthetic.

Monofilament A single filament of a fiber (normally synthetic).

Mullen Burst Hydraulic bursting strength of textiles.

Multifilament A yarn consisting of many continuous filaments or strands.

Needlepunched In geotextiles, mechanical bonding of staple or filament fibers with barbed needles to form a compact fabric.

Nonwoven Fabric A textile structure produced by bonding or interlocking fibers, or both, accomplished by mechanical, thermal, or chemical means.

Permittivity   Of geotextiles and related products, the volumetric flow rate of water per unit cross-sectional area per unit head under laminar flow conditions, in the normal direction through a geotextile.

Plasticizer   A plasticizer is a material, frequently solvent-like, incorporated into a plastic or a rubber to increase its ease of workability, its flexibility, or distensibility.

Polyester Fiber   Generic name for a manufactured fiber in which the fiber-forming substance is any long-chain synthetic polymer composed of an ester of a dihydric alcohol and terephthalic acid.

Polyethylene   A polyolefin formed by bulk polymerization (for low-density) or solution polymerization (for high-density) where the ethylene monomer is placed in a reactor under high pressure and temperature.

Polymer   A macromolecular material formed by the chemical combination of monomers having either the same or different chemical composition. Plastics, rubbers, and textile fibers are all high-molecular-weight polymers.

Polyolefin   A family of polymeric materials that includes polypropylene and polyethylene, the former being very common in geotextiles, the latter in geomembranes.

Polyvinyl Chloride (PVC)   A synthetic thermoplastic polymer prepared from vinyl chloride.

Quality Assurance (QA)   A planned system of activities whose purpose is to provide a continuing evaluation of the quality control program, initiating corrective action where necessary. It is applicable to both the manufactured product and its field installation.

Quality Control (QC)   Actions that provide a means of controlling and measuring the characteristics of (both) the manufactured and the field-installed product.

Separation   The function of geosynthetics as a partition between two adjacent materials (usually dissimilar) to prevent mixing of the two materials.

Specification   A precise statement of a set of requirements to be satisfied by a material, product, system or service that indicates the procedures for determining whether each of the requirements is satisfied.

Spun-Bonded Fabrics   Fabric formed by continuous filaments which have been spun (extruded), drawn, laid into a web and bonded (chemical, mechanical, or thermal bonding) together in one continuous process.

Staple Fibers   Fiber of short lengths frequently used to make needle-punched nonwoven fabrics.

Subgrade Intrusion   Localized aggregate penetration of a soft cohesive subgrade and resulting displacement of the subgrade into the cohesionless material.

Subgrade Pumping   The displacement of cohesive or low-cohesion fines from a saturated subgrade into overlying aggregate as the result of hydraulic forces created by transmittal of wheel-load stresses to the subgrade.

Survivability   The ability of a geosynthetic to be placed and to perform its intended function without undergoing degradation.

Tensile Strength   The maximum resistance to deformation developed for a specific material when subjected to tension by an external force.

Trapezoid Tear Test   Test method used to measure the tearing strength of geotextiles.

Transmissivity   For a geosynthetic, the volumetric flow rate per unit thickness under laminar flow conditions, in the in-plane direction of the fabric or geocomposite.

Ultraviolet Degradation   The breakdown of polymeric structure when exposed to natural light.

Woven Geotextile   A planar geotextile structure produced by interlacing two or more sets of elements such as yarns, fibers, rovings of filaments where the elements pass each other usually at right angles and one set of elements is parallel to the fabric axis.

Wide-Width Strip Tensile Test   A uniaxial tensile test in which the entire width of a 200-mm-wide specimen is gripped in the clamps and the gauge length is 100 mm.

Yarn   A generic term for continuous-strand strands (1 or more) of textile filaments, monofilaments, or slit form suitable for knitting, weaving, or otherwise intertwining or bonding to form a textile fabric.

Sources for these and other definitions of terms can be found in ASTM (1997) and Koerner (1994).

