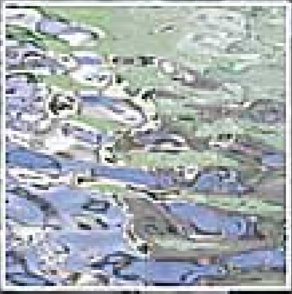
McGraw-Hill **HANDBOOKS**

URBAN WATER SUPPLY HANDBOOK



- Single source of information on wide-ranging topics involved in water infrastructure
- Illustrated with charts, tables, and case studies in the field
- Includes delivery systems information

LARRY W. MAYS, Ph.D, P.E., P.H.

P • A • R • T • 1

HISTORY STRATEGIC PLANNING, AND OUTSOURCING

CHAPTER 1
URBAN WATER INFRASTRUCTURE: A HISTORICAL PERSPECTIVE

Larry W. Mays
Department of Civil and Environmental Engineering
Arizona State University
Tempe, Arizona

Attention to water supply and drainage is the sine qua non for urbanization, and hence for that human condition we call civilization. In fact, development of water supply, waste removal, and drainage made dense settlement possible. Crouch (1993)

Cities are systems within systems of cities.
Berry (1964), given in
Lees and Hohenberg (1988)

1.1 CITIES AND WATER KNOWLEDGE

1.1.1 The Beginning

Humans have spent most of their history as hunting and food gathering beings. Only in the last 9000 to 10,000 years have we discovered how to raise crops and tame animals. Such revolution probably first took place in the hills to the north of present-day Iraq and Syria. From there the agricultural revolution spread to the Nile and Indus valleys. During this time of agricultural revolution, living in permanent villages took the place of a wandering existence. About 6000 to 7000 years ago, farming villages of the Near and Middle East became cities. The first successful efforts to control the flow of water were made in Mesopotamia and Egypt. Remains of these prehistoric irrigation canals still exist.

About 5000 years ago the science of astronomy began and observation of other natural phenomena was leading to knowledge about water resulting in advances for control and use. In the third millennium B.C. time period the Indus civilization had bathrooms in houses and sewers in streets. The Mesopotamians were not far behind (Adams, 1981). In the second millennium B.C. the Minoan civilization on Crete had running water and flushing latrines (Evans, 1964). The Minoan and Mycenaean settlements used cisterns 1000 years before the classical and Hellenistic Greek cities. Water runoff from rooftops was stored in the cisterns which supplied water for the households through the dry summers of the Mediterranean. Between the time of the fall of the Minoan civilization and before the (flowering) growth of the Greek culture (1100–700 B.C.), the Aegean societies were in disarray.

Around 3000 B.C. the first true urban settlements appeared in ancient Mesopotamia, Egypt, and the Indus Valley. These settlements (societies) had elaborate religious, political, and military hierarchies. The areas devoted to the activities of the elite were often highly planned and regular in form, whereas the residential areas often grew by a slow process of accretion resulting in complex and irregular patterns. Greek cities did not follow a single pattern, but grew from old villages. Figure 1.1 shows locations of selected Greek sites, some of which are discussed in this chapter.

In consolidating their empire, the Romans engaged in extensive building of cities. Rome resulted from centuries of irregular growth with particular temple and public districts that were highly planned. The Roman military and colonial towns were laid out in a variation of the grid. As an example, the layout of London, Paris, and many European cities resulted from these Roman origins. Because cities needed a healthy water supply, locations along rivers and streams or underground watercourses were always favored. When cities were small, obtaining clean water and disposing of wastes was not a major problem; however, as cities grew to larger populations and much higher densities there was a much greater need for public infrastructure. Figure 1.2 illustrates the extent of the Roman Empire with selected sites, some of which are discussed in this chapter.

Historically, settlements and communities relied on natural sources to obtain their water. Supplying large quantities such as for fountains (e.g., the Treni Fountain in Rome) was a luxury few communities and states could afford before the Roman era. The most common method of collecting water was saving rainwater in rooftop reservoirs and cisterns. This method was used by the Minoans and Mycenaeans and later by the classical and Hellenistic Greeks, and then the Romans. In fact cisterns are still used throughout the world for storing rainwater for various purposes and as the most common method of providing water at locations without adequate or safe on-site supplies. Before running-water supplies were made possible by conduits and aqueducts, many Roman cities relied upon cisterns and storage tanks. Cisterns ranged from individual use for houses to communal cisterns. Probably the most impressive and immense cistern ever built by the Romans was the Piscina Mirabilis near Pozzuoli in the bay of Naples, Italy.

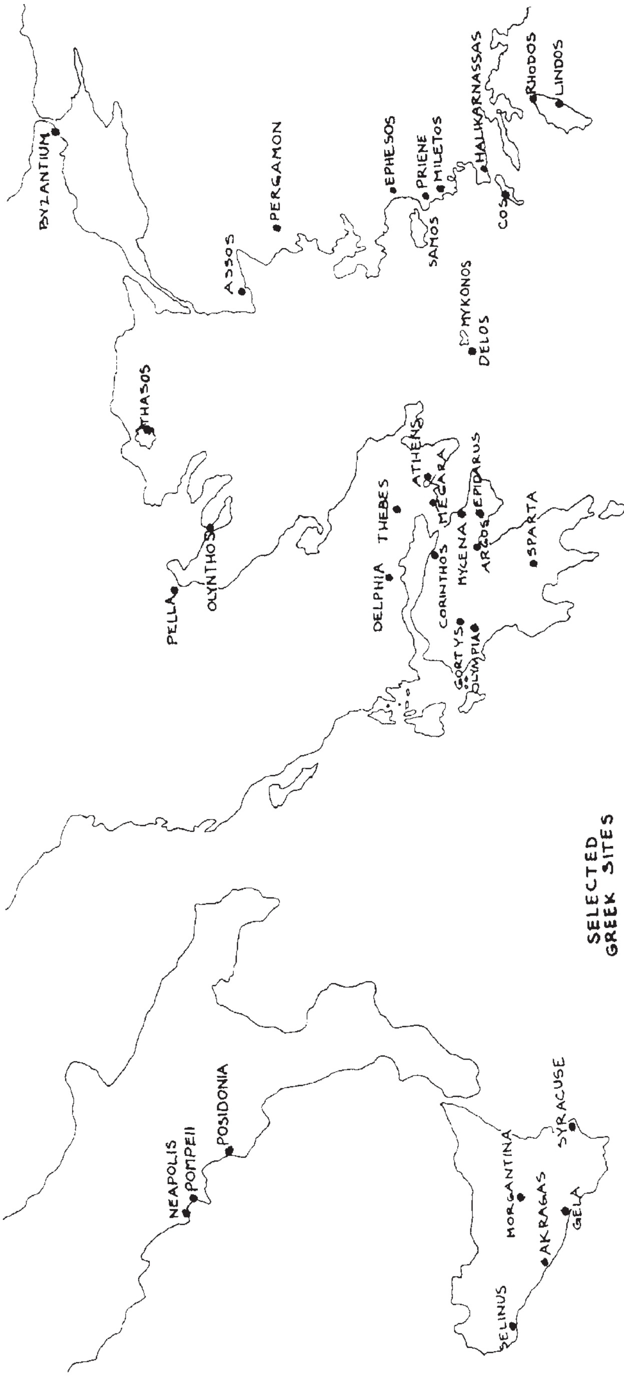


FIGURE 1.1 Map of selected ancient Greek sites. (Crouch, 1993)

FIGURE 1.2 Map of sites in the Roman Empire. (Garraty and Gay, 1972)


B
Pharsalus
Actium
Delphi
Athens
Corinth
Sparta
0 50
Miles
C
GALILEE
Caesarea
Jordan R.
Jabneh
Jerusalem
Dead Sea
0 30 Miles
(RUSSIA)
DACIA
MOESIA
THRACE
Philippi
MACEDON
BLACK SEA
CAUCASUS
CASPIAN SEA
PONTUS
BITHYNIA
Hellespont
AEGEAN SEA
GREECE
Pergamum
ASIA
ASIA MINOR
PHRYGIA
(TURKEY)
Athens
ARMENIA
MESOPOTAMIA
Tigris R.
Euphrates R.
PERSIA
CILICIA
RHODES
CYPRUS
SYRIA
CRETE
SEA
See inset C above
PALESTINE
TRANSJORDAN
Jerusalem
Coastline, 1970 A.D.
NABATAEANS
Alexandria
THE FAIYUM
SINAI
EGYPT

Table 1.1 presents a chronology of water knowledge, from Crouch (1993), who points out that traditional water knowledge relied on geologic and meteorologic observation plus social consensus and administrative organization, particularly in the ancient Greek world.

1.1.2 Contrast of Past and Present

Figure 1.3 shows the functional components of a modern-day water utility, and Fig. 1.4 shows the hierarchical relationship of components, subcomponents, and sub-subcomponents for a modern-day-water distribution system. In comparison Fig. 1.5 shows the functional components of a Roman urban water distribution system based upon the Pompeii system. This chapter attempts to provide an insight into the Greek and Roman era water systems to provide a better understanding of our present-day urban water supply systems. A comparison of the ancient and the modern-day aqueduct systems is shown in Figs. 1.6 and 1.7. Figure 1.6 shows the Central Arizona Project aqueduct through a residential area

TABLE 1.1 Chronology of Water Knowledge

Prehistoric period	Springs
3d–2d millennium B.C.	Cisterns
3d millennium B.C.*	Dams
3d millennium B.C.	Wells
Probably very early (?)	Reuse of excrement as fertilizer
2d millennium B.C.*	Gravity flow supply pipes or channels and drains, pressure pipes (subsequently forgotten)
8th–6th century B.C.	Long-distance water supply lines with tunnels and bridges, as well as intervention in and harnessing of karst water systems
6th century B.C. at latest	Public as well as private bathing facilities, consisting of bathtubs or showers, footbaths, washbasins, latrines or toilets, laundry and dishwashing facilities
6th century B.C. at latest	Utilization of definitely two and probably three qualities of water: potable, subpotable, and nonpotable including irrigation using storm runoff, probably combined with wastewaters
6th–3d centuries B.C.	Pressure pipes and siphon systems

*Indicates an element discovered, probably forgotten, and then rediscovered later.
(?) indicates an educated guess.
Source: Crouch, 1993.

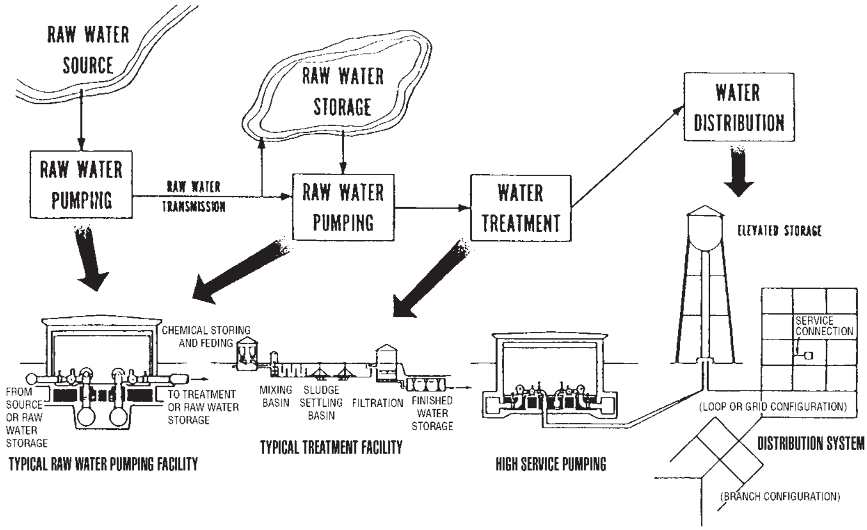


FIGURE 1.3 Functional components of a modern-day water utility. (Cullinane, 1989)

in Scottsdale, Arizona. In contrast Fig. 1.7 shows a Roman aqueduct bridge located in Izmir, Turkey.

1.2 ROMAN PREDECESSORS

1.2.1 The Minoans

Knossos, located approximately 5 kilometers (km) from Herakleion, the modern capital of Crete, was one of the most ancient and most unique cities of the Aegean region and Europe. Knossos was first inhabited shortly after 6000 B.C., and within 3000 years it had become the largest neolithic (circa 5700–28 B.C.) settlement in the Aegean world. During the bronze age (circa 2800–1100 B.C.) the Minoan civilization developed and reached its culmination as the first Greek cultural miracle of the Aegean world. During the neopalatial period (1700–1400 B.C.), Knossos was at the height of its splendor. The city extended over an area of 75,000 to 125,000 square meters (m^2) and had an estimated population in the order of tens of thousands of inhabitants. The water supply system at Knossos was most interesting. An aqueduct supplied water through tubular conduits from the Knunavoi and Archanes regions and branched out into the city and the palace. Figure 1.8 shows the type of pressure conduits used within the palace for water distribution.

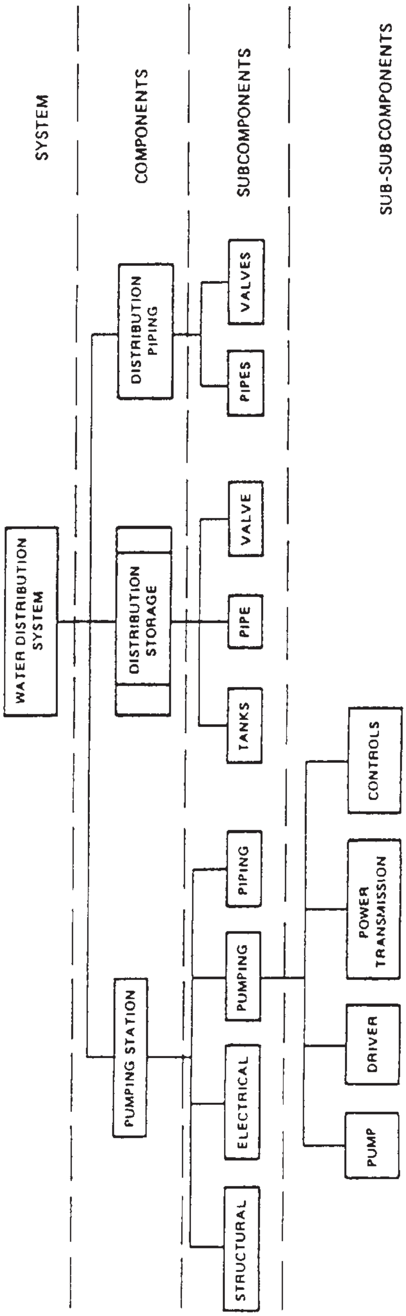


FIGURE 1.4 Hierarchical relationship of components, subcomponents, and sub-subcomponents for a water distribution system. (Cullinane, 1989)

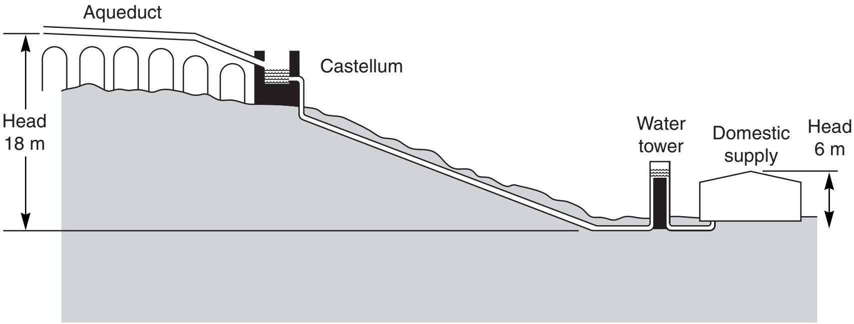


FIGURE 1.5 Functional components of a Roman urban water distribution system based upon the Pompeii system. (Hodge, 1992)

FIGURE 1.6 Central Arizona Project (CAP) aqueduct through a residential area in Scottsdale, Arizona. (Courtesy of Central Arizona Project)

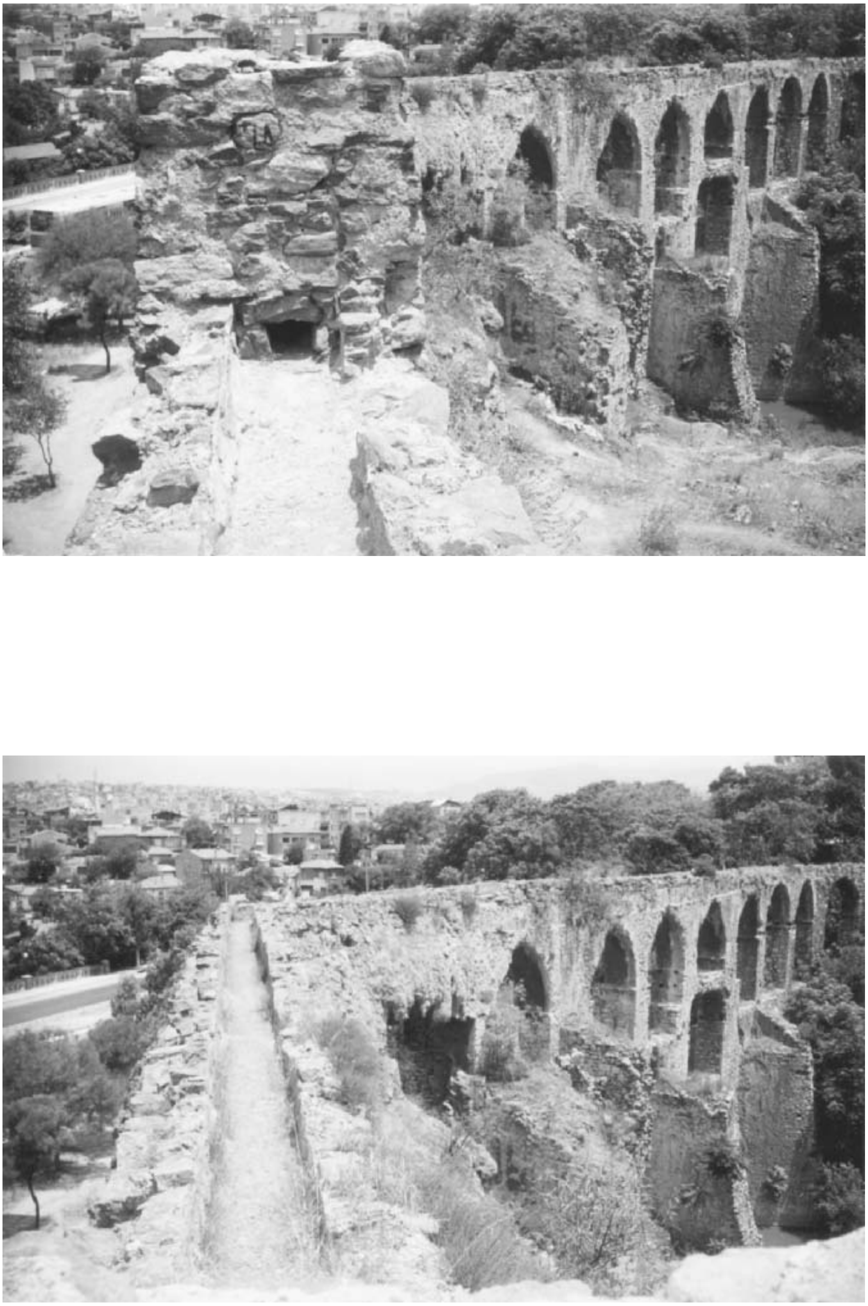


FIGURE 1.7 Roman aqueduct bridge in Izmir, Turkey. (Photo by Koksal B. Celik)

FIGURE 1.8 Water distribution pipe in Knossos, Crete, built around 2000 B.C. by the Minoans. (Photo by Larry W. Mays and copyright by Larry W. Mays)

Unfortunately, around 1450 B.C. the Mycenean palace was destroyed by an earthquake and fire, as were all the palatial cities of Crete.