APPENDIX A

Mathematical Appendices (A), Figures (F), and Tables (T)

APPENDIX B

List of Figures (F) and Tables (T) of Properties, Methods, and Techniques

Appendix C

Mathematics, Symbols, and Physical Constants

Greek Alphabet

	Greek letter		Greek name	English equivalent		Greek letter		Greek name	English equivalent
A	α		Alpha	a	N	ν		Nu	n
B	β		Beta	b	Ξ	ξ		Xi	x
Γ	γ		Gamma	g	O	o		Omicron	ŏ
Δ	δ		Delta	d	Π	π		Pi	p
E	ε		Epsilon	ĕ	P	ρ		Rho	r
Z	ζ		Zeta	z	Σ	σ		Sigma	s
H	η		Eta	ē	T	τ		Tau	t
Θ	θ	ϑ	Theta	th	Y	υ		Upsilon	u
I	ι		Iota	i	Φ	ϕ	φ	Phi	ph
K	κ		Kappa	k	X	χ		Chi	ch
Λ	λ		Lambda	l	Ψ	ψ		Psi	ps
M	μ		Mu	m	Ω	ω		Omega	ō

International System of Units (SI)

The International System of units (SI) was adopted by the 11th General Conference on Weights and Measures (CGPM) in 1960. It is a coherent system of units built form seven *SI base units,* one for each of the seven dimensionally independent base quantities: they are the meter, kilogram, second, ampere,

kelvin, mole, and candela, for the dimensions length, mass, time, electric current, thermodynamic temperature, amount of substance, and luminous intensity, respectively. The definitions of the SI base units are given below. The *SI derived units* are expressed as products of powers of the base units, analogous to the corresponding relations between physical quantities but with numerical factors equal to unity.

In the International System there is only one SI unit for each physical quantity. This is either the appropriate SI base unit itself or the appropriate SI derived unit. However, any of the approved decimal prefixes, called *SI prefixes,* may be used to construct decimal multiples or submultiples of SI units.

It is recommended that only SI units be used in science and technology (with SI prefixes where appropriate). Where there are special reasons for making an exception to this rule, it is recommended always to define the units used in terms of SI units. This section is based on information supplied by IUPAC.

Definitions of SI Base Units

Meter—The meter is the length of path traveled by light in vacuum during a time interval of 1/299, 792, 458 of a second (17th CGPM, 1983).

Kilogram—The kilogram is the unit of mass; it is equal to the mass of the international prototype of the kilogram (3rd CGPM, 1901).

Second—The second is the duration of 9, 192, 631, 770 periods of the radiation corresponding to the transition between the two hyperfine levels of the ground state of the cesium-133 atom (13th CGPM, 1967).

Ampere—The ampere is that constant current which, if maintained in two straight parallel conductors of infinite length, of negligible circular cross-section, and placed 1 meter apart in vacuum, would produce between these conductors a force equal to 2×10^{-7} newton per meter of length (9th CGPM, 1948).

Kelvin—The kelvin, unit of thermodynamic temperature, is the fraction 1/273.16 of the thermodynamic temperature of the triple point of water (13th CGPM, 1967).

Mole—The mole is the amount of substance of a system which contains as many elementary entities as there are atoms in 0.012 kilogram of carbon-12. When the mole is used, the elementary entities must be specified and may be atoms, molecules, ions, electrons, or other particles, or specified groups of such particles (14th CGPM, 1971).

Examples of the use of the mole:

1 mol of H_2 contains about 6.022×10^{23} H_2 molecules, or 12.044×10^{23} H atoms

1 mol of HgCl has a mass of 236.04 g

1 mol of Hg_2Cl_2 has a mass of 472.08 g

1 mol of Hg_2^{2+} has a mass of 401.18 g and a charge of 192.97 kC

1 mol of $Fe_{0.91}S$ has a mass of 82.88 g

1 mol of e^- has a mass of 548.60 µg and a charge of –96.49 kC

1 mol of photons whose frequency is 10^{14} Hz has energy of about 39.90 kJ

Candela—The candela is the luminous intensity, in a given direction, of a source that emits monochromatic radiation of frequency 540×10^{12} hertz and that has a radiant intensity in that direction of (1/683) watt per steradian (16th CGPM, 1979).

Names and Symbols for the SI Base Units

Physical quantity	Name of SI unit	Symbol for SI unit
length	meter	m
mass	kilogram	kg
time	second	s
electric current	ampere	A
thermodynamic temperature	kelvin	K
amount of substance	mole	mol
luminous intensity	candela	cd

SI Derived Units with Special Names and Symbols

Physical quantity	Name of SI unit	Symbol for SI unit	Expression in terms of SI base units	
frequency[1]	hertz	Hz	s^{-1}	
force	newton	N	$m\ kg\ s^{-2}$	
pressure, stress	pascal	Pa	$N\ m^{-2}$	$= m^{-1}\ kg\ s^{-2}$
energy, work, heat	joule	J	$N\ m$	$= m^2\ kg\ s^{-2}$
power, radiant flux	watt	W	$J\ s^{-1}$	$= m^2\ kg\ s^{-3}$
electric charge	coulomb	C	$A\ s$	
electric potential, electromotive force	volt	V	$J\ C^{-1}$	$= m^2\ kg\ s^{-3}\ A^{-1}$
electric resistance	ohm	Ω	$V\ A^{-1}$	$= m^2\ kg\ s^{-3}\ A^{-2}$
electric conductance	siemens	S	Ω^{-1}	$= m^{-2}\ kg^{-1}\ s^3\ A^2$
electric capacitance	farad	F	$C\ V^{-1}$	$= m^{-2}\ kg^{-1}\ s^4\ A^2$
magnetic flux density	tesla	T	$V\ s\ m^{-2}$	$= kg\ s^{-2}\ A^{-1}$
magnetic flux	weber	Wb	$V\ s$	$= m^2\ kg\ s^{-2}\ A^{-1}$
inductance	henry	H	$V\ A^{-1}\ s$	$= m^2\ kg\ s^{-2}\ A^{-2}$
Celsius temperature[2]	degree Celsius	°C	K	
luminous flux	lumen	lm	$cd\ sr$	
illuminance	lux	lx	$cd\ sr\ m^{-2}$	
activity (radioactive)	becquerel	Bq	s^{-1}	
absorbed dose (of radiation)	gray	Gy	$J\ kg^{-1}$	$= m^2\ s^{-2}$
dose equivalent (dose equivalent index)	sievert	Sv	$J\ kg^{-1}$	$= m^2\ s^{-2}$
plane angle	radian	rad	1	$= m\ m^{-1}$
solid angle	steradian	sr	1	$= m^2\ m^{-2}$

[1] For radial (circular) frequency and for angular velocity the unit rad s^{-1}, or simply s^{-1}, should be used, and this may not be simplified to Hz. The unit Hz should be used only for frequency in the sense of cycles per second.

[2] The Celsius temperature θ is defined by the equation:
$$\theta/°C = T/K - 273.15$$

The SI unit of Celsius temperature interval is the degree Celsius, °C, which is equal to the kelvin, K. °C should be treated as a single symbol, with no space between the ° sign and the letter C. (The symbol °K, and the symbol °, should no longer be used.)

Units in Use Together with the SI

These units are not part of the SI, but it is recognized that they will continue to be used in appropriate contexts. SI prefixes may be attached to some of these units, such as milliliter, ml; millibar, mbar; megaelectronvolt, MeV; kilotonne, ktonne.

Physical quantity	Name of unit	Symbol for unit	Value in SI units	
time	minute	min	60 s	
time	hour	h	3600 s	
time	day	d	86 400 s	
plane angle	degree	°	$(\pi/180)$ rad	
plane angle	minute	′	$(\pi/10\ 800)$ rad	
plane angle	second	″	$(\pi/648\ 000)$ rad	
length	ångstrom[1]	Å	10^{-10} m	
area	barn	b	10^{-28} m²	
volume	litre	l, L	dm^3	$= 10^{-3}$ m³
mass	tonne	t	Mg	$= 10^3$ kg
pressure	bar[1]	bar	10^5 Pa	$= 10^5$ N m⁻²

Physical quantity	Name of unit	Symbol for unit	Value in SI units
energy	electronvolt[2]	eV (= $e \times$ V)	$\approx 1.60218 \times 10^{-19}$ J
mass	unified atomic mass unit[2,3]	u (=m_a(^{12}C)/12)	$\approx 1.66054 \times 10^{-27}$ kg

[1] The ångstrom and the bar are approved by CIPM for "temporary use with SI units," until CIPM makes a further recommendation. However, they should not be introduced where they are not used at present.