1.2.2 The Greeks

From the viewpoint of water supply in ancient Greece there are two periods before the Hellenistic period, the archaic period and the classical period, during which time nothing built could compare with the grandiosity of the Roman aqueducts. The city of the archaic and classical Greek periods typically had a spring at its center from which it grew, without any aqueducts, at least in comparison to what the Romans built. Terra-cotta pipelines probably were the usual method of conveying water during the classical Greek period. These terra-cotta pipes [20 to 25 centimeters (cm) in diameter] fit into each other (see Fig. 1.9). Cities were served by fountain(s) in central location(s) receiving water either from a local source or by a conduit made of terra-cotta pipes. Pipes were laid along the bottom of trenches or tunnels, allowing for both protection and access. Two or more pipes in parallel were used depending upon the flow to be conveyed.

During the Hellenistic period the political and economic situation changed, leading to much more architectural development and urban beautification, of which aqueducts played a major role. The progress in science during the

FIGURE 1.9 Terra-cotta pipes found in Ephesus, Turkey. (Photo by Koksal B. Celik)

Hellenistic period provided a new technical expertise. Hellenistic aqueducts usually used pipes, as compared to the Roman masonry conduit. The Hellenistic people did not have the Roman's engineering skill especially in the use of the arch and the building of aqueduct bridges. Greek and Hellenistic aqueducts generally followed the contours of the land, without using any major engineering structures. The one exception was the use of the siphon, which was the method used by the Hellenists to convey water across valleys. Locations of siphons included Ephesus, Methumna, Laodicea (see the parallel siphons in Fig. 1.10), Pergamon, and many others. There are difficulties in dating these siphons, but they may be of the early Roman or Hellenistic period and obviously provided models for the later Roman work. Hellenistic pipelines were built of stone (see Fig. 1.11) or terra-cotta (see Fig. 1.9) whereas the Romans used pipes made of lead (see Fig. 1.12).

Acropolis. The Acropolis at Athens, Greece, had been the focus of settlement starting with the earliest times. Not only its defensive capabilities, but also its water supply, made it the logical location for groups who dominated the region. The Acropolis is located on a rock outcropping and has naturally occurring water. This and the ability of its inhabitants to save rain and spring water resulted in a number of diverse water sources being available, including cisterns, wells, and springs. Figure 1.13 shows the shaft of one of the archaic water holders.

(a)

(b)

FIGURE 1.10 Siphon at Laodicea, Turkey. (*a*) View of the two parallel siphon pipes, and (*b*) closer view of one of the siphons. (Photos by Koksal B. Celik)

FIGURE 1.11 Stone pipes in Ankara, Turkey. (Photo by Koksal B. Celik)

FIGURE 1.12 Lead pipe with marmor joint elements. Pipe is located in the Ephesus Museum. (Photo by Koksal B. Celik)

FIGURE 1.13 Shaft of a water holder at the Acropolis at Athens, Greece. (Photo by Larry W. Mays and copyright by Larry W. Mays)

Syracuse. The karst areas of the Italian peninsula and Sicily are what interested the Greek colonists during the archaic period. An excellent example of this was the founding of Syracuse (on Sicily) as a colony of Corinth in 734 B.C. Among the many things that transferred from the Corinthian culture, such as language, religion, government, and farming, was the water management. As Crouch (1993) points out, "the transfer of knowledge about managing water was facilitated by the similarity of geology and climate between the two sites." During the eighth to first century B.C., the knowledge of locating and collecting water was coupled with the increasing knowledge of transporting both fresh and used water.

Figure 1.14 illustrates the water elements of Syracuse during the Greek times, and Fig. 1.15 shows the later water elements during the Roman times along with further

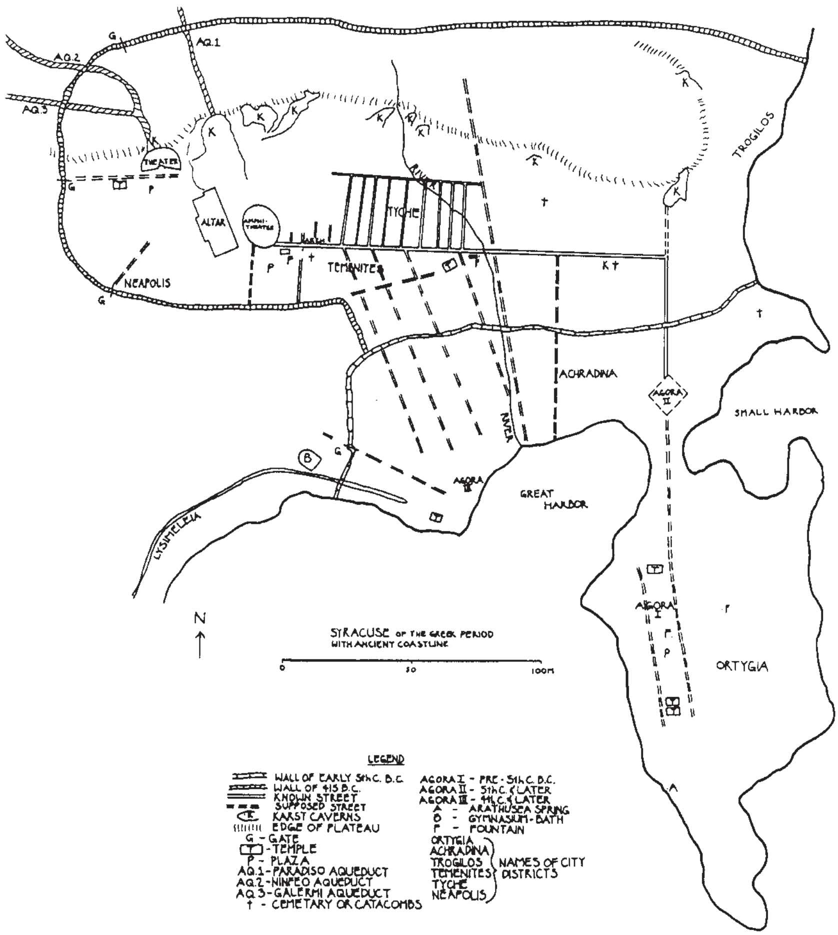


FIGURE 1.14 Map of Syracuse during the Greek times showing the water elements. (Crouch, 1993)

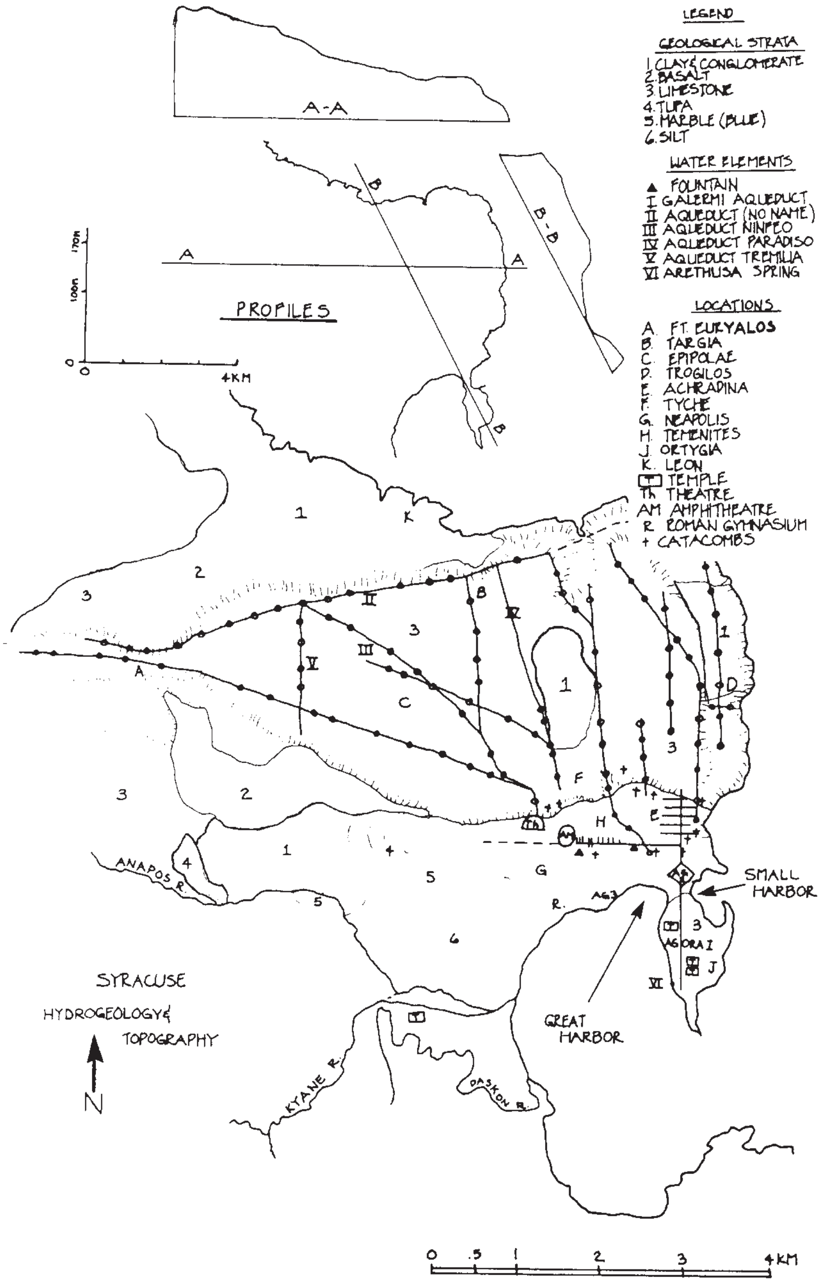


FIGURE 1.15 Map of Syracuse during the Roman times showing the water elements. (Crouch, 1993)

illustration of the geology of the area. The geology of the site, with the earlier and later limestone layers above the clay, created an abundance of water. The Arethusa spring, located at the edge of the sea (refer to Fig. 1.15), was the first settlement on Ortygia (Crouch, 1993). The water supply came from many surface and subsurface openings in the limestone, particularly where the limestone lay above impermeable strata such as marl. The series of grottoes above the Greek theater (see Fig. 1.16) was probably a major factor in the development of Syracuse, because the early Greeks found water flowing here. After a time, possibly a couple of centuries, water found a new path further downhill. Because of this and an increased demand for water (by the increased population), new supplies to this downhill location were developed, using the same outlets. These were the Galermi and Ninfeo aqueducts (routes are shown in Figs. 1.14 and 1.15). Figure 1.16 shows the Greek theater with the grotto formations in the background. Figure 1.17*a* and *b* shows the outlet of these two aqueducts inside the grotto formation, and Fig. 1.18 shows the aqueduct above the grottoes.

The Romans conquered Sicily in the late third century creating an early Greco-Roman society. During the Roman period new water system elements were added to compensate for old elements going out of use and to supply water for bath-gymnasiums and other uses for an increased population (Crouch, 1993). The grottoes and underground tunnels became tomb sites as early as the last century B.C., after centuries of use as water lines. By the second or third century, the water systems began to deteriorate because of little or no maintenance.

FIGURE 1.16 Greek theater in Syracuse showing the grotto formation in the background. Early Greeks found water flowing from the grottoes. (Photo by Larry W. Mays and copyright by Larry W. Mays)

(a)

(b)

FIGURE 1.17 Outlet (of the Galermi and Ninfeo aqueducts) located inside the grotto formation above the Greek theater in Syracuse. (*a*) View looking up at the entrance, and (*b*) floor of the entrance. (Photo by Larry W. Mays and copyright by Larry W. Mays)

FIGURE 1.18 Views of the aqueduct above the grottoes near the Greek theater in Syracuse. (Photo by Larry W. Mays and copyright by Larry W. Mays)

1.2.3 Anatolia

Anatolia, also called Asia Minor, which is part of the present-day Republic of Turkey, has been the crossroads of many civilizations during the last 10,000 years. During the last 4000 years, going back to the Hittite period (2000–200 B.C.), there are many remains of ancient urban water supply systems, including pipes, canals, tunnels, inverted siphons, aqueducts, reservoirs, cisterns, and dams. The majority of the Anatolian waterworks belong to urban water supply systems, in contrast to the large irrigation projects found in Egypt, Mesopotamia, and Indus. Ozis and Harmancioglu (1979) discussed the systems located in Side, Aspendos, Hierapolis, and Ephesus. Ozis (1987, 1996) discussed the history of ancient waterworks in Anatolia starting with the Hittite period (2000–700 B.C.), the Urartu period (900–600 B.C.) in eastern Anatolia, the Ionian to Roman periods (1000 B.C.–A.D. 395) in western and central Anatolia, and continuing through the Byzantine (395–1453), the Seljukian (1071–1308), and the Ottoman (1281–1922) periods. Ephesus was founded during the tenth century B.C. as an Ionian city out of the Artemis Temple. During the sixth century B.C. Ephesus was reestablished at the present site where it further developed during the Roman period.

Baths were unique in ancient cities. One example, the Skolacctica baths in Ephesus, had a salon and central heating. These baths had a hot bath (*caldarium*), a warm bath (*tepidarium*), a cold bath (*frigidarium*), and a dressing room (*apodyterium*). The first building, in the second century, that housed these baths had three floors. A woman named Skolacticia modified the baths in the fourth century making them amiable to hundreds of people. There were public rooms and private rooms, and those who wished could stay for many days. A furnace and a large boiler were used to provide hot water.

Perge, located in Anatolia, is another ancient city that had a unique urban water infrastructure. Figure 1.19*a* illustrates the Majestic Fountain (*nymphaion*), which consisted of a wide basin and a richly decorated architectural facade. Because of the architecture and statues of this fountain, it was one of Perge's most magnificent edifices. A water channel (shown in Fig. 1.19*a* and *b*) ran along the middle, dividing each street and bringing life and coolness to the city. A cover to the underground drainage system for the water channel is shown in Fig. 1.20. The baths of Perge were magnificent and of a rather large size as shown in Fig. 1.21. As in other ancient cities in Anatolia, three separate baths existed (a *caldarium, tepidarium,* and *frigidarium*).

1.3 ROMAN WATER SUPPLY: AQUEDUCTS AND AQUEDUCT BRIDGES

The early Romans devoted much of their time to useful public works projects, building boats, harbor works, aqueducts, temples, forums, town halls, arenas, baths, and sewers. The prosperous early-Roman bourgeois typically had a dozen-room house, with a square hole in the roof to let rain in, and a cistern beneath the roof to store the

(a)

(b)

FIGURE 1.19 Views of the Majestic Fountain and downstream channel for water flow in Perge, Anatolia, Turkey. (*a*) Majestic Fountain (*nymphaion*), and (*b*) water channel dividing street. (Photos by Larry W. Mays and copyright by Larry W. Mays)

FIGURE 1.20 Cover to underground drainage system for water channel from the Majestic Fountain. (Photo by Larry W. Mays and copyright by Larry W. Mays)

water. The Romans built many aqueducts; however, they were not the first to do so. King Sennacherio built aqueducts, as did both the Phoenicians and the Hellenes. The Romans and Hellenes needed extensive aqueduct systems for their fountains, baths, and gardens. They also realized that water transported from springs was better for their health than river water and did not need to be lifted to street level, as did river water. Roman aqueducts were built on elevated structures to provide the needed slope for water flow. Knowledge of pipe making—using bronze, lead, wood, tile, and concrete—was in its infancy, and the difficulty of making pipes was a hindrance. Most Roman piping was made of lead, and even the Romans recognized that water transported by lead pipes was a health hazard.

The water source for a typical water supply system of a Roman city was a spring or a dug well, usually with a bucket elevator to raise the water. If the well water was clear, and of sufficient quantity, it was conveyed to the city by aqueduct. Also, water from several sources was collected in a reservoir, and then conveyed by aqueduct or pressure conduit to a distributing reservoir (*castellum*).

Flow in the Roman aqueducts was obtained by gravity. Water flowed through an enclosed conduit (*specus* or *rivus*), which was typically underground, from the source to a terminus or distribution tank (*castellum*). Above ground aqueducts were built on a raised embankment (*substructio*) or on an arcade or bridge. Settling tanks (*piscinae*) were located along the aqueducts to remove sediments and foreign matter. Subsidiary lines (*vamus*) were built at some locations along the aqueduct to supply additional water. Also subsidiary or branch lines (*ramus*) were used. At distribution points water

FIGURE 1.21 Views of the Roman baths at Perge, Anatolia, Turkey. (Photos by Larry W. Mays and copyright by Larry W. Mays)

was delivered through pipes (*fistulae*) made of either tile or lead. These pipes were connected to the *castellum* by a fitting or nozzle (*calix*). These pipes were usually placed below the ground level along major streets. See Garbrecht (1982), Evans (1994), Robbins (1946), and Van Deman (1934) for additional reading on the water supply of the City of Rome and of other locations in the Roman Empire.

To properly discuss Roman water supply we must at least be aware of the treatises of Vitruvius (*De Architectura*) (Morgan, 1914) and Sextus Julius Frontinus (*De aqueaductu urbis Romae*) (translation 1973). Vitruvius (84 B.C.) and Frontinus (A.D. 40–103) did not contribute to the scientific development of hydraulics; however, the treatises that they authored do give us insight to the planning, construction, operation, and management of Roman hydraulic structures. The Greeks gave us the great achievements in science, and the Romans gave us the great achievements in the improvement of technology.

1.3.1 Vitruvius and Frontinus

Vitruvius discussed the various elements of water supply in his book VIII, which also is an interesting source for information on springs, the uses and quality of water, and some of the techniques involved. In chapters 5 and 6 of book VIII, Vitruvius addressed the quality of water is some of his passages.

Chapter 5:

20. Some springs appear to be mixed with wine; as that in Paphlagonia, which, when taken, inebriate as wine.
21. In Arcadia, at the well-known city of Clitorium, is a cave flowing with water, of which those who drink become abstemious.
22. There is also in the island of Chios, a fountain, of which those who imprudently drink become foolish.
23. At Susa, the capital of Persia, there is a fountain, at which those who drink lose their teeth.
24. The quality of the water, in some places, is such that it gives the people of the country an excellent voice for singing, as at Tarsus, Magnesia, and other countries.

Chapter 6:

1. Water is conducted in three ways, either in streams by means of channels built to convey it, in leaden pipes or in earthen tubes, according to the following rules.…
10. Water conducted through earthen pipes is more wholesome than that through lead; indeed that conveyed in lead must be injurious, because from it white lead is obtained, and this is said to be injurious to the human system.

Frontinus was a retired army officer who, in A.D. 97 took over as director of the Rome Metropolitan Waterworks. He declared the Roman aqueducts as the real mark of civilized living. Rome was declared as the first civilization to set a proper priority on decent sanitation and abundant drinking water. Actually drinking water was a by-product of the aqueducts with the real purpose being to supply baths (Hauck, 1988).

1.3.2 Aqueducts of Rome

Rome's aqueduct system evolved over a 500-year time period. As did most of our modern-day urban water distribution systems, Rome's system evolved in a piecemeal fashion. Frontinus's treatise *De aquaeductu urbis Romae* (On the water supply of the city of Rome), described Rome's water supply in the form of a notebook and was written for himself, or for a possible successor, to serve as a rule or guide (*formulas administrationis*). This treatise is not a comprehensive discussion of the aqueduct system, but does give us a picture of various aspects of Rome's aqueduct system.

The major sources and routes of the aqueducts to ancient Rome are illustrated in Fig. 1.22. In 312 B.C. the Adile Appius Claudius constructed the first aqueduct, Aqua Appia, a simple underground channel. The next aqueduct was Aqua Anio Vetus, constructed in 272 B.C. Eventually there were 11 aqueducts that supplied water to Rome. Table 1.2 lists the aqueducts in Rome along with other information such as the date the aqueduct was built, its length, and its origin. Frontinus discussed nine of the aqueducts; the two built after his treatise was written are the Traiana and Alexandrina. The two Anio aqueducts received water directly from the river, and the Alsietina aqueduct received water directly from Lake Alsietina. The remaining aqueducts (Appia, Marcia, Tepula, Julia, Virgo, Claudia, Traiana, and Alexandrina) received water from springs. Several feeder branches were commonly necessary to collect enough water for the aqueducts. For example, the Virgo and Marcia aqueducts used feeder branches to collect water in a collecting basin where the water entered the main channel.