[2] The values of these units in terms of the corresponding SI units are not exact, since they depend on the values of the physical constants e (for the electronvolt) and N_a (for the unified atomic mass unit), which are determined by experiment.

[3] The unified atomic mass unit is also sometimes called the dalton, with symbol Da, although the name and symbol have not been approved by CGPM.

Conversion Constants and Multipliers

Recommended Decimal Multiples and Submultiples

Multiples and submultiples	Prefixes	Symbols	Multiples and submultiples	Prefixes	Symbols
10^{18}	exa	E	10^{-1}	deci	d
10^{15}	peta	P	10^{-2}	centi	c
10^{12}	tera	T	10^{-3}	milli	m
10^{9}	giga	G	10^{-6}	micro	μ (Greek mu)
10^{6}	mega	M	10^{-9}	nano	n
10^{3}	kilo	k	10^{-12}	pico	p
10^{2}	hecto	h	10^{-15}	femto	f
10	deca	da	10^{-18}	atto	a

Conversion Factors—Metric to English

To obtain	Multiply	By
Inches	Centimeters	0.3937007874
Feet	Meters	3.280839895
Yards	Meters	1.093613298
Miles	Kilometers	0.6213711922
Ounces	Grams	$3.527396195 \times 10^{-2}$
Pounds	Kilogram	2.204622622
Gallons (U.S. Liquid)	Liters	0.2641720524
Fluid ounces	Milliliters (cc)	$3.381402270 \times 10^{-2}$
Square inches	Square centimeters	0.155003100
Square feet	Square meters	10.76391042
Square yards	Square meters	1.195990046
Cubic inches	Milliliters (cc)	$6.102374409 \times 10^{-2}$
Cubic feet	Cubic meters	35.31466672
Cubic yards	Cubic meters	1.307950619

Conversion Factors—English to Metric[*]

To obtain	Multiply	By
Microns	Mils	**25.4**
Centimeters	Inches	**2.54**
Meters	Feet	**0.3048**
Meters	Yards	**0.9144**
Kilometers	Miles	**1.609344**
Grams	Ounces	28.34952313
Kilograms	Pounds	**0.45359237**
Liters	Gallons (U.S. Liquid)	**3.785411784**
Millimeters (cc)	Fluid ounces	29.57352956
Square centimeters	Square inches	**6.4516**
Square meters	Square feet	**0.09290304**
Square meters	Square yards	**0.83612736**
Milliliters (cc)	Cubic inches	**16.387064**
Cubic meters	Cubic feet	$2.831684659 \times 10^{-2}$
Cubic meters	Cubic yards	0.764554858

Conversion Factors—General[*]

To obtain	Multiply	By
Atmospheres	Feet of water @ 4°C	2.950×10^{-2}
Atmospheres	Inches of mercury @ 0°C	3.342×10^{-2}
Atmospheres	Pounds per square inch	6.804×10^{-2}
BTU	Foot-pounds	1.285×10^{-3}
BTU	Joules	9.480×10^{-4}
Cubic feet	Cords	**128**
Degree (angle)	Radians	57.2958
Ergs	Foot-pounds	1.356×10^{7}
Feet	Miles	**5280**
Feet of water @ 4°C	Atmospheres	33.90
Foot-pounds	Horsepower-hours	1.98×10^{6}
Foot-pounds	Kilowatt-hours	2.655×10^{6}
Foot-pounds per min	Horsepower	3.3×10^{4}
Horsepower	Foot-pounds per sec	1.818×10^{-3}
Inches of mercury @ 0°C	Pounds per square inch	2.036
Joules	BTU	1054.8
Joules	Foot-pounds	1.35582
Kilowatts	BTU per min	1.758×10^{-2}
Kilowatts	Foot-pounds per min	2.26×10^{-5}
Kilowatts	Horsepower	0.745712
Knots	Miles per hour	0.86897624
Miles	Feet	1.894×10^{-4}
Nautical miles	Miles	0.86897624
Radians	Degrees	1.745×10^{-2}
Square feet	Acres	**43560**
Watts	BTU per min	17.5796

---

[*] Boldface numbers are exact; others are given to ten significant figures where so indicated by the multiplier factor.

Temperature Factors

$$°F = 9/5 \ (°C) + 32$$
Fahrenheit temperature = 1.8 (temperature in kelvins) − 459.67
$$°C = 5/9 \ [(°F) − 32]$$
Celsius temperature = temperature in kelvins − 273.15
Fahrenheit temperature = 1.8 (Celsius temperature) + 32

Conversion of Temperatures

From	To	
°Celsius	°Fahrenheit	$t_F = (t_C \times 1.8) + 32$
	Kelvin	$T_K = t_C + 273.15$
	°Rankine	$T_R = (t_C + 273.15) \times 18$
°Fahrenheit	°Celsius	$t_C = \frac{t_F - 32}{1.8}$
	Kelvin	$T_k = \frac{t_F - 32}{1.8} + 273.15$
	°Rankine	$T_R = t_F + 459.67$
Kelvin	°Celsius	$t_C = T_K − 273.15$
	°Rankine	$T_R = T_K \times 1.8$
°Rankine	Kelvin	$T_K = \frac{T_R}{1.8}$
	°Fahrenheit	$t_F = T_R − 459.67$

Physical Constants

General

Equatorial radius of the earth = 6378.388 km = 3963.34 miles (statute).
Polar radius of the earth, 6356.912 km = 3949.99 miles (statute).
1 degree of latitude at 40° = 69 miles.
1 international nautical mile = 1.15078 miles (statute) = 1852 m = 6076.115 ft.
Mean density of the earth = 5.522 g/cm^3 = 344.7 lb/ft^3
Constant of gravitation $(6.673 \pm 0.003) \times 10^{-8}$ cm^3 gm^{-1} s^{-2}.
Acceleration due to gravity at sea level, latitude 45° = 980.6194 cm/s^2 = 32.1726 ft/s^2.
Length of seconds pendulum at sea level, latitude 45° = 99.3575 cm = 39.1171 in.
1 knot (international) = 101.269 ft/min = 1.6878 ft/s = 1.1508 miles (statute)/h.
1 micron = 10^{-4} cm.
1 ångstrom = 10^{-8} cm.
Mass of hydrogen atom = $(1.67339 \pm 0.0031) \times 10^{-24}$ g.
Density of mercury at 0°C = 13.5955 g/ml.
Density of water at 3.98°C = 1.000000 g/ml.
Density, maximum, of water, at 3.98°C = 0.999973 g/cm^3.
Density of dry air at 0°C, 760 mm = 1.2929 g/l.
Velocity of sound in dry air at 0°C = 331.36 m/s – 1087.1 ft/s.
Velocity of light in vacuum = $(2.997925 \pm 0.000002) \times 10^{10}$ cm/s.
Heat of fusion of water 0°C = 79.71 cal/g.
Heat of vaporization of water 100°C = 539.55 cal/g.
Electrochemical equivalent of silver 0.001118 g/s international amp.
Absolute wavelength of red cadmium light in air at 15°C, 760 mm pressure = 6438.4696 Å.
Wavelength of orange-red line of krypton 86 = 6057.802 Å.

π Constants

$$\pi = 3.14159\ 26535\ 89793\ 23846\ 26433\ 83279\ 50288\ 41971\ 69399\ 37511$$
$$1/\pi = 0.31830\ 98861\ 83790\ 67153\ 77675\ 26745\ 02872\ 40689\ 19291\ 48091$$
$$\pi^2 = 9.8690\ \ 44010\ 89358\ 61883\ 44909\ 99876\ 15113\ 53136\ 99407\ 24079$$
$$\log_e\pi = 1.14472\ 98858\ 49400\ 17414\ 34273\ 51353\ 05871\ 16472\ 94812\ 91531$$
$$\log_{10}\pi = 0.49714\ 98726\ 94133\ 85435\ 12682\ 88290\ 89887\ 36516\ 78324\ 38044$$
$$\log_{10}\sqrt{2\pi} = 0.39908\ 99341\ 79057\ 52478\ 25035\ 91507\ 69595\ 02099\ 34102\ 92128$$