The location and routes of the aqueducts in ancient Rome are illustrated in Fig. 1.23*a,* with more details shown in Fig. 1.23*b.* The area map shown in Fig. 1.24 is the location (area of Spes Vetus near the two major roads, Via Labicana and Via Praenestina) where all of the eastern aqueducts entered the city. Figure 1.25*a* to *e* illustrates various views taken near and of the Porta Maggiore (the double-arched gate which carries the Aqua Claudia and Aqua Novus), with the views (directions pointed to by the camera) indicated in Fig. 1.24. Figure 1.25*a* and *b* shows the aqueducts Claudia (above) and Anio Novus (lower) on top of the Porta Maggiore. The gate is reminiscent of a triumphal arch. Figure 1.25*c* points to the location of where the two aqueducts changed direction. Figure 1.25*d* shows the three aqueducts [Julia (top), Tepula (center), and Marcia (lower)] located in the Aurelian Wall. Figure 1.25*e* shows the branch aqueduct, the Aqua Claudia, that supplied water to the Trophies of Marius *nymphaeum*.

As pointed out by Evans (1994), throughout the history of Rome, aqueduct construction was generally not planned in an orderly manner. During Republican Rome the city fathers tended to allow needs to become critical before aqueducts were built, similar to modern-day practice. Available funds for the construction were also needed as for Anio Vetus and Aqua Marcia as pointed out by Frontinus (6.1, 7.4). Rome's natural supply from abundant springs, wells

(Frontinus, 4.1) supplied by a high water table, household cisterns, and water from the Tiber were probably adequate for several centuries, so the first aqueduct to bring water into the city from outside, Appia, was not completed until 312 B.C. Some of the springs of Rome were said by Frontinus (4.2) to have curative powers, "The memory of the springs is still considered holy and revered; indeed they are believed to retore sick bodies to health, such as the spring of the Camenae, and…that of Juturna." (See Evans, 1994.)

Table 1.3 lists the number of *castella* for each of the aqueducts and the volume in *quinariae* according to Frontinus's statistics (Evans, 1994). Unfortunately, we do not know the definition of the *quinaria.* As pointed out by Evans (1994), Frontinus[1] *quinaria* cannot be converted into modern units of measurement: "Frontinus regarded it as an accepted unit internally consistent and applicable to the statistics he reports in *De aquaeductu* (34.2–3)." The *quinaria* can only be used to compare relative capacities and deliveries within Rome's water system as a whole (Bruun, 1991). Table 1.4 provides standardized measures of pipes, providing some insight to the *quinaria.*

The total distribution is 14,018 *quinariae,* with 4063 *quinariae* outside the city and 9955 1/2 *quinariae* within the city through 247 *castella.* Of the 4063 *quinariae* distributed outside the city, 1718 *quinariae* were in Caesar's name and 2345 *quinariae* in private customer names. The 9955 1/2 *quinariae* distributed within the city had the following purposes according to Frontinus (Evans, 1994):

In Caesar's name	1707 1/2
To private consumers	3847
For public functions	4401
To 18 camps	279
To 95 public works	2301
To 39 fountains	386
To 591 basins	1335

Frontinus (18) discussed elevations of the aqueducts:

> All the aqueducts arrive in the city at different elevations. As a result, certain ones serve higher places, and others cannot be raised to more lofty areas; indeed even the hills have grown up little from rubble on account of the great number of fires. The height of five aqueducts permits them to be raised into every part of the city, but of these, some are forced by greater pressure, others by less. The highest of all is the Anio Novus, the next highest, the Claudia; the Julia holds third place, the fourth the Tepula, and after this the Marcia, which even equals the height of the Claudia at its source. But the old aqueduct builders constructed their lines at lower elevation, either because the fine points of the leveling art had yet been ascertained or because they deliberately made it their practice to bury aqueducts underground to prevent them from being cut easily by enemies, since a good many wars were still being fought against the Italians.

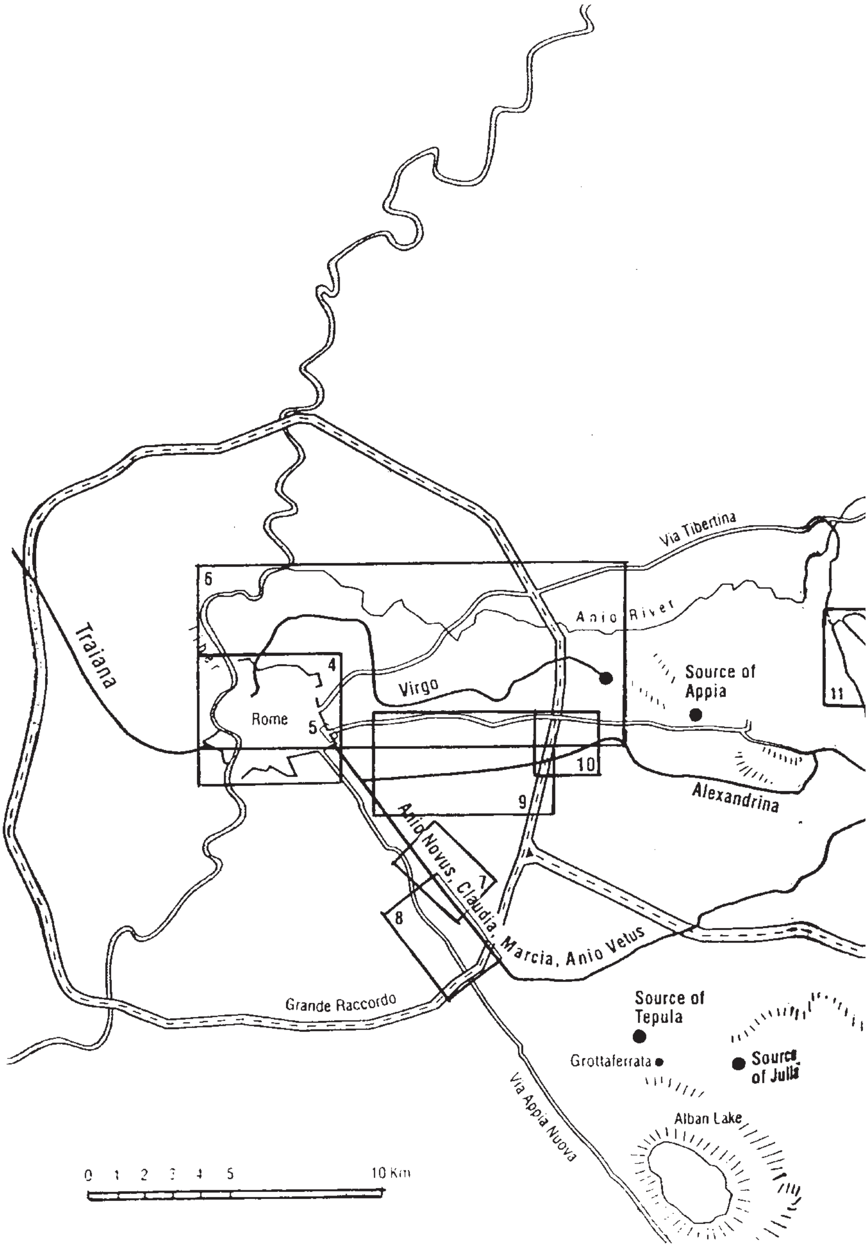
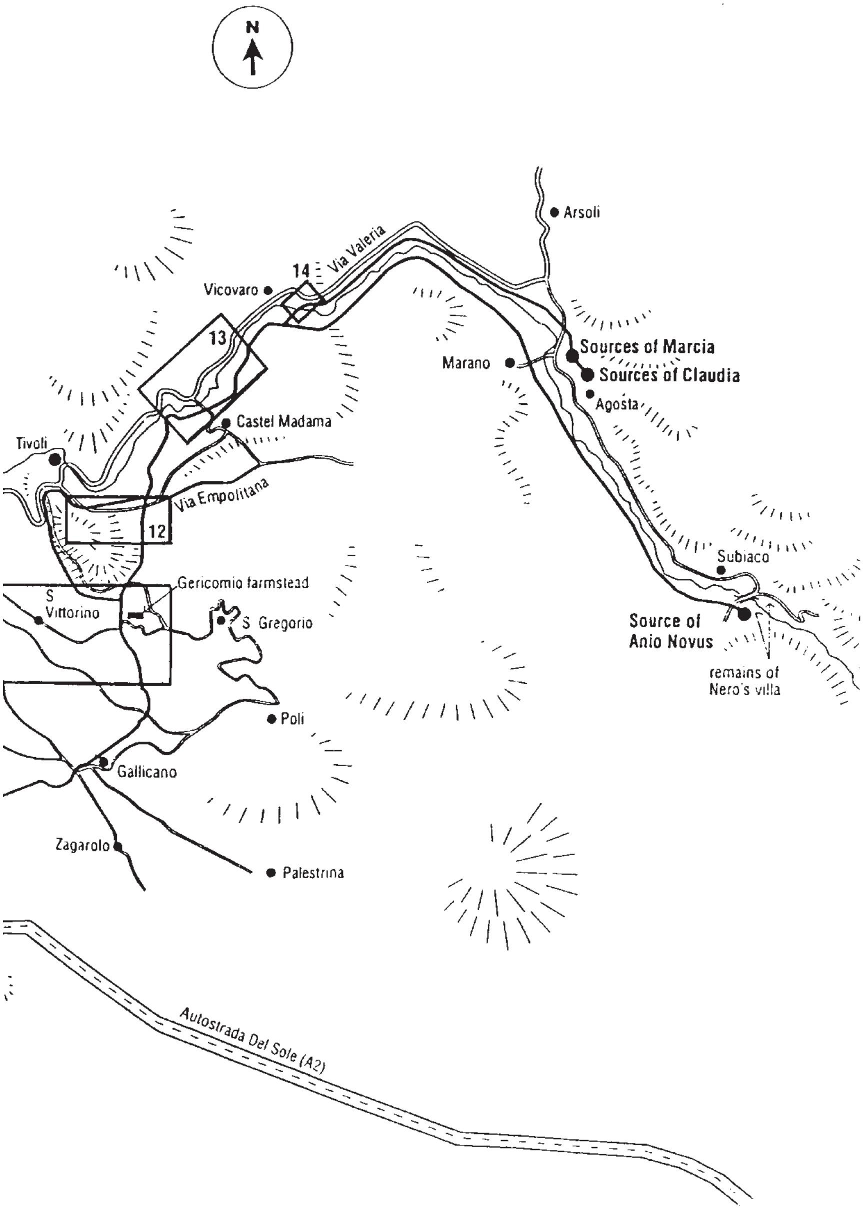


FIGURE 1.22 Map of aqueducts to Rome. (Aicher, 1995)



Overview of Maps

TABLE 1.2 Rome's Aqueducts According to N. Smith (Fahlbusch, 1987)

Name	Date	Length, km				Cross-section, m^2	Height, mMSL	Origin
		Total	Below surface	Above surface	Bridge			
Appia	312 B.C.	17.6	16.8	0.8	0.09	0.69 × 1.68	20	Springs in Anio Valley
Anio Vetus	272	64	63.6	0.4	—	0.91 × 2.29	48	Anio River
Marcia	144–140	91.2	80	0.8	10.4	1.52 × 2.59	59	Springs in Anio Valley
Tepula	126	18.4	8.4	0.8	9.2	0.76 × 1.07	61	Springs near Alban Mountains
Julia	33	22.8	12.4	0.8	9.6	0.61 × 1.52	65	Springs near Alban Mountains
Virgo	19–21	20.8	19.2	0.4	1.2	0.61 × 1.75	20	Springs in Anio Valley
Alsietina	10–2	32.8	32.4	—	0.4	1.75 × 2.59	17	Lake Alsietinus
Claudia	38–52 A.D.	68.8	53.6	1.2	14	0.91 × 1.98	67	Springs in Anio Valley
Anio Novus	38–52	86.4	72.8	2.4	11.2	1.22 × 2.74	70	Anio River
Traiana	109–117	59.2	59.2	—	—	1.30 × 2.29	73	Springs near Lake Sabatina
Alexandrina	226	22.4	12.8	7.2	2.4	—	—	Springs at Sasso Bello

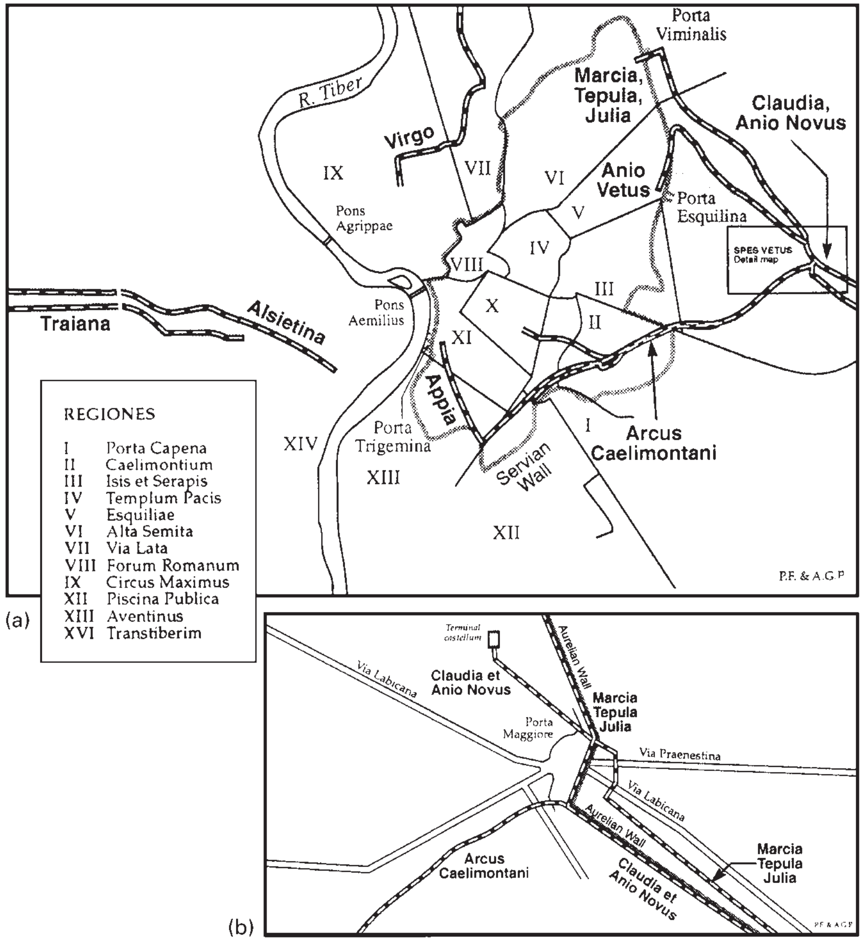


FIGURE 1.23 Aqueducts in ancient Rome. (*a*) Termini of the major aqueducts (Evans, 1994), and (*b*) the area of the Spes Vetus showing the courses of the major aqueducts entering the city above ground. (From R. Lanciani, *Forma Urbis Romae,* as presented in Evans, 1994)

1.3.3 Aqueduct of Nîmes (Ancient Nemausus) and the Pont du Gard

The aqueduct of Nemausus (built circa 20 B.C. by Agrippa) conveyed water approximately 50 km from the Fontaine d'Eure at Uzes to the *castellum* in Nîmes (see the location of the route in Fig. 1.26). The most spectacular feature of this aqueduct is the Pont du Gard (Bridge of the Gard) discussed at the end of this section. Of almost equal importance from an engineering viewpoint, however, is the not-so-obvious gradient of the aqueduct from Uzes to Nîmes and the high quality of surveying that would have been required to maintain the gradient during con-

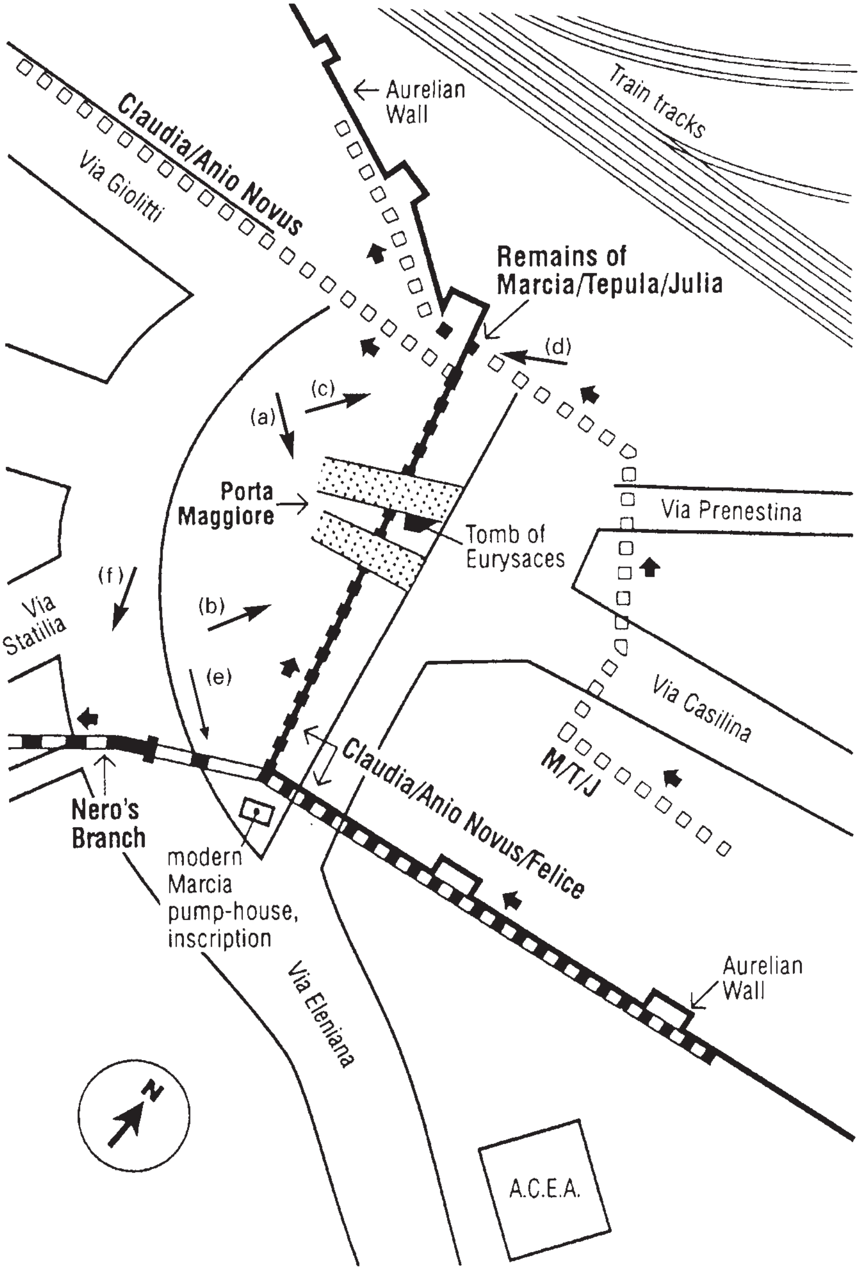


FIGURE 1.24 Map of the Porta Maggiore area showing the direction of the photographs in Fig. 1.25*a* to *e*. (Modified from Aicher, 1995)

(a)

(b)

FIGURE 1.25 Views (direction shown in Fig. 1.24) of the aqueducts of Rome at the Porta Maggiore (double-arched gate) on the Aurelian Wall where all the eastern aqueducts entered Rome. (*a*) View of the Porta Maggiore, (*b*) aqueducts Claudia (top) and the Anio Novus (bottom) above the Porta Maggiore, (*c*) view of the Aurelian Wall, (*d*) three aqueducts Julia (top), Tepula (center), and Marcia (lower) located on the Aurelian Wall, and (*e and f*) Nero's branch of the Aqua Claudia that supplied water to the Trophies of Marius *nymphaeum*. (Photos by Larry W. Mays and copyright by Larry W. Mays)

(c)

(d)

FIGURE 1.25 (Continued)

(e)

(f)

FIGURE 1.25 (Continued)

TABLE 1.3 Frontinus's Statistics on the Aqueducts of Rome

Aqueduct	Number of *castella*	Volume, *quinariae*
Appia	20	699
Anio Vetus	35	1,508 1/2
Marcia	51	1,472
Tepula	14	331
Julia	17	597
Virgo	18	2,304
Claudia/Anio Novus	92	3,498
Total	247	10,409 1/2

*Frontinus (78.3) states a total of 9,955 *quinariae,* but the sum of individual deliveries by aqueduct in Frontinus (79–86) results in a total of 10,409 1/2 *quinariae.*

struction. Elevation of the source at Uzes was 76 m and of the castellum at Nîmes was 59 m, for a difference of 17 m. Even if the aqueduct had been constructed on a straight line between the two points, the distance was 20 km making the slope or overall gradient 0.85 m/km, 0.00085 m/m, or 0.085 percent.