Constants Involving e

$$e = 2.71828\ 18284\ 59045\ 23536\ 02874\ 71352\ 66249\ 77572\ 47093\ 69996$$
$$1/e = 0.36787\ 94411\ 71442\ 32159\ 55237\ 70161\ 46086\ 74458\ 11131\ 03177$$
$$e^2 = 7.38905\ 60989\ 30650\ 22723\ 04274\ 60575\ 00781\ 31803\ 15570\ 55185$$
$$M = \log_{10}e = 0.43429\ 44819\ 03251\ 82765\ 11289\ 18916\ 60508\ 22943\ 97005\ 80367$$
$$1/M \cdot = \log_e 10 = 2.30258\ 50929\ 94045\ 68401\ 79914\ 54684\ 36420\ 67011\ 01488\ 62877$$
$$\log_{10}M = 9.63778\ 43113\ 00536\ 78912\ 29674\ 98645\ -10$$

Numerical Constants

$$\sqrt{2} = 1.41421\ 35623\ 73095\ 04880\ 16887\ 24209\ 69807\ 85696\ 71875\ 37695$$
$$\sqrt[3]{2} = 1.25992\ 10498\ 94873\ 16476\ 72106\ 07278\ 22835\ 05702\ 51464\ 70151$$
$$\log_e 2 = 0.69314\ 71805\ 59945\ 30941\ 72321\ 21458\ 17656\ 80755\ 00134\ 36026$$
$$\log_{10}2 = 0.30102\ 99956\ 63981\ 19521\ 37388\ 94724\ 49302\ 67881\ 89881\ 46211$$
$$\sqrt{3} = 1.73205\ 08075\ 68877\ 29352\ 74463\ 41505\ 87236\ 69428\ 05253\ 81039$$
$$\sqrt[3]{3} = 1.44224\ 95703\ 07408\ 38232\ 16383\ 10780\ 10958\ 83918\ 69253\ 49935$$
$$\log_e 3 = 1.09861\ 22886\ 68109\ 69139\ 52452\ 36922\ 52570\ 46474\ 90557\ 82275$$
$$\log_{10}3 = 0.47712\ 12547\ 19662\ 43729\ 50279\ 03255\ 11530\ 92001\ 28864\ 19070$$

Conversion Factors for Intrinsic Permeability, k [L²]

	cm²	ft²	darcy
1 cm²	1	1.076391E-3	1.01325 E+8*
1 ft²	9.290304 E+2*	1	9.413401E+10
1 darcy	9.869233 E-9	1.062315E-11	1

* Indicates an exact value.

Example: to covert 0.02 ft² in darcy multiply by 9.413401E+10 or 0.02 × 9.413401E+10 = 1.882680E+9 darcy.

$$k = \frac{\mu(Q/A)}{dp/dl}$$

$$1\,darcy = \frac{\dfrac{1\,centipoise \times 1\,cm^3/\sec}{1\,cm^2}}{1\,atmosphere/\text{cm}} = \frac{\dfrac{(0.01\,dyne\ \sec/cm^2) \times 1\,cm^3/\sec}{1\,cm^2}}{(1.01325 \times 10^6\,dynes/cm^2)/\text{cm}} = 9.869233 \times 10^{-9}\,cm^2$$

Conversion Factors for Transmissivity [L²/T]

	m²/s	*m²/min*	*m²/day*	*ft²/s*	*ft²/day*	*USgpd/ft*	*UKgpd/ft*
m²/s	1	6.00E+1*	8.64E+4*	1.076391E+1	9.300019E+5	6.956897E+6	5.792826E+6
m²/min	1.666667E-2	1	1.44E+3*	1.793985E-1	1.550003E+4	1.159483E+5	9.654710E+4
m²/day	1.157407E-5	6.944444E-4	1	1.245823E-4	1.076391E+1	8.051964E+1	6.704660E+1
ft²/s	9.290304E-2*	5.574182	8.026823E+3	1	8.64E+4*	6.463169E+5	5.381703E+5
ft²/day	1.075267E-6	6.4516E-5*	9.290304E-2*	1.157407E-5	1	7.480519	6.228823
USgpd/ft	1.437423E-7	8.624535E-6	1.241933E-2*	1.547229E-6	1.336806E-1	1	8.326725E-1
UKgpd/ft	1.726273E-7	1.035764E-5	1.491500E-2	1.858148E-6	1.605440E-1	1.200952	1

* Indicates exact value

Example: to convert a transmissivity of 2.5 ft²/day to m²/min multiply by 6.4516E-5 or 2.5 × 0.000064516E-5 = 0.00016129 m²/min.