The actual length of the aqueduct constructed was around 50 km, and the aqueduct was built with the profile shown in Fig. 1.27 showing the variation of slopes. Hauck (1988) states: "The responsibilities of the chief *aquilex* consisted of two unequal legs of the aqueduct. One was the northern portion, upstream from the river, a good ten *milia passus* long, with a normal gradient of four-fifths of an *uncia* per one-hundred *pedes* (0.00067)." The *aquilex* is a person employed to find water. *Milia* (or *mille* for singular) refers to a thousand, and *passus* refers to approximately 1.48 m, so that a *mille passus* is 1478.5 m or 1 mile and 10 *milia passus* refers to 10 miles. A *pedes* is 1 foot (ft), so 1 *passus* is 5 *pedes. Unicia* refers to a twelfth or 1 inch (in).

Figure 1.28 shows a portion of the aqueduct several hundred meters southwest (downstream) of the Pont du Gard. Figure 1.29 shows the aqueduct tunnel immediately on the upstream (in the direction of the aqueduct) side of the Pont du Gard. Figure 1.30 shows portions of the aqueduct to the northeast of the upstream side of the Pont du Gard.

The Pont du Gard, shown in Fig. 1.31 is one of the more spectacular aqueduct bridges ever built and is the most photographed aqueduct in existence.

1.3.4 Aqueduct of Segovia

The Segovia aqueduct received water from the Rio Acebeda (also referred to as the Frio River on La Acebeda), a small river approximately 12 km south of the

TABLE 1.4 Standardized Measures of Roman Pipes According to Frontinus (Fahlbusch, 1987)

Name of pipe	Inner diameter		Circumference		Area of cross-section		
	Digiti	cm	*Digiti*	cm	*Digiti*	*Quinariae*	cm^2
5 quinaria	5/4*	2.31	3.93	7.27	1.23	1.00	4.20
6 senaria	6/4*	2.78	4.72	8.72	1.77	1.44	6.05
7 septenaria	7/4*	3.24	5.50	10.18	2.41	1.96	8.22
8 octonaria	8/4*	3.70	6.29	11.63	3.14	2.56	10.75
10 denaria	10/4*	4.63	7.86	14.54	4.71	4.00	16.80
12 duodenaria	12/4*	5.55	9.43	17.44	7.07	5.76	24.19
15 quinum denum	15/4*	6.94	11.79	21.80	11.04	9.00	37.80
20 vicenaria	20/4*	9.25	15.72	29.07	19.63	16.00	67.20
20 vicenaria	5.05	9.34	15.85	29.32	20*	16.26	68.45
25 vicenum quinum	5.64	10.44	17.73	32.80	25*	20.37	85.56
30 tricenaria	6.18	11.44	19.42	35.92	30*	24.43	102.62
35 tricenum quinum	6.67	12.35	20.98	38.81	35*	28.51	119.74
40 quadragenaria	7.14	13.20	22.42	41.47	40*	32.58	136.85
45 quadragenum quinum	7.57	14.00	23.79	44.00	45*	36.65	153.94
50 quinquagenaria	7.99	14.76	25.07	46.39	50*	40.73	171.05
55 quinquagenum quinum	8.37	15.48	26.29	48.64	55*	44.80	188.16
60 sexagenaria	8.74	16.17	27.46	50.80	60*	48.87	205.26
65 sexagenum quinum	9.09	16.82	28.58	52.88	65*	52.94	222.37
70 septuagenaria	9.44	17.46	29.67	54.88	70*	57.02	239.47
75 septuagenum quinum	9.77	18.08	30.71	56.81	75*	61.09	256.58
80 octogenaria	10.09	18.67	31.71	58.65	80*	65.17	273.70
85 octogenum quinum	10.40	19.24	32.69	60.47	85*	69.24	290.79
90 nonagenaria	10.70	19.80	33.64	62.23	80*	73.31	307.90
95 nonagenum quinum	11.00	20.34	34.56	63.93	95*	77.38	325.01
100 centenaria	11.28	20.87	35.46	65.60	100*	81.45	342.10
120 centenum vicenum	12.36	22.86	38.83	71.84	120*	97.75	410.55

*Source of name.

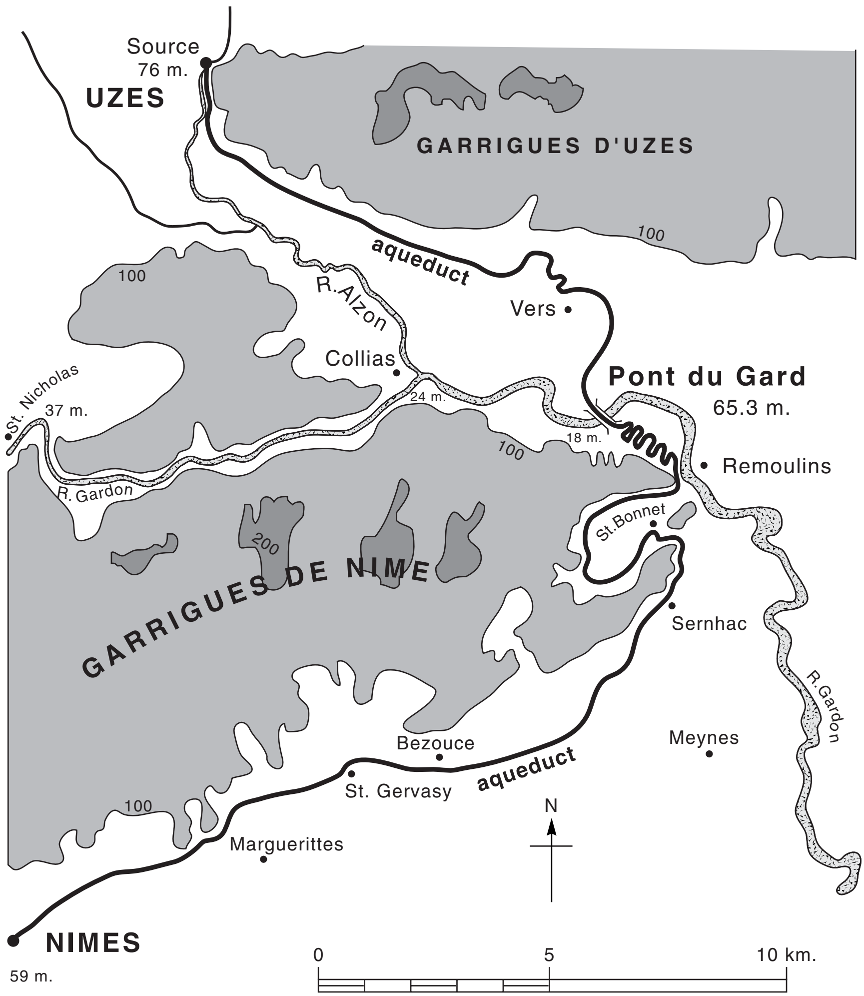


FIGURE 1.26 Route of the aqueduct of Nîmes. (Hodge, 1992)

city. Water was diverted into the aqueduct by a weir set at an angle across the river. Figure 1.32 shows the Roman aqueduct bridge in Segovia with the arches at two levels. The aqueduct makes a 90 degree turn in the Plaza of Diax Sanz. This aqueduct, built during the second half of the first century or early years of the second century, has a maximum height of 28.9 m in the Plaza of Azoguejo. This masterpiece of engineering consisted of around 20,400 stone blocks that were not held together by mortar or cement, making it very unique from an engineering perspective.

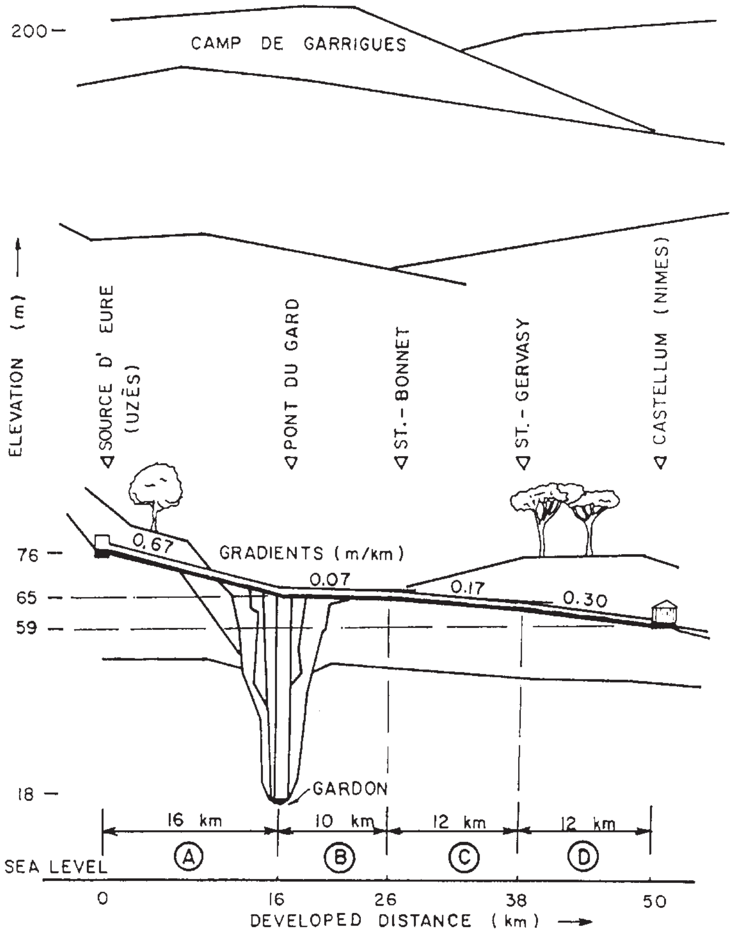


FIGURE 1.27 Profile of the aqueduct of Nîmes. (Hauck, 1988)

1.3.5 Aqueducts of Ephesus

The water supply system to Ephesus consisted of four systems (see Fig. 1.33): (1) the Sirince system from the east, (2) the Derbentdere system from the southeast, (3) the Degirmendere system from the southwest, and (4) the Kayapinar system from the northeast. Ephesus (the metropolitan city of Roman Asia Minor) received its first water by aqueduct (built by C. Sextius Pollio) about the same time as Nîmes did.

The Sirince system conveyed groundwater from the hills of the village Sirince, east of Selcuk. This system probably supplied the Artemis Temple in Ephesus (Ozis, 1996). A collection system, consisting of a main and three lateral galleries

FIGURE 1.28 Aqueduct of Nîmes to the southeast of the Pont du Gard. (Photo by Larry W. Mays and copyright by Larry W. Mays)

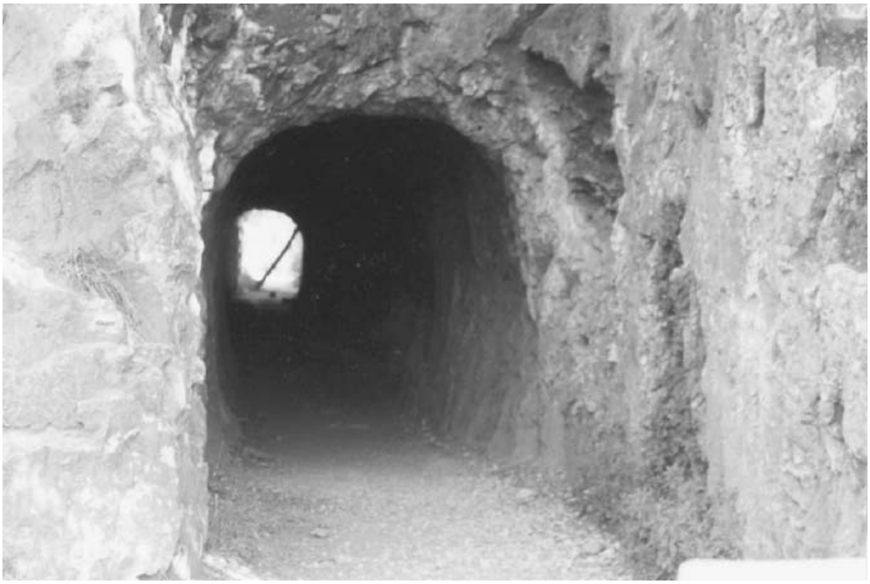
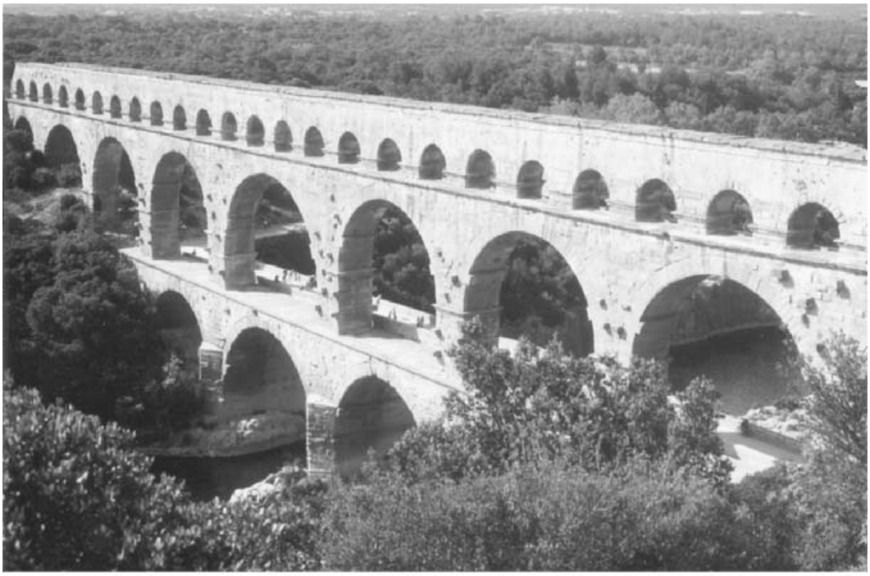


FIGURE 1.29 Aqueduct tunnel adjacent to the Pont du Gard. (Photo by Larry W. Mays and copyright by Larry W. Mays)

FIGURE 1.30 Aqueduct of Nîmes to the northeast of the Pont du Gard. (Photo by Larry W. Mays and copyright by Larry W. Mays)



(a)

(b)

(c)

FIGURE 1.31 Views of the Pont du Gard aqueduct bridge. (*a*) Pont du Gard bridge showing the three levels, (*b*) view of the bridge from the river, (*c*) view of the arches, and (*d*) view of the top of the aqueduct bridge. (Photo by Larry W. Mays and copyright by Larry W. Mays)

(d)

FIGURE 1.31 (Continued)

(3.5 m by 0.45 m), was used to bring water to the 8-km-long conveyance system which consisted of baked clay pipes of 12- to 22-cm outer diameter and 10- to 16-cm inner diameter (Ozis, 1996). Figure 1.34 shows an aqueduct bridge used to support the pipes on the system from Pirango.

Water entered the foundation of the Artemis Temple by seven lead pipes (still remaining in situ), each having an inner diameter of 8 cm, a wall thickness of 4.5 cm, and a length of 60 cm. Joints were set up by marmor joint elements (shown in Fig. 1.12) of 13-cm inner diameter, and 35-cm outer diameter and length. Figure 1.12 illustrates a lead pipe with marmor joint element's displayed in the Ephesus museum in Selcuk.

The Derbentdere system (also called Marnas) from the southeast was 6 km long and consisted of three parallel lines of baked clay pipes of different diameters, laid partly on rock-cut terraces. Ozis (1996) discusses these systems in more detail and provides many references for further reading. Figure 1.35 shows the 16-m-high Sextilius-Pollio aqueduct bridge on the Marnas conveyance system, built around A.D. 4 to 14.

The Degirmendere system (also called Kenchrios) from the southwest of Ephesus was a 43-km-long Roman aqueduct system consisting of 15 aqueduct bridges crossing the valleys. This system dates from the first century, conveyed water from the Degirmendere Springs (east of Kusadasi) at a rate of 60 liters per

FIGURE 1.32 Two views of the Roman aqueduct bridge in Segovia, Spain. The aqueduct was built on two levels without the use of any mortar or cement. (Photo by Larry W. Mays and copyright by Larry W. Mays)

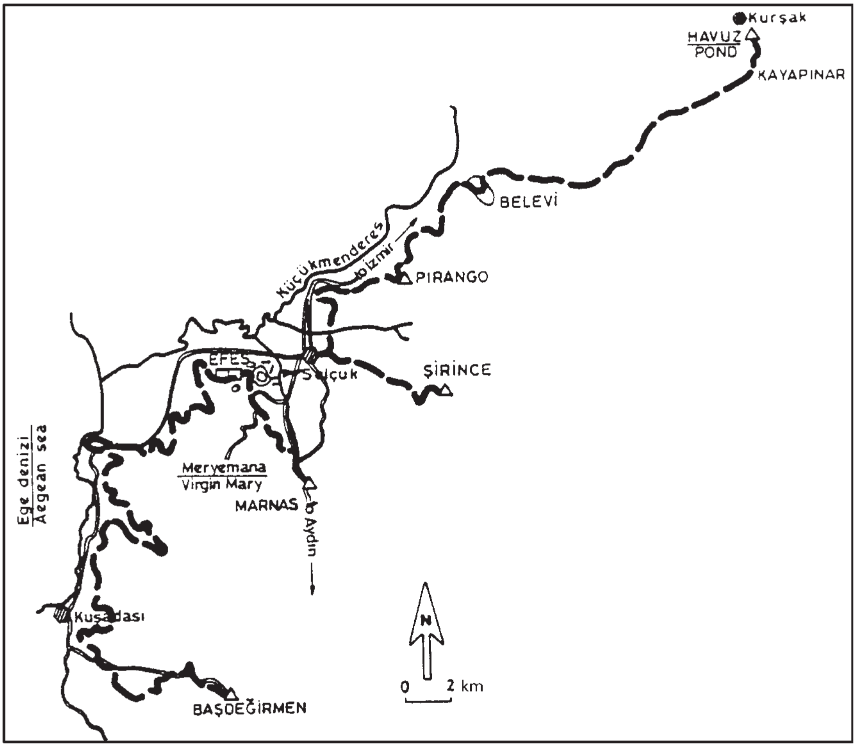


FIGURE 1.33 Locations of the aqueduct system to Ephesus. (Ozis, 1996)

second (L/s) and the Keltape Springs (north of Kusadasi) at a rate of 18 L/s (Ozis, 1996). This Roman water conveyance system was 43 km long and crossed the valley by means of 15 aqueducts. Figure 1.36 shows a portion of the conveyance through Kusadasi.

The highest aqueduct was Bachcecik, which was 20 m high, and the longest aqueduct was the Arvalya, which was 400 m long. This aqueduct system also consisted of tunnel sections, the major one being between the Baskemer and Bahcecik aqueducts, and was constructed by the kanat technique. The other two systems are described in more detail by Ozis (1996).

Water was distributed throughout Ephesus by a dense network of baked clay pipes (see Fig. 1.37). A large number of these still remain in their original positions. A sewerage network under the main streets of Ephesus discharges to the sea (the present-day port is 5 km from the sea). Figure 1.38 shows the street near the Celsus Library with the sewage network under the street (marmor road—the main road leading to the port).

FIGURE 1.34 Aqueduct bridge of the Sirince system from Pirango to Ephesus. (Photo by Koksal B. Celik)

1.3.6 Siphons

Siphons were briefly mentioned in Sec. 1.2.2, and the parallel siphons near Laodicea in Turkey were shown in Fig. 1.10. Siphons were built throughout the Roman Empire, being numerous in Gaul and relatively rare in Rome (Hodge, 1992). Many more were built in Roman times than in Greek times. The best-preserved and largest examples of siphons are found in Lyon. Most siphons were V- or U-shaped for valley crossings. In some situations two or three siphons were used in series. An example of a double siphon is at Les Tourillons on the Craponne, Lyon. A triple siphon was built at Aspendos with two open tanks separating them at the top of the two (north and south) towers. Figure 1.39 shows a general view of the wide area without a true valley, with the south tower in the foreground and the north tower in the distance.

1.4 ROMAN WATER SUPPLY: URBAN DISTRIBUTION SYSTEMS

A diagram of a simple Roman urban distribution system (as based on the Pompeii system) is shown in Fig. 1.40. The main aqueduct ends at the main *castellum,* or *castellum divisorium.* The *castellum divisorium* is a junction where the main aqueduct ends and the urban distribution system begins. A lead pipe or smaller aqueduct

FIGURE 1.35 Bridge (16 m high) of the aqueduct Sextilius-Pollio of the Derbentdere system to Ephesus. This is part of the Marnas conveyance system to Ephesus built around A.D. 4–14. (Photo by Koksal B. Celik)

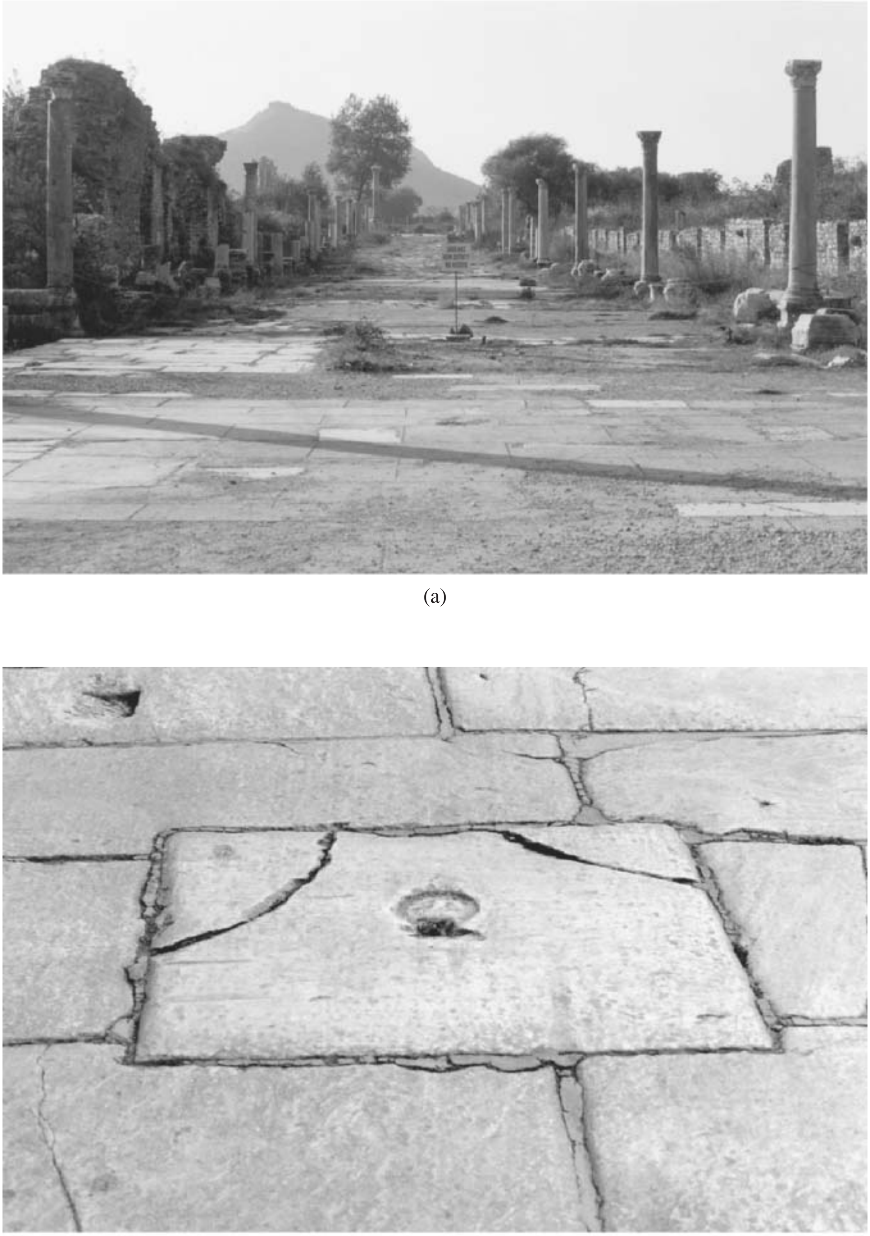
FIGURE 1.36 Aqueduct of the Degirmendere system to Ephesus located in Kusadasi, Turkey. (Photo by Koksal B. Celik)

FIGURE 1.37 Typical baked clay pipe found in Ephesus. (Photo by Larry W. Mays and copyright by Larry W. Mays)

was then used to transport the water from the main *castellum* to a secondary *castellum* or water tower when the secondary *castellum* was raised to the top of a brick pier. From the water tower (secondary *castellum*) lead pipes were used to branch the supply to individual customers and to public fountains for the domestic supply.

1.4.1 Pompeii

The Greco-Roman city of Pompeii is located on the bay of Naples, south-southeast of Mt. Vesuvius in Italy. Pompeii is one of the most significant proofs of the magnificence of Roman civilization that was originally founded by Greek colonists probably around the ninth to eighth century B.C. The city was influenced by others including the Etruscans for almost 50 years (until 474 B.C.), after which it came back under Greek rule. During the fifth century it came part of the Samnite area of expansion and saw tremendous growth. Pompeii became under the influence of Rome after three long and bitter wars around 290 B.C. In 80 B.C. it became a Roman colony with the name of *Colonia Cornelia Veneria Pompeii.* Pompeii had a flourishing economy and widespread affluence after it became a colony but later experienced a devastating earthquake in about 62 A.D. A few years later on August 24, A.D. 79, Mt. Vesuvius erupted and destroyed the city.



(a)

(b)

FIGURE 1.38 Ephesus sewage network drained to the port under the street. (*a*) Main street leading to the Celsus Library in Ephesus, and (*b* inlet to underground sewage network. (Photo by Larry W. Mays and copyright by Larry W. Mays)

FIGURE 1.39 Remnants of triple siphon at Aspendos, Turkey. The wide valley is illustrated with the south tower in the foreground and the north tower in the distance. Open tanks were located at the top of each of the two towers, commonly referred to as pressure towers in the literature. (Photo by Koksal B. Celik)

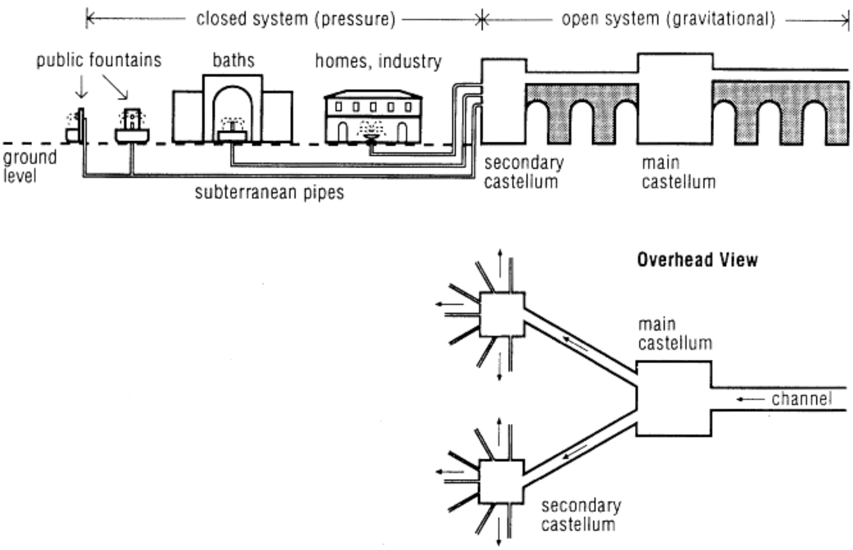


FIGURE 1.40 Typical Roman urban water distribution system.

Sources of water for Pompeii included wells, cisterns, and other reservoirs, and a long-distance water supply line (Crouch, 1993, p. 178). According to Richardson (1988, p. 51) there were no springs within the city of Pompeii. The water table was tapped within Pompeii using wells as deep as 38 m below the surface (Maiuri, 1931, pp. 546–557). A long-distance aqueduct from the hills to the east and northeast also supplied the city. This aqueduct received water from springs at Serino, near Avellino, and then was routed via Sarno around the north side of Mt. Vesuvius to serve Naples and two large cisterns of Cento Camerelle (Baiae) and the Piscina Mirabilis (Misenum). From Sarno a branch aqueduct was routed to Pompeii terminating at the *castellum* at Porta Vesuvii (Hodge, 1992). Figure 1.41 illustrates the water distribution system elements of Pompeii (circa A.D. 79).

The households and public buildings both had very interesting systems to collect and store rainwater. Buildings with peaked roofs had gutters along the eaves to collect the rainwater and downspouts to carry the water to the cisterns located under the buildings. Downspouts were made of terra-cotta pipes and were often set inside the wall (see Fig. 1.42).

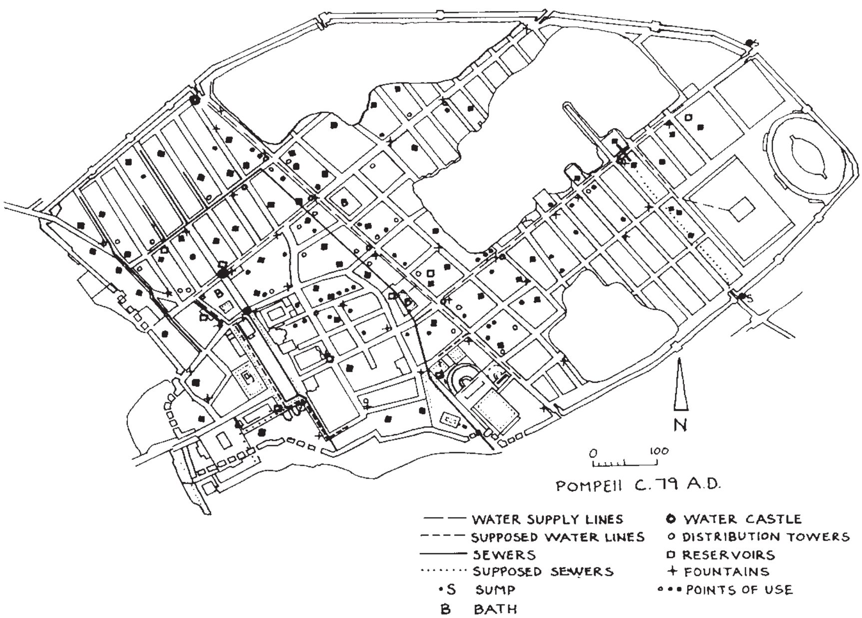


FIGURE 1.41 Plan showing all the known water system elements of Pompeii (circa A.D. 79. (As presented in Crouch, 1993, compiled from maps published by Escherbach, Larsen, and Richardson)

FIGURE 1.42 Downspout tile made of terracotta pipe draining to a cistern under a house in Pompeii. (Photo by Larry W. Mays and copyright by Larry W. Mays)

1.4.2 The *Castellum Divisorium*

The following quote from Vitruvius's treatise on architecture as translated by Morgan (1914) describes how the aqueduct *castellum* worked (as presented in Evans, 1994):

> When it (the water) has reached the city, build a reservoir with a distribution tank in three compartments connected with the reservoir to receive the water, and let the reservoir have three pipes, one for each of the connecting tanks, so that when the water runs over from the tanks at the ends, it may run into the one between them. From this central tank, pipes will be laid to all the basins and fountains; from the second tank, to baths, so that they yield an annual income to the state; and from the third, to private houses, so that water for public use will not run short; for people will be unable to divert it if they have only their own supplies from headquarters. This is the reason why I have made these divisions, and in order that individuals who take water into their houses may by their taxes help to maintain the conducting of the water by the contractors.

This quote from Vitruvius's treatise indicates that three pipes conveyed the water, the first to pools (basins) and fountains (*lacus et salientes*), the second to the public baths (*balneae*), and the third to private houses (*privatae domus*) for revenue to maintain the aqueducts. Even though this has been repeated in the literature on Roman hydraulics and accepted as a canonical arrangement by many scholars (see Hodge, 1992, and Evans, 1994, for references), it has been suggested that Vitruvius's treatise is somewhat in conflict with the actual practice in the Roman world.

Frontinus's treatise, *De aquaeductu urbis Romae,* written 100 years after Vitruvius's treatise, does not agree with Vitruvius's writings. Frontinus recorded 247 *castella* in Rome suggesting that the geographic distribution by regions was the arrangement used for the distribution systems. Unfortunately little remains of these *castella* in Rome. The *castellum divisorium* was a relatively small tank usually located at the edge of a city. Two of the most interesting *castellums* that remain today are those at Pompeii in Italy and at Nîmes in France.

Both the systems at Pompeii and Nîmes indicate that water was not distributed strictly according to function or use, but was distributed based upon the topography and geography of the areas served. On the other hand though, the water distribution systems found in Pompeii and Nîmes were relatively simple compared to those that existed in Rome. The *castellum* at Pompeii, located at Porta Vesuvii, is housed in a large brick building (see Fig. 1.43*a*), and the aqueduct connects from the rear of the *castellum* (see Fig. 1.43*b*). The plan and elevation of the circular basin along with the various components that make up the distribution arrangement of the *castellum* are shown in Fig. 1.44. Water first flowed through two transverse mesh screens (a coarse screen and a fine screen) to remove objects in the water, after which the water flowed into one of the three channels to the outlets. The entrance to each channel had a wooden gate, none of which still exist; however, the bronze fastenings do exist (Hodge, 1992). The three gates were of different heights, with the highest gate on the channel serving the private houses, hydraulically eliminating their supply first. The lowest gate was on the channel serving the public fountains, obviously giving them first priority to the water. This system reflects the principles presented by Vitruvius, but is not the same as pointed out by Hodge (1992).

From the outlet, water flowed into lead pipes. The center one serving public fountains had an external diameter of approximately 30 cm; the two outside pipes had external diameters of approximately 25 cm. Figure 1.45*a* to *e* is a group of photos taken looking into the *castellum* from the center hole on the outside of the building housing the *castellum.* Figure 1.45 *a, b,* and *c* are, respectively, photos of the center channel (for the pipe leading to the public fountains), the left channel (for the pipe leading to the baths and theaters), and the right channel (for the pipe leading to the private houses). Note that by referring to Fig. 1.45*d* it is obvious that the center channel is placed higher than the other two channels. Figure 1.45*e* shows in the background the entrance opening into the *castellum.*

Hodge (1996) reappraised the *castellum divisorium* at Pompeii to conclude that the pipes delivering water to the city were far too large to have been filled by the small aqueduct. He illustrates that Vitruvius's description of dividing up the flow would be mathematically impractical and instead each of the three pipes must have been filled up one at a time using the sluice gate as the control mechanism, thus instituting a system of water rationing. The urban distribution system included secondary *castella,* mounted on brick piers (see Fig. 1.46) throughout the city, and served as storage tanks to supply local demand when the main supply was shut off. The taps found in Pompeii were possibly used to regulate supply. Hodge concluded that this system may have been peculiar only to Pompeii and should not be taken (as it generally has been) as the model of standard Roman urban practice.

The *castellum* at Nîmes, France (see Fig. 1.47*a* and *b*), is located on high ground at the north end of the city. Water entered the circular basin through the approximately 1.2-m-wide by 1.10-m-high (Hodge, 1992) opening shown in Fig. 1.47*a.* A sluice gate, located near the outlet opening and consisting of two verti-

(a)

(b)

FIGURE 1.43 *Castellum divisorium* at Pompeii located at Porta Vesuvii and housed in a large brick building. (*a*) Large brick building housing the *castellum.* Note the three holes at the base of the front. (*b*) Rear of the *castellum* showing the aqueduct. (Photo by Larry W. Mays and copyright by Larry W. Mays)

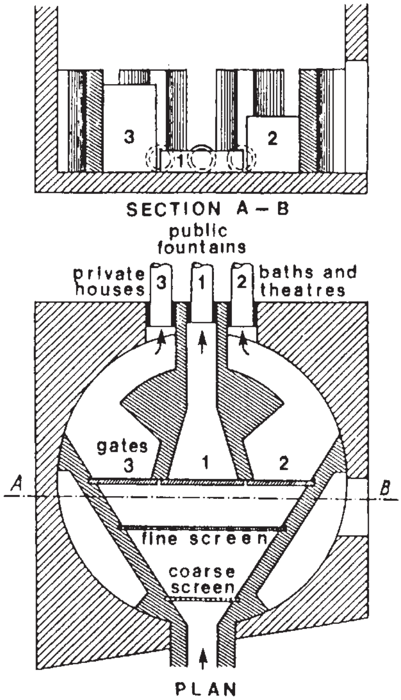


FIGURE 1.44 Plan of the distribution arrangements inside the *castellum divisorium* at Pompeii showing the three gates. (As shown in Hodge, 1992, from Kretzschmer).

cal gates, was used to control flow into the *castellum.* The upper sluice gate was movable, and the lower gate was fixed (Hodge, 1992). The schematic of the *castellum* in Fig. 1.48 shows the holes for 10 large lead pipes and three additional drains in the floor.

Evans (1994) feels that the remains of the distribution tanks (*castella*) that survive at Pompeii and Nîmes (see Figs. 1.43 and 1.47, respectively) indicate that the tanks distributed water according to geography as opposed to use. The pipes from the *castellum,* located along the main streets, carried water to designated neighborhoods, with branched pipes supplying both public basins and private homes (Richardson, 1988).

1.4.3 Pipes and Fountains

Standardized Measures of Roman Pipes. The best approach to explain the system of measures for Roman pipes is to use some quotes from *De aquaeductu urbis Romae* by Sextus Julius Frontinus.

Frontinus (24):

Units of measurement have been established according to digits or inches. That of the digit is followed in Campania and most places of Italy; inches are still followed in Apulia. (2) A digit, moreover, is agreed to be the sixteenth part of a foot, an inch the twelfth part. (3) But unlike the difference between the inch and the digit, there is a double rule for the digit itself. (4) One type is called the square digit, another the round digit. (5) The square digit is larger than the round one by $^3/_{14}$, the round digit smaller than the square one by $^3/_{11}$, precisely because the corners are subtracted.

Frontinus (25):

At a later period another unit of measure developed, which is called the *quinaria* or 5-pipe, taking its origin neither from the inch nor from either type of digit.… The most plausible explanation of the name *quinaria* is that it is derived from its diameter of five quarter-digits, a system that is maintained in units of measure that follow,

(a)

(b)

(c)

FIGURE 1.45 Photos of the three gated channels inside the *castellum divisorium* at Porta Vesuvii in Pompeii. (*a*) Middle channel; (*b*) left channel facing the front of the *castellum*; (*c*) right channel facing front of *castellum.* (Photos by Larry W. Mays and copyright by Larry W. Mays)

FIGURE 1.46 Brick tower on which secondary *castella* (lead storage tanks) were mounted on top of the tower. Lead pipes were used for the flow of water to and from the lead tanks and were placed in the vertical recessed portion shown in the tower. The exiting lead pipes (*calices*) branched off to supply individual customers and also supplied the public fountain shown at the base of the tower. (Photo by Larry W. Mays and copyright by Larry W. Mays)

> up to the *vicenaria,* or 20-pipe, the diameter in each unit increasing by the addition of individual quarter-digits, as in the *senaria,* or 6-pipe, which has a diameter of six-fourths, and the *septenaria,* or 7-pipe, which has seven, and so on by similar increases, up to the *vicenaria.*

Frontinus defined 25 pipe sizes with the first being the *quinaria.*

Frontinus (39):

> The *quinaria* pipe: diameter, 1 digit plus $^{1}/_{4}$; circumference, 3 digits plus $^{11}/_{12}$ plus $^{3}/_{288}$ [3.9272 digits]; capacity, 1 *quinaria.*

The *quinaria* has an inner diameter of 2.31 cm, a circumference of 7.27 cm, and a cross-sectional area of 4.2 cm^2. The *quinaria* was used as the unit for rate of flow (41.5 m^3/day) (Hauck, 1988).

Lead Pipes. The use of lead pipes was the most common method of conveying water from the *castella* to the public fountains and private houses throughout the Roman Empire. It is interesting to note that the Latin word for lead workers was *plumbarii,* which found its way into the English as plumbers and into French as plombiers. Lead pipes were recognized as a health hazard by the Romans, and Vitruvius warned against their use. However, because of the calcium carbonate buildup inside the pipes and the fact that water was moving continuously in the pipes indicates that the Romans most likely did not contract lead poisoning from the lead pipes in their water supply systems (Hodge, 1992). Lead had many advantages including: (1) it was cheap, (2) it was readily available in large quantities, (3) it was easy to handle and malleable enough to form sheets, (4) it had a low melting temperature so it was easy to cast and to solder, (5) it was flexible enough for pipes to be bent around obstructions, and (6) it was strong enough to handle water pressures developed in the water supply systems.

Both Vitruvius and Frontinus discussed the process of making lead pipes. First, the lead was melted and poured out from a melting pot onto a flat surface.

FIGURE 1.47 Views of the *castellum divisorium* at Nîmes. Water enters the circular basin through the rectangular aqueduct opening. Refer to Fig. 1.48 for the plan view with dimensions. (*a*) Shows 6 of the 10 openings for lead pipes used to distribute the water, and (*b*) shows a different view illustrating two of the openings. (Photos by Larry W. Mays and copyright by Larry W. Mays)

Lengths were standard, 10 Roman feet. The molten lead hardened into a long, narrow sheet that was rolled and hammered to develop the sheet. While the lead was still flexible, it was bent up around a wooden or bronze cylinder that was centered in the middle of the sheet. The cylinder core was then removed forming a pipe with the two pipe edges sticking up. These edges were then soldered, folded, welded, or hammered together forming a seam. As shown in Fig. 1.49*a* the pipe was placed so that the seam was on top. This obviously was done to facilitate pipe repairs such as leaking and failed joints and seams. Also note the joint shown in Fig. 1.49*a,* which appears to be a heavily soldered joint. Other methods of joining pipes were to use an overlapping male-female joint or to use a covered close-fitted sleeve.

During the manufacturing process, inscriptions in raised lettering were placed on the sheets of lead before they were bent together. Movable molds of individual letters were used to form the large letters. Keep in mind these letters were placed on the lead sheets at the same time the pipes were being made.

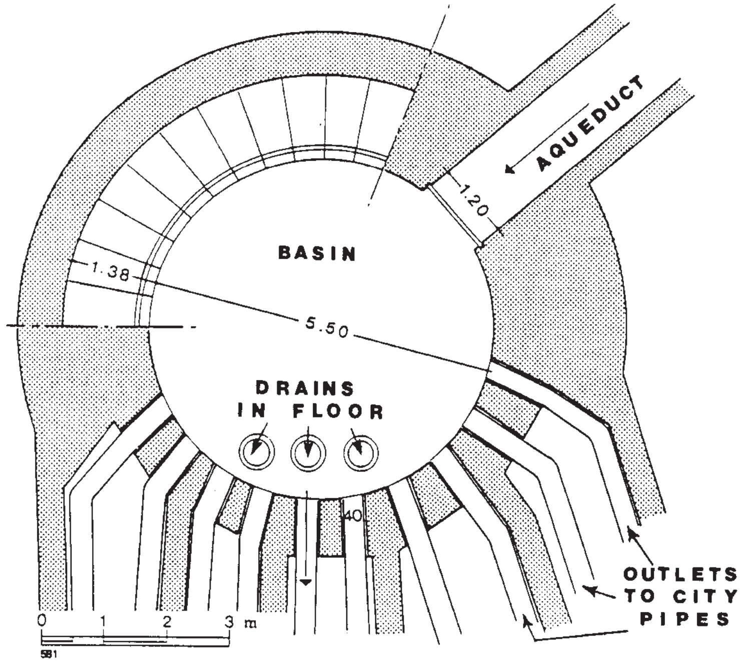
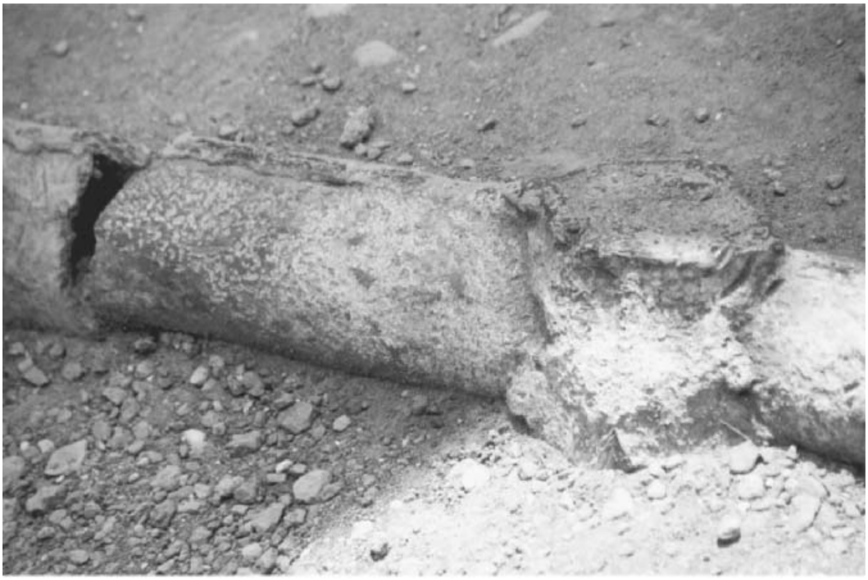


FIGURE 1.48 Plan view of the Nîmes *castellum divisorium* showing the 10 outlets and the 3 drains in the floor of the 1.5-m-diameter basin. (Hodge, 1992, from Adam).

Junctions were used throughout the systems for joining smaller pipes to larger ones or for branching the system to individual customers. Some junction boxes were simply a cylinder laid horizontally with pipes entering and leaving each end, such as the junction shown in Fig. 1.49*b.* Other junction boxes had one entering pipe at a right angle to form a T with the junction box and two or more smaller pipes leaving at right angles on the other side of the junction box.

Layout of Street Pipes. Two approaches to laying out the piping network were: (1) using a main pipe from the secondary *castellum* with smaller branch pipes attached to serve individual customers, and (2) not using main pipes but using individual pipes laid from the secondary *castellum* to the individual customer. The second approach may have been the normal Roman practice (see Hodge, 1992, p. 320).

Pompeii's water distribution system consisted of pipes along the main streets connecting the main *castellum* at Porta Vesuvii to the various water towers (secondary *castella*), from which smaller pipes were placed under the sidewalks and streets and served the various customers. Not all customers had individual lines to a secondary *castellum* but instead received their supply from taps into the system at their houses.



(a)

(b)

FIGURE 1.49 Lead pipes in Pompeii. (*a*) Lead pipe and joint, and (*b*) lead pipe with enlarged section of a junction box. (Photo by Larry W. Mays and copyright by Larry W. Mays)

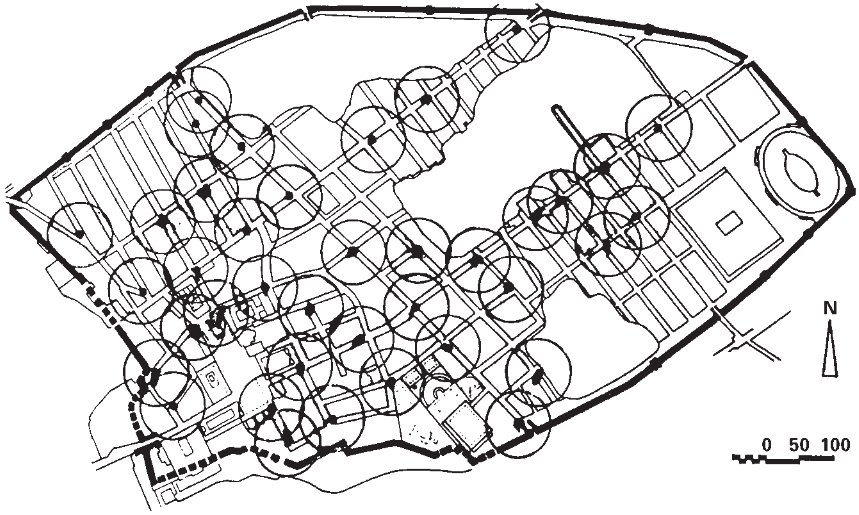


FIGURE 1.50 Pompeii, showing locations of public water fountains. Circles have a radius of 50 m. [As presented in Crouch (1993) from Eschebach (1983)]

Public Fountains. Figure 1.46 illustrates a public fountain at the base of a water tower (secondary *castellum*) located in Pompeii. The public fountains in Pompeii were placed at somewhat evenly spaced locations (Hodge, 1992) as shown in Fig. 1.50. The radius of each circle in the figure is 50 m, so fountains were commonly spaced about 100 m apart in Pompeii. Figure 1.51 shows two public water fountains placed back to back along a street. Because there is no water tower at these fountains, water was supplied from a water tower at another location.

The fountains in Pompeii were oblong-shaped stone basins (typically 1.5 m × 1.8 m and 0.8 m high, as described by Hodge, 1992, p. 306). As shown in Figs. 1.46 and 1.51 the delivery spout was on a stone pier overlooking the stone basin. Examining Fig. 1.51 you can see the overflow weir from which water normally flowed into the streets and then into the drainage system. Most likely, overflow water was used for street cleaning and helped flush the sewer system. Drain holes were also placed at the bottom of the stone tanks and were plugged. The plug was removed for purposes of cleaning the stone basins.

1.5 AFTER THE ROMANS

The fall of the Roman Empire extended over a 1000-year transition period called the Dark Ages. During this period, the concepts of science related to water resources probably retrogressed. After the fall of the Roman Empire, water sani-

FIGURE 1.51 Public fountain in Pompeii. (Photo by Larry W. Mays and copyright by Larry W. Mays)

tation and public health declined in Europe. Historical accounts tell of incredibly unsanitary conditions—polluted water, human and animal wastes in the streets, and water thrown out of windows onto passersby. Various epidemics ravaged Europe. During the same period, Islamic cultures, on the periphery of Europe, had religiously mandated high levels of personal hygiene, along with highly developed water supplies and adequate sanitation systems.

The Hakali water conveyance system, the first of three main systems feeding Istanbul, Turkey, dates back to the early Byzantine period. According to Ozis (1987) the most interesting waterworks of the Byzantine period were the several dozens of cisterns, some of which covered 70 m $\times$ 140 m, such as the Yerebatan cistern in Istanbul. Some of these systems collected precipitation water, and others served the seasonal regulation of water from the large conveyance systems. Others were used to level off the topography for foundations of buildings and to provide additional height to the buildings.

During the Seljukian period (1071–1308) there were no large water systems built. However, the architecture is well known for its impressive and highly ornamented buildings and arched bridges. The Ottoman period (1281–1922) was when the Istanbul water system was developed. After the conquest of Istanbul in 1453, the Roman Hakali water conveyance system was restored, with expansions made until the middle of the eighteenth century.

For additional information on ancient urban water supply see Ashby (1935) and Hodge (1998).

1.6 REFERENCES

Adams, R. McC., *Heartland of Cities: Surveys of Ancient Settlements and Land Use on the Central Floodplain of the Euphrates,* University of Chicago Press, Chicago, 1981.

Aicher, P. J., *Guide to the Aqueducts of Rome,* Bolchazy-Carducci Publisher's Inc., Wauconda, Ill., 1995.

Ashby, T., *The Aqueducts of Rome*, McGrath, Oxford, 1935.

Berry, B. J. L., "Cities as Systems within Systems of Cities," *Papers and Proceedings of the Regional Science Association,* vol. 13, 1964, pp. 146–163.

Bruun, C., *The Water Supply of Ancient Rome: A Study in Roman Imperial Administration,* Commentationes Humanarum 93, Helsinki, 1991.

Crouch, D., *Water Management in Ancient Greek Cities,* Oxford University Press, New York, 1993.

Cullinane, Jr., M. J., "Methodologies for the Evaluation of Water Distribution System Reliability/Availability," Ph.D. dissertation, University of Texas at Austin, May 1989.

Eschebach, H., "Die Innerstädtische Gebrauchwasserversorgung, Dargestellt am Beispiel Pompejis," in J.-P. Boucher (ed.), *Journées d'Études sur les Aqueducs Romains,* Paris, 1983, pp. 79–132.

Evans, A. *The Palace of Minos,* vol. I., Biblo and Tannen, New York, 1964. See especially pp. 142–143 and 226–230. Reprint of 1921–1935 edition.

Evans, H. B., *Water Distribution in Ancient Rome: The Evidence of Frontinus,* University of Michigan Press, Ann Arbor, 1994.

Fahlbusch, H., "Vitruvius and Frontinus—Hydraulics in the Roman Period," in G. Garbrecht (ed.), *Hydraulics and Hydraulic Research: A Historical Review,* A. A. Balkema, Boston, 1987.

Frontinus, Sextus Julius, *The Two Books on the Water Supply for the City of Rome, AD 97,* translated by C. Hershel, republished by the New England Water Works Association, Boston, 1973.

Garbrecht, G., *Wasserversorgung Stechnik in romischer Zeit,* R. Oldenbourg Verlag, München, Wein, 1982.

Garraty, X., and X. Gay, The Columbia History of the World, Harper and Row, Publishers, Inc., 1972.

Hauck, G., *The Aqueduct of Nemausus,* McFarland & Company, Inc., Jefferson, N.C., 1988.

Hodge, A. T., *Roman Aqueducts & Water Supply,* Gerald Duckworth & Co., Ltd., London, 1992.

Hodge, A. T., "In Vitruvium Pompeianum: Urban Water Distribution Reappraised," *American Journal of Archaeology,* vol. 100, no. 2, April 1996.

Hodge, A. T., *Ancient Greek France,* University of Pennsylvania Press, Philadelphia, 1998.

Lees, L. H., and P. M. Hohenberg, "How Cities Grew in the Western World: A Systems Approach," in J. H. Ausubel and R. Herman (eds.), *Cities and Their Vital Systems; Infrasructure Past, Present, and Future,* Series on Technology and Social Priorities, National Academy of Engineering, National Academy Press, Washington, D. C., 1988.

Maiuri, A., "Pompeii Pozzi e Condotture d'Acqua…" and "Scoperta di Grandi Conditture in Piombo dell'Acquedoto Urbano," *Notizie degli Scavi di Antichita,* vol. 7, 1931, pp. 546–576.

Morgan, M. H. (translator), *Vitruvius, The Ten Books on Architecture,* Cambridge, Mass., 1914.

Ozis, U., "Ancient Water Works in Anatolia," *Water Resources Development,* vol. 3, no. 1, Butterworth & Co. Ltd., 1987.

Ozis, U., "Historical Water Schemes in Turkey," *Water Resources Development,* vol. 12, no. 3, 1996, pp. 347–383.

Ozis, U., and N. Harmancioglu, "Some Ancient Water Works in Anatolia," *Proceedings International Seminar on Karst Hydrogeology,* Oymapinar, Anatolia, Turkey, October 1979, pp. 380–385.

Richardson, Jr., L., *Pompeii: An Architectural History,* Johns Hopkins University Press, Baltimore, 1988.

Robbins, F. W., *The Story of Water Supply,* Oxford University Press, London, 1946.

Van Deman, E. B., *The Building of Roman Aqueducts,* Carnegie Institute of Washington,, D.C., 1934.

1.7 INTERNET SOURCES

Crane, G. R. (ed.), The Perseus Project, *http://www.perseus.tufts.edu*

James, W., "A Historical Perspective on the Development of Urban Water Systems," *http://www.eos.uoguelph.ca/webfiles/james/homepage/Teaching/437he.html*

"The History of Plumbing—CRETE," *http://www.theplumber.com/crete.html*

Yegul, F., "Roman Building Technology and Architecture," *http://id-archserve.ucsb.edu/AH152K/index.html*

"Roman Earthworks," *http://www.sky.net/~eichmann/fortplnl.html*

"Sextus Julius Frontinus: The Aqueducts of Rome," *http://www.ukans.edu_exts/Frontinus/De_Aquis/text*.html*

"Water Supply Systems: Cisterns, Reservoirs, Aqueducts," *http://archserve.id.ucsb.edu/AH152K/water.htm*

CHAPTER 2
STRATEGIC PLANNING FRAMEWORK FOR SMALL WATER SYSTEMS

Janice A. Beecher
Institute of Public Utilities
Michigan State University
East Lansing, Michigan

Peter E. Shanaghan
U.S. Environmental Protection Agency
Washington, D.C.

2.1 INTRODUCTION

Strategic planning is a disciplined effort to guide an organization's future in terms of its purpose, structure, and functions. Strategic planning for small water utilities is not unlike strategic planning for small organizations in general (public, private, or nonprofit). Of course, the inputs and outcomes of the planning process are specific to the essential nature of supplying drinking water to the public.

The imperative for strategic planning by water systems has never been clearer. As the twenty-first century dawns, the drinking water industry in the United States finds itself in a period of rapid and tumultuous change. Much of the basic water utility infrastructure in the United States is reaching the end of its useful life and must be rehabilitated or replaced. Development of new water supply sources is becoming increasingly difficult, and great emphasis is being placed on protecting source waters from contamination. Drinking water utilities are facing unprecedented and increasing competitive pressure. Ongoing implementation of the 1996 Safe Drinking Water Act (SDWA) includes a flurry of new or tightened regulatory standards. Finally, public expectations have never been higher both in terms of product quality and service quality.

The environment in which small water systems operate today is exceptionally dynamic. Among the many dynamic forces at work, SDWA regulatory standards are likely to first drive small systems to respond. This is quite simply the result of the fact that SDWA regulatory standards carry with them a date certain by which systems must take action. Systems usually have 3 to 5 years from the date of promulgation of a SDWA regulation to achieve regulatory compliance. To varying degrees, other pressures on water systems—to replace the infrastructure, development new water sources, and to respond to competitive forces—can, to varying degrees, be ignored or dealt with in a "patch and get by" mode for a somewhat longer time horizon. Forestalling attention to these forces, however, can diminish the quality of service and raise the cost of the ultimate solution.

Over the next several years, water systems must comply with numerous regulations. Depending upon the specific circumstances of a given system, compliance will involve varying degrees of (1) investment in capital, (2) enhancement of operations, and (3) improvement to management practices. This regulatory mandate for action provides a perfect opportunity for systems to think and plan strategically for their future.

2.2 STRATEGIC PLANNING

The many types of plans that water systems might prepare include business or financial plans, capital-facility plans, operation and maintenance plans, watershed protection plans, resource or conservation plans, and emergency or contingency plans. Certainly, larger water utilities engage in a variety of planning processes. Yet many water systems, especially small water systems, have not engaged in systematic planning, strategic or otherwise.[1] Historically, planning by small water systems has been limited to that associated with major capital projects and required by various state regulatory or financial-assistance agencies. However, the requirements for and practice of planning have been limited.

The traditional method of long-range planning for water systems generally involves development of a set of actions to accomplish a goal or set of goals over a period of years, with the assumption that the future will be relatively stable and somewhat predictable. *Traditional long-range planning* does not necessarily prepare an organization to successfully respond to a dynamic or changing environment.

The focus of strategic planning is on preparing organizations to successfully meet challenges and opportunities, both known and unknown. Strategic planning does not make decisions for the future, but shapes the environment in which future decisions will be made (Steiner, 1979). Strategic planning does not deal with future decisions but the future of present decisions and preparing today for an uncertain tomorrow (Drucker, 1997).

As noted, strategic planning is disciplined. Strategic planning also is goal-oriented, comprehensive, and adaptive. An organization's goals must be articulated in order to guide the entire planning process and evaluate results. Planning should be comprehensive, as well as integrative, in terms of the entire organization and its elements. The resultant plan must be adaptive to a dynamic environment and changing circumstances.

Strategic planning is the basis for strategic management. The organization's leaders must be strategic thinkers and committed to the planning process. They must engage in an uncompromisingly critical and continuous assessment of what the organization does and why and how it does it. *Strategic management* helps organizations survive and maintain relevance. During dynamic times, organizations are more likely to succeed if they engage in strategy rather than accept the inertia of "if it ain't broke, don't fix it" (Bryson, 1995). Over time, organizations should internalize strategic thinking into day-to-day operations.

The potential benefits of planning for water systems are far-reaching in terms of ensuring compliance with drinking water standards, enhancing water system capacity and performance, and promoting both sustainability and continuous improvement.

2.3 PLANNING FRAMEWORK FOR SMALL WATER SYSTEMS

The planning process need not be daunting, even to systems with limited resources. Although grounded in the rational-choice and systems models of public administration, planning is as much common sense as theory. The planning process can be subdivided into various logical steps. A simple, commonsense framework can guide strategic planning for small water systems:

1. Specify the system's goals and objectives relative to its mission.
2. Assess the water system's organizational structure and service roles.
3. Identify external influences (challenges and opportunities).
4. Evaluate internal capacity (technical, financial, managerial).
5. Analyze strategic options for achieving compliance and other goals.
6. Implement the preferred planning alternative.
7. Monitor and evaluate outcomes and make adjustments as needed.

Each step can be encapsulated by a few critical questions (Table 2.1) and each represents a progression in a cumulative planning process. Although laid out in a linear format, planning should be approached as a nonlinear exercise. As planners work on subsequent steps, they may need to revisit assumptions made and con-

TABLE 2.1 Planning Steps and Key Questions

Planning step	Key questions to guide planners
Step 1: Specify mission and goals	• What is the water system's mission? • What values guide the water system? • What are the system's immediate and long-term goals? • Are values and goals established in an open and participatory process that includes employees, customers, and other stakeholders?
Step 2: Assess structure and roles	• How is the water system structured in terms of the hierarchy from governance, to management, to operations? • What service functions does the water system presently provide? • What operational tasks does the water system perform? • What is the role of the water system for each service function?
Step 3: Identify challenges and opportunities	• What are the principal change factors or drivers affecting the water system? • What challenges are presented? • What opportunities are presented?
Step 4: Evaluate system capacity	• Does the water system have adequate technical, financial, and managerial capacity? • What are the systems strengths (performance-enhancing factors) and weaknesses (performance-limiting factors)? • Can the system manage change and effectively respond to external challenges and opportunities?
Step 5: Identify strategic options	• What strategic options are available to the system for achieving its goals? • What benefits and costs are associated with each option? • How are the system's technological and structural options interrelated?
Step 6: Choose a strategy	• Which strategic option (or combination of options) can best provide the system's service roles and functions? • How do options compare in terms of cost-effectiveness? • Which alternative is optimal in terms of the selection criteria?
Step 7: Implement and monitor	• What implementation issues are presented by the strategy and how will they be addressed? • How will the strategy be monitored over time to ensure success? • Is the plan producing desired outputs and achieving desired outcomes?

clusions reached in earlier steps. The planning process includes continuous feedback loops from later steps and planning outcomes back to the earlier steps.

2.3.1 Step 1: Specify Mission and Goals

A positive first step in the planning process is to contemplate the water system's mission and goals. Like all organizations, water systems should be guided by a mission statement.[2] Writing a mission statement is a meaningful exercise because it should play a significant role in shaping the organization and what it does.

Establishing goals, objectives, and priorities for a specified planning horizon is a critical early planning step. Water systems generally have multiple goals. These might include compliance with applicable standards, maintenance of a reliable water source, efficient management and operations, affordable rates for water customers, and excellent customer service. Ideally, the water system's goals will be operationalized or defined in measurable terms so that managers can gauge progress. An example is a goal related to customer satisfaction that can be measured through a survey. Generally, the planning horizon should be a minimum of 10 years.

The goals and objectives of a water system will be shaped not only by its overarching mission but also by its core values. These values often are community-specific and can be discovered and refined through participatory planning processes. Involving the public in the planning process has a number of potential benefits for water systems. In a direct sense, public involvement can help the water system increase the public's awareness of water issues, expand the range of viable planning options, and build support. Public involvement may also help reduce the public's reluctance to pay for water service and improve demand-side behavior.

2.3.2 Step 2: Assess Structure and Roles

In this step, the planner assesses the water system's organizational structure and service roles, as presently constituted and anticipated for the planning horizon. The water system must be structured in a manner that will facilitate the fulfillment of its mission and goals.

The water system's basic organizational structure can be represented by a three-part hierarchy from governance, to management, to operations (Fig. 2.1). Governance is at the top of the hierarchy because it refers to the ultimate accountability for the water system, which may rest with a board of directors. Management focuses on responsibility and provides the link between governance and actual operations; strategic planning is a management function. Operation focuses on performance and involves the direct performance of functional tasks. For water systems, one or more entities must provide governance and assume ultimate responsibility for the provision of service. The same entity or entities may or may not provide management and operation of the water system or its various functions.

At the operational level, water utilities are similar to other utilities in terms of basic utility functions: source-water development and protection, drinking water treatment, treated water storage, transmission and distribution, retail customer services, and regulatory monitoring and reporting. Performance of the physical delivery functions requires both long-term investment in capital facilities and specific operational and maintenance tasks over time. Traditional water utilities tend to be vertically integrated; that is, the utility operates all functions and provides them on a bundled basis to customers. The separation of functions in this manner (or *unbundling*) is useful for strategic planning purposes because it encourages the consideration of alternative service roles for the water utility.

Assuming a role, however, involves choice; not every system must assume every role for every service function. Through restructuring, some roles and responsibilities may be retained while others are shifted. Over time, roles can change and evolve, and responsibilities for some service roles and functions can be assigned to others. Some systems, for example, might purchase wholesale treated water and concentrate efforts on water distribution and customer services. Others might contract with a private company for specific projects or for general

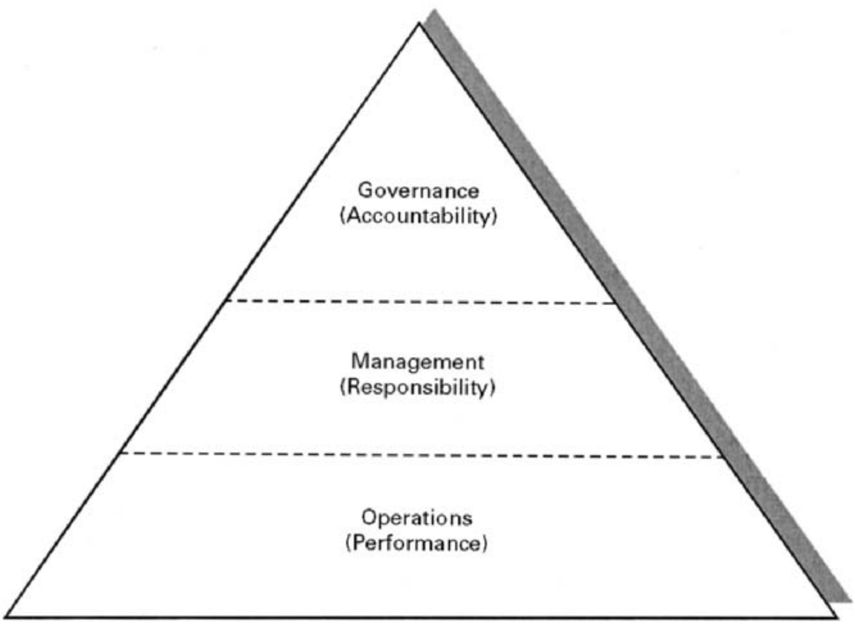


FIGURE 2.1 Water system organizational hierarchy.

operations. Still others might enter into long-term partnership agreements with other systems (such as satellite management) or merge utilities altogether to harness scale economies and improve service quality. The water system's structure and roles should be revisited throughout the strategic planning process.

2.3.3 Step 3: Identify Challenges and Opportunities

In the third step, planners look outside of the utility organization and systematically consider the full range of challenges and opportunities for the water system. By accounting for challenges and opportunities through the planning process, the water system can position itself strategically to meet challenges effectively and take full advantage of opportunities.

The particular needs of individual water systems vary, but some challenges are common to many water systems today. Some of the more significant sources of pressure are compliance with regulatory standards (drinking water and other), financial constraints and rate concerns, rising consumer expectations, water supply quality or quantity issues, substantial infrastructure needs, changing demand and demographic patterns, and competition from other providers. For some water systems, the external environment can exert significant pressure for tactical or structural change. Water planners should strive to identify the particular pressures that will affect the water system across its planning horizon.

Drinking water regulations are a major force driving the need for strategic planning because of the certainty of the SDWA compliance schedule. Many external forces do not follow a predictable time line, but regulations contain fixed dates by which compliance must be achieved. The specific types of activities that utilities will be required to undertake in order to comply with SDWA regulations will vary greatly from one utility to the next. Compliance with different rules requires focus on different types of activities. Most rules involve some monitoring, some involve optimization of existing treatment processes; others involve installation of major new treatment equipment; and others require strong management attention.[3]

In addition to the SDWA, water systems must also comply with the Clean Water Act, as well as a number of other federal, state, and local regulatory standards. These include regulations concerning environmental impacts, resource protection, endangered species, public health, local finances, historic preservation, occupational safety and health, minority businesses, Americans with disabilities. Water systems must understand the full range of existing and emerging regulations with which they must comply and incorporate compliance planning in the strategic planning.

Despite the many challenges they face, many water systems also enjoy a number of opportunities. These include, but are not limited to, programs for educating and training personnel, technical assistance from governmental agencies and professional organizations, grant and loan programs to build financial capacity, partnerships and alliances with other utilities, techniques for improving operational efficiency, integrated-resource management, and technological innovation. Planners should strive to identify the broadest range of opportunities available to the water system for enhancing the achievement of its mission and goals.

2.3.4 Step 4: Evaluate System Capacity

A critical evaluation of the utility's internal capacity plays a key role in effective strategic planning. Likewise, strategic planning is a basic and essential tool for capacity development.[4] *Water system capacity* is defined in terms of three interrelated elements: technical, financial, and managerial. Each element of capacity is necessary but not sufficient by itself to sustain the water system over time. The numerous tools used in capacity assessment and development can be readily adapted to the strategic planning process. Table 2.2 provides an overview of capacity-development concepts.

Technical capacity is defined in terms of three areas: source-water adequacy, infrastructure adequacy, and operations and maintenance. The water system will need to create and maintain a detailed inventory of the physical infrastructure used to deliver drinking water, organized according to functional areas: source-water development, water treatment, treated water storage, and transmission and distribution. Facilities used for general administration and retail service should also be assessed. The inventory and plan should include an appraisal of physical conditions, as well as a review of operation and maintenance practices.

Financial capacity is defined in terms of revenue sufficiency, creditworthiness, and fiscal management and controls. Revenue sufficiency or financial sustainability generally requires systems to implement cost-based rates for water service. Creditworthiness indicates that the system can fund capital improvements through appropriate financial markets. Fiscal management and controls ensure that revenues and costs are accounted for through generally accepted procedures.

Managerial capacity is defined in terms of ownership accountability, staffing and organization, and external linkages. Water system owners must be identifiable and accountable. The staffing and organization of the utility must be appropriate to the system's mission and goals, as well as the challenges it must meet. The concept of external linkages refers to the ability of the water system to interact effectively with all relevant stakeholders, including customers and regulators.

Strategic planning can enhance all three areas of capacity. For each element of capacity, the planners should make a broad assessment across the utility to identify strengths and weaknesses in each of the elements of capacity. Strengths are performance-enhancing factors; weaknesses are performance-limiting factors. The assessment should consider strengths and weaknesses in light of anticipated challenges and opportunities for the water system identified in the previous plan-

TABLE 2.2 Overview of Water System Capacity

Elements of capacity	Key indicators
	Technical
Source-water adequacy	• Does the system have a reliable source of drinking water? • Is the source of generally good quality and adequately protected?
Infrastructure adequacy	• Can the system provide water that meets SDWA standards? • What is the condition of its infrastructure, including wells and intakes, treatment, storage, and distribution? • Does the system have a capital improvement plan?
Technical knowledge and implementation	• Is the system operator certified? • Does the operator have sufficient technical knowledge of applicable standards? • Can the operator effectively implement technical knowledge? • Does the operator understand the system's technical and operational characteristics? • Does the system have an effective operation and maintenance program?
	Financial
Revenue sufficiency	• Are the system's costs and revenues known and measurable? • Are system assets properly valued and reflected in rates? • Do revenues from rates and charges cover system costs?
Credit worthiness	• Is the system financially healthy, as measured through indicators, ratios, and ratings? • Does it have a credit record and access to capital through public or private sources? • Can it provide assurance of repayment?
Fiscal management and controls	• Are adequate books and records maintained? • Are appropriate budgeting, accounting, and financial planning methods used? • Does the system manage its revenues effectively?
	Managerial
Ownership structure and accountability	• Are the system owners clearly identified? • Can owners be held accountable for the system?
Staffing and organization	• Are the system operators and managers clearly identified? • Is the system properly organized and staffed? • Do personnel have adequate expertise to manage operations? • Do personnel understand the regulatory requirements? • Do personnel have the necessary licenses and certifications?

TABLE 2.2 Overview of Water System Capacity (*Continued*)

Elements of capacity	Key indicators
	Managerial
External	• Does the system interact well with customers, regulators, and other entities? • Is the system aware of available external resources, such as technical and financial assistance?

ning step. Planners should consider whether the system is capable of managing change and responding effectively to external challenges and opportunities.

2.3.5 Step 5: Identify Strategic Options

Water systems that have a clear understanding of their service roles, as well as their external environment and internal capacity, are well positioned to identify *strategic options* that will best fulfill their mission and goals. Strategic options may be *tactical* (which can be implemented within the water system's existing organizational framework) or *structural* (which require fundamental changes in the organization of the water system or its service roles).

The basic planning model can be used to identify strategic options for a number of planning issues, including but not limited to compliance with drinking water standards. The SDWA identifies a number of alternative paths to compliance for small water systems. These include conventional and centralized water treatment options, decentralized treatment (point of use and point of entry), water supply alternatives (ground and surface sources), interconnection with another system (for water purchasing), and restructuring (changes ownership or operations). Planning can—and should—expand beyond SDWA compliance. An array of strategic options can be identified for other planning issues. These might include aesthetic and quality issues, customer service issues, supply shortages or unreliability, infrastructure challenges, and conservation and efficiency. For any water system, planners should identify strategic options based on particular issues that apply to the water system for its planning horizon.

When identifying options, some general guidelines are helpful. First, planners should think *comprehensively* and consider the widest possible range of alternatives. Second, planners should consider options across a broad *spatial* horizon. This requires looking to the regional context, including the needs and circumstances of nearby water systems. Third, planners should consider potential options over a long *temporal* horizon. Some options that may not seem feasible in the near term may be decidedly more feasible in the longer term. A long-term time frame is essential for identifying the best long-term solutions. By expanding the spatial and temporal planning horizons, more solutions and opportunities may be revealed in the planning process (Fig. 2.2).

2.3.6 Step 6: Choose a Strategy

Choosing a planning strategy for the water system is facilitated by an options analysis, in which planners fully consider all tactical and structural options and identify the optimum solution for a particular utility. For many water systems, a combination of strategies will yield the best plan of action.

Integrative decision making involves the joint consideration of tactical and structural options. Doing so requires a degree of *nonlinear* thinking on the part of planners. In other words, the best solutions for the long term might not simply line up in order or reflect a clear sequence of steps for implementation. In some instances, a structural option (such as a change in management or ownership) may open the door to other tactical options (such as a change in technical approach) for fulfilling one or more of the system's basic service functions.

The analysis of options requires an evaluation framework. The planner will need to establish a framework that is reasonable for the water system. Some potential evaluation criteria include:

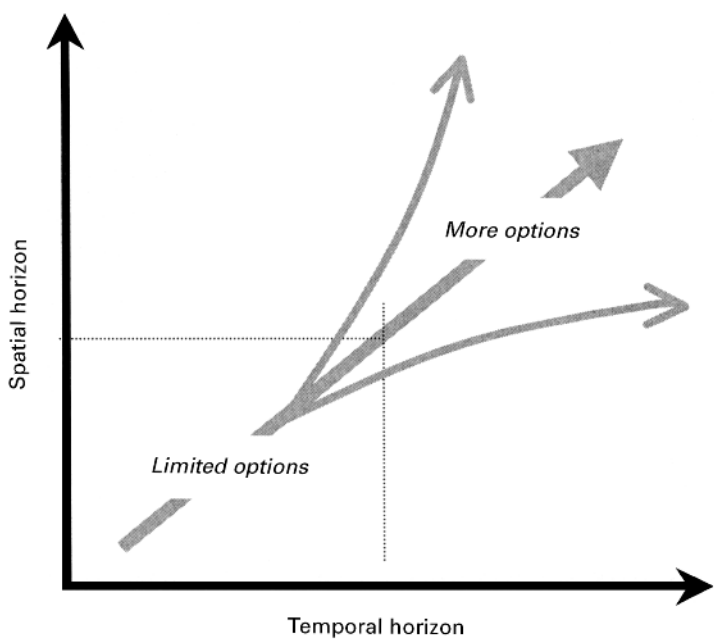


FIGURE 2.2 Temporal and spatial planning horizons.

- Consistency with the water system's overarching mission and goals
- Compliance with drinking water and other standards and regulations
- Impact of capacity development (technical, financial, and managerial)
- Economic feasibility in terms of the total cost of implementation
- Operational efficiency in terms of the unit cost or production
- Structural efficiency in terms of total long-run societal cost
- Quality of service for customers of the water system
- Reliability of service for customers of the water system
- Practicality of implementation for water system managers
- Political acceptance in the local community by all stakeholders
- Regulatory acceptance by drinking water and other agencies
- Customer acceptance throughout the water system's service territory

Optimization involves choosing the strategy (the option or combination of options) that best meets the evaluation criteria established for the system. Strategic options can be evaluated using a number of analytical tools. Qualitative assessment involves a simple scoring or ranking of options according to the specified planning criteria. Extra weight can be assigned to criteria as deemed appropriate by decision makers and in accordance with prevailing values.

Other assessment tools are more quantitative, including the many variations of benefit-cost analysis. Avoided-cost analysis is a metric for evaluating the cost-effectiveness of planning alternatives for achieving a specified benefit (such as compliance with standards or the provision of a desired level of service). The analysis compares the cost of alternatives to a benchmark that reflects the cost associated with achieving the desired benefit through typical or conventional means. The difference between the conventional option and a better option is the cost avoided by implementing the better option. Avoided-cost analysis is used widely to measure the potential benefits of water conservation.

2.3.7 Step 7: Implement and Monitor

Once the best option or options are identified, planners need to identify the steps necessary for implementation. An action plan can help guide implementation by specifying key dates and actions that must be taken.

Water systems may face both internal and external implementation issues. First, for some organizations, implementation can be thwarted by *inertia,* or a generalized resistance to change associated with uncertainty or other issues. Second, the water system must have adequate technical, financial, and managerial *resources* for implementation. Third, the *leadership* of the organization must be prepared for and committed to the implementation process. Fourth, implementation may require *organizational or personnel changes,* including special training for technical or managerial staff members. Finally, over time, successful implementation may depend on how change is managed, as well as how organizational *conflicts are resolved.*

External implementation issues may be significant as well. First, some options may require environmental or economic *regulatory approvals,* including certification or permitting. Second, implementation may raise any special legal or *liability issues.* Third, the success of the process may rely on whether *stakeholders* are informed about, involved in, and supportive of the process. Fourth, implementation of the strategy may require special funding for implementation from external public or private sources. Finally, the organization must be prepared to adapt to change and uncertainty in the external environment (such as the economy).

The successful utility also will constantly reassess its strategic plan and strive for continuous improvement. The *optimum* strategy must be continually reevaluated to ensure its continued optimality. The evaluation of a plan should distinguish between outputs (the actions taken to implement the plan) and outcomes (the actual consequences or results of the strategy).[5] Outcomes can be direct or indirect, expected or unexpected, and intended or unintended. Performance measures based on the water system's goals and objectives, as well as the evaluation criteria used to develop the strategy, can be used to evaluate outcomes. The feedback loop connects planning outcomes to the formulation of goals and supports the process of continuous improvement. Reassessment will lead to adjustments to the strategic plan in response to changes in the external environment, as well as changes in the internal capacity of the water system.

2.4 CONCLUSIONS

Strategic planning is a dynamic and ongoing process that supports the continuous improvement of water systems. Planning encourages strategic thinking by managers on a day-to-day basis, with internalization of goals and commitment to the process. Planning requires continual assessment and adaptation. Fortunately, many tools and resources are available to support the planning process. Clearly, the benefits of strategic planning—including the discipline that the planning process brings—should outweigh the costs.

The basic planning framework presented here may seem overly complex to some and overly simplified to others. The framework provides a good starting point for systems that have not engaged in any form of strategic planning. For most water systems, the goal of planning is not perfection but improvement, that is, to move in a positive direction despite limited resources and an uncertain future. In the long term, it is more important that the system makes informed choices than flawless ones.

2.5 ENDNOTES

1. About 54,400 community water systems (CWSs) operate in the United States. (CWSs serve at least 15 service connections used by year-round residents or regularly serve at least 25 year-round residents) (USEPA, 1997). Nearly 95 percent of these systems serve populations of 10,000 persons or fewer.
2. Examples of mission statements can be found on water system web sites.
3. The universe of systems subject to various rules is different, but generally fewer than 20 percent of systems subject to any rule are expected to have to undertake major capital investment to comply.
4. The SDWA places significant emphasis on capacity development for small water systems. States must ensure that all new systems have adequate capacity and also implement a strategy for improving the capacity of existing systems. Federal funding cannot be provided to systems that lack capacity or that will not achieve capacity with the benefit of the funding.
5. As an example, an output might be the hiring of a customer service representative; an outcome might be the change in customer satisfaction that results from an improvement in service.

CHAPTER 3
IMPROVING URBAN WATER INFRASTRUCTURE THROUGH PUBLIC-PRIVATE PARTNERSHIPS

Robin A. Johnson
Adrian T. Moore
Reason Public Policy Institute
Los Angeles, California

Water issues facing urban communities are part of a growing worldwide debate over the challenges of providing safe drinking water supplies and treated wastewater. While oil was described as the key commodity of the twentieth century, water is increasingly seen as the substance that will influence the course of human progress and economic development in the twenty-first century.

Recent issues have elevated the debate over water infrastructure to the top of the national agenda. President George W. Bush's proposed changes in acceptable arsenic levels in drinking water in March 2001 provoked a debate about the proper balance between regulatory mandates and the costs of compliance. The Water Infrastructure Network (WIN), a coalition of local elected officials, drinking water and wastewater providers, state environmental and health administrators, and others, proposed a multibillion dollar federal grant program to help finance local water infrastructure needs. In addition, the Water Infrastructure Caucus was formed in the U.S. Congress to highlight the need for water and wastewater improvements.

At the same time, many cities have been turning to the private sector for assistance in meeting water infrastructure needs. A variety of factors are contributing to increased private sector involvement, including deteriorating water infrastructures, unfunded federal and state environmental mandates, costly capital improvements, and the desire among many public officials to operate water facilities more efficiently. Changes in federal rules during the Bush and Clinton administrations also helped to facilitate more public-private partnerships for water and wastewater services. These trends have combined to fuel growth in the number of cities entering into public-private partnerships to enhance daily operations and improve water infrastructures.

This chapter explores the growth of water and wastewater public-private partnerships and examines the extent of partnerships in urban communities. We look at some factors causing the growth in public-private partnerships and obstacles that must often be overcome. Next is an examination of the process involved in considering a public-private partnership, including legal, financial, economic, and political concerns. Finally, we include a number of case studies from urban communities that have successfully implemented water and wastewater public-private partnerships.

3.1 BACKGROUND OF WATER AND WASTEWATER PARTNERSHIPS

In the United States, most urban residents have long received their water from public water utilities and had wastewater treated by local government agencies. For example, private water utilities currently serve about 15 percent of the U.S. population, although that figure is growing. In other parts of the world, private firms have long provided water management services. In France, for example, privatization goes back to charters issued by Napoleon III.

The present-day structure of the water and wastewater industries had its genesis in the early 1970s. Passage of the Clean Water Act and the Safe Drinking Water Act led to increasingly stringent standards for municipal water and wastewater treatment facilities. The need for new and upgraded treatment plants placed enormous financial burdens on cities. In addition, the legislation provided funding for the largest public works program in the nation since the construction of the interstate highway system in the 1950s. Thousands of water and wastewater treatment plants were built with most of the cost (up to 75 percent) paid by the federal government.

While cities welcomed the federal assistance in constructing new facilities, many small communities lacked properly trained personnel to operate the plants and had difficulty attracting qualified staff. The regulatory burden led some municipalities to hire private firms to operate their water and wastewater facilities. Under operations and maintenance (O&M) contracts, cities retain ownership of the treatment facilities and control rates while the private contractor has the responsibility for day-to-day management, environmental compliance, and overall performance. O&M contracts eventually spread to medium-sized cities as the water and wastewater industries became larger and more competitive.

Since the 1980s, many large cities entered into O&M agreements for water and wastewater services. Oklahoma City was one of the first, contracting for wastewater treatment in 1984. Other large cities followed, including Houston for water treatment in 1990, New Orleans for wastewater treatment in 1992, and Indianapolis for wastewater treatment in 1994.

An initial burst in private sector activity was slowed in the mid-1980s by the passage of the Tax Reform Act of 1986. Tax benefits that had fueled the growth in private ownership and operation of water and wastewater facilities, such as rapid (5-year) depreciation and an investment tax credit equaling 10 percent of the capital investment, were all but eliminated by the law.

While the momentum slowed for private ownership of water and wastewater facilities, the market for contract operations continued to increase. Many of the initial contracts were renewed as the industry renewal rate topped 90 percent. By the mid-1990s, the water and wastewater O&M market had grown into a billion dollar business and more than 1500 water and wastewater facilities were operated by private firms.

The growth in contract public-private partnerships was spurred by two executive orders signed in the 1990s. President George Bush issued an executive order in 1992 that removed a major obstacle to public-private partnerships by enabling ownership of publicly owned facilities to be cost-effectively transferred to private firms.[1] Public agencies can sell or lease federally funded wastewater facilities and recoup the local share of the cost before payback of funds to the federal government. The payback is limited to the depreciated amount of the federal share—no payback is required if the sale price is less than the local share and the amount of accumulated depreciation.

More recently, President Bill Clinton signed an executive order in 1997 that provided a further boost to public-private partnerships for water and wastewater services.[2] Under the order, federal agencies must seek private sector participation in ownership, financing, construction, and operation of infrastructure projects such as wastewater facilities. Federal agencies were also directed to work with state and local governments to reduce any legal and regulatory barriers to private sector participation in infrastructure development.

Another boost for water and wastewater privatization occurred in 1997 when the U.S. Conference of Mayors endorsed public-private partnerships as an effective way for cities to "realize significant operational cost savings, and to attract private capital investment for needed infrastructure development."[3] The endorsement came 2 years after the conference formed the Urban Water Council to serve as a forum for local governments to share information on technology, innovative management methods, and infrastructure development.

Prior to 1997, contracts for water and wastewater O&M were limited to 5 years and required a termination clause allowing cancellation after 3 years. With such a narrow time frame, operators were restricted in their ability to secure new capital improvements and increase operating efficiencies. Short-term contracts were designed to keep vendors honest—rebidding contracts every 3 or 5 years kept continuous competitive pressures on vendors to reduce costs and improve operations.

3.2 LONG-TERM CONTRACTS

The market for water and wastewater public-private partnerships was transformed in 1997 when the Internal Revenue Service (IRS) issued new regulations that made possible long-term contracts for public utility operations while maintaining the tax-exempt status of bonds used to finance the facility. Under Revenue Procedure 97-13, private operators can enter into contracts for at least 15 years and up to 20 years for public utility properties, which include water and wastewater treatment systems.

The new IRS rules also imposed constraints on long-term contracts that are designed to prevent the abuse of tax-exempt financing. Contractors can have no interest in the net revenues of the system but may share in cost savings or revenue enhancements. These measures have led to a wave of recent contracts with gain sharing provisions that provide incentives for private operators to achieve additional cost savings.

Another incentive-based provision of the rule changes allows private operators to have 20 percent of the total contract amount to be in the form of a variable payment from cost savings or revenue sharing. These provisions create a win-win outcome for private operators and municipalities—contractors continually search for further operating efficiencies and cities benefit from additional cost savings.

The net effect of the changes in IRS rules has been to give the private sector more flexibility to meet local water and wastewater needs. Public-private partnerships can now run the gamut from a short-term O&M contract to a long-term contract that includes design, build, and operate (DBO) features, all without private ownership (Table 3.1). Long-term contracts have proven to be an effective method for addressing municipal water and wastewater needs, especially for capital improvements or in cases where new facilities are required.

3.3 EXTENT OF PRIVATIZATION

While the current extent of privatization of water and wastewater facilities is somewhat limited, recent trends suggest that more cities will be examining private sector alternatives in the future.

TABLE 3.1 The Public-Private Partnership Industry

Type	Scope	Typical term, years	Responsibility for capital improvements	Fee structure
Concession	Full service and system responsibility including billing and collection	20–30	Private company handles debt refinance, facility renewal, concession fee	Revenue from billing and collection
Lease/delegated services	Full service and system responsibility	10–20	Moderate investment by private company, facility renewal fund	Fixed fee, ASA service ROE
Design, build, operate	Design, build, operate and maintain new facility	15–25	Municipality provides capital for new facility	Fixed price for plant, fixed fee for operations
Contract operations with capital improvement	Full service O&M, CIP management	5–15	Moderate investment by private firm, facility renewal fund	Fixed fee, ASA service ROE
Contract operations	Full service O&M, plant management	3–5	Municipality	Fixed fee, ASA service
Operations assistant	Particular service, e.g., plant startup	<1	Municipality	Fixed fee
Asset transfer	Facility ownership transfer; wholesale service agreement with municipality	15–20	Private firm	Base = take or pay, variable component
BOOT	Build, own, operate new facility	15–30	Private firm	Wholesale service agreement
BOOT	Build, own, operate, transfer at specified date	15–30	Private firm	Wholesale service agreement
Purchase, renovate, operate transfer	Facility returned after plant upgrade and operations period	5–10	Private firm	Base = take or pay, variable component
Joint ownership	"Mixed company" or water authority with private service provider	30	Private and public	Percentage of additional revenue

ASA = additional services agreement; ROE = return on equity.

Source: Bob Siemak, Presentation at "Water/Wastewater" session at the Washington Institute Foundation conference, Bellevue, Wash., June 22, 1999.

3.3.1 Water

According to the Environmental Protection Agency (EPA) approximately one-third of drinking water systems nationwide are private regulated utility systems.[4] Also, 15 percent of urban water utilities are regulated investor-owned companies.[5] Of those that are publicly owned, few contract operations and maintenance with private firms. A 1997 International City/County Managers Association (ICMA) found that 5.7 percent of responding cities nationwide privatize water distribution and 3.7 percent contract water treatment.[6] Few state sources of data exist on levels of privatization. One of the few, a 1995 survey of Illinois municipalities, reported that 5.1 percent of responding cities contract for water distribution and 7.9 percent contract for water treatment.[7]

3.3.2 Wastewater

There are fewer sources of information regarding the extent of wastewater public-private partnerships. A 1994 article reported there are approximately 40 large facilities [greater than 10 million gallons per day (mgd)], 350 midsized facilities (1 to 5 mgd), and 1000 small plants under some form of contract management in the United States.[8] In the ICMA survey, 6.2 percent of responding cities have privatized wastewater collection and treatment.[9] The Illinois survey showed that 3.9 percent of municipalities reported contract operations for wastewater collection and 6.5 percent for wastewater treatment.[10]

3.3.3 Trends

Several recent reports indicate that the amount of privatization is increasing and will continue to grow in the future. A survey of selected firms in the water industry by *Public Works Financing* revealed that the municipal contract operations market increased 16 percent in 2000.[11] Although the rate of increase was down from the previous year, it still indicates an upward trend in water privatization. Contract operations have been increasing as much as 25 percent annually.[12] In 2000, Moody's Investor Service also predicted more privatization, saying that "public policymakers will turn to the private sector for financial, technical, and operating assistance when the municipal water system receives reliable and reasonably priced services." An analysis of ICMA survey data shows that water distribution was one of the municipal services with the largest increase in privatization between 1988 and 1997.[13] Finally, a 1995 Illinois survey of cities of more than 5000 showed that wastewater treatment is among the services most likely to be considered for privatization in the future.[14]

Since the new IRS regulations were announced in 1997, numerous cities, large and small, have entered into long-term water and wastewater contracts. The size of municipalities with long-term contracts ranges from major cities, such as Atlanta, to small towns, such as Port Byron, Illinois (pop. 1350). In the first 2 years after the regulation went into effect, more than 80 cities began the competitive process for contracts with initial terms of more than 10 years. During the same period, 45 municipalities agreed to O&M contracts of more than 10 years.

The trend of slow but steady growth in the number of long-term water and wastewater contracts continues into the 2000s. In 2000, another 25 cities entered into long-term contracts at least 10 years in length. Table 3.2 lists some of the water and wastewater public-private partnerships currently in existence across the United States.

3.4 FACTORS CAUSING WATER AND WASTEWATER PARTNERSHIPS

While philosophical or ideological factors formerly played a role in officials' decisions to privatize, more practical considerations have emerged in recent years that have increased the attractiveness of privatization. There are a variety of factors forcing local officials to consider private sector alternatives.

3.4.1 Cost Savings

According to the U.S. Conference of Mayors, expenditures for water and wastewater services are among the largest facing local governments today.[15] Thus, there are more opportunities for cost savings from public-private partnerships. Water companies can utilize advanced technology, more flexible management practices, and streamlined procurement and construction practices to lower costs. In addition, larger firms that operate several facilities can use economies of scale to achieve better prices for chemicals, capital equipment, and supplies. For example, energy costs, which can comprise approximately two-thirds of a water utility's budget, are one area where efficiencies can be gained in a short period of time.[16]

The White River Environmental Partnership between the city of Indianapolis and United Water provides an example of some of the cost savings possible through a public-private partnership. Officials with the partnership estimated that the contract would reduce utility costs by 20 percent through engineering process control and improved performance; lower personnel costs by 30 percent through better training, streamlined management, and lower overhead costs; and reduce maintenance costs by 30 percent through increased preventive maintenance and bulk purchasing.

Private firms can not only generate significant O&M savings, they can also reduce capital costs between 10 to 50 percent through design and build techniques rather than the traditional design, bid, and build approach used by many munici-

TABLE 3.2 Selected Cities With Water and Wastewater Public-Private Partnerships

City	State	Description	Length (years)	Type
Athens	N.Y.	Water/wastewater	10	O&M
Atlanta	Ga.	Water	20	
Atwater	Calif.	Wastewater	15	O&M
Bartlesville	Okla.	Wastewater	10	
Bessemer	Ala.	Water	20	DBO
Beverly Hills	Calif.	Water	20	DBO
Black River Falls	Wis.	Wastewater	5	O&M
Boonville	Ind.	Wastewater	10	
Bowling Green	Mo.	Water/wastewater	10	
Brockton	Mass.	Water/wastewater	20	
Buffalo	N.Y.	Water	10	
Chester Borough	N.J.	Wastewater	20	
Cranston	R.I.	Wastewater	25	DBO
Danbury	Conn.	Wastewater	20	
Edison	N.J.	Water	20	
Evansville	Ind.	Water	10	
Floyd County	Ky.	Wastewater	20	DBO
Franklin	Ohio	Water	20	BOT
Freeport	Tex.	Water/wastewater	15	
Fulton County	Ga.	Wastewater	10	O&M
Gladewater	Tex.	Water/wastewater	10	
Glen Falls	N.Y.	Water/wastewater	20	O&M
Hoboken	N.J.	Wastewater	20	
Indianapolis	Ind.	Wastewater	10	
Jackson	Ala.	Water/wastewater	10	
Kenner	La.	Wastewater	10	
Manalapan	N.J.	Water	20	
Milwaukee	Wis.	Wastewater	10	
Monmouth	Ill.	Water/wastewater	10	
Mount Vernon	Ill.	Wastewater	20	
New Haven	Conn.	Wastewater	15	
Newport	R.I.	Wastewater	20	DBO
No. Adams	Mass.	Water	10	O&M
No. Brunswick	N.J.	Water	20	
Norwalk	Conn.	Wastewater	20	O&M
Pine River	Minn.	Wastewater	10	
Plymouth	N.C.	Water/Wastewater	10	

TABLE 3.2 (*Continued*)

City	State	Description	Length (years)	Type
Quincy	Wash.	Wastewater	20	DBO
Reidsville	N.C.	Wastewater	10	O&M
Rockland	Mass.	Wastewater	10	
Seattle	Wash.	Water	25	DBO
Springfield	Mass.	Wastewater	20	O&M
Tampa Bay	Fla.	Wastewater	20	DBO
Toronto	Ohio	Wastewater	10	
W. Melbourne	Fla.	Wastewater	10	
Wildwood	N.J.	Water	20	O&M
Wilmington	Del.	Wastewater	20	

Sources: *Public Works Financing,* March 1998, p. 5, and March 2001, pp. 8–9.

palities.[17] For example, Seattle used a design, build, and operate approach to build a new water treatment facility and saved more than 40 percent.[18]

Long