Conversion Factors for Flows, Q[L³/T]

	m³/sec	*m³/day*	*liter/sec*	*ft³/sec*	*ft³/day*	*ac-ft/day*	*gal/min*	*gal/day*	*mgd*
m³/sec	1	8.64E+4*	1.0E+3*	3.531466E+1	3.051187E+6	7.004521E+1	1.585032E+4	2.282446E+7	2.282446E+1
m³/day	1.157407E-5	1	1.157407E-2	4.087345E-4	3.531466E+1	8.107085E-4	1.834528E-1	2.641720E+2	2.641720E-4
liter/sec	1.0E-3*	8.64E+1*	1	3.531466E-2	3.051187E+3	7.004521E-2	1.585032E+1	2.282446E+4	2.282446E-2
ft³/sec	2.831685E-2	2.446576E+3	2.831685E+1	1	8.6400E+4*	1.983471	4.488312E+2	6.463169E+5	6.463169E-1
ft³/day	3.277413E-7	2.831685E-2	3.277413E-4	1.157407E-5	1	2.295684E-5	5.194805E-3	7.480519	7.480519E-6
ac-ft/day	1.427649E-2	1.233489E+3	1.427649E+1	5.041667E-1	4.356E+4*	1	2.262857E+2	3.258514E+5	3.258514E-1
gal/min	6.309020E-5	5.450993	6.309020E-2	2.228009E-3	1.925E+2	4.419192E-3	1	1.440E+3*	1.44E-3*
gal/day	4.381264E-8	3.785412E-3	4.381264E-5	1.547229E-6	1.336806E-1	3.068883E-6	6.944444E-4	1	1.0E-6*
mgd	4.381264E-2	3.785412E+3	4.381264E+1	1.547229	1.336806E+5	3.068883	6.944444E+2	1.0E+6*	1

* Denotes an exact value

Example. To convert 7.5 gal/min to liter per sec: multiply by 6.309020E-2 or 7.5 × 0.06309020 = 0.473177 liter/sec.

Conversion Factors for Hydraulic Conductivity, K [L/T]

	cm/s	m/s	m/day	ft/s	ft/day	ft/yr	USgpd/ft²	UKgpd/ft²
cm/s	1	1.00E-2*	8.64E+2*	3.280840E-2	2.834646E+3	1.034646E+6	2.120462E+4	1.765651E+4
m/s	1.00E+2*	1	8.64E+4*	3.280840	2.834646E+5	1.034646E+8	2.120462E+6	1.765651E+6
m/day	1.157407E-3	1.157407E-5	1	3.797268E-5	3.280840	1.197507E+3	2.454238E+1	2.043577E+1
ft/s	3.048E+1*	3.048E-1*	2.633472E+4	1	8.64E+4*	3.1536E+7*	6.463169E+5	5.381703E+5
ft/day	3.527778E-4	3.527778E-6	0.3048*	1.157407E-5	1	3.65E+2*	7.480519	6.228823
ft/yr	9.665145E-7	9.665145E-9	8.350685E-4	3.170979E-8	2.739726E-3	1	2.049457E-2	1.706527E-2
USgpd/ft²	4.715953E-5	4.715953E-7	4.074584E-2	1.547229E-6	1.336806E-1	4.879340E+1	1	8.326725E-1
UKgpd/ft²	5.663635E-5	5.663635E-7	4.893381E-2	1.858148E-6	1.605440E-1	5.859855E+1	1.200952	1

* Indicates an exact value

Example. Convert a hydraulic conductivity of 2.5 ft/day to cm/s : multiply by 3.527778E-4 or 2.5 × 3.527778E-4 = 8.819445E-4 = 0.0008819445 cm/sec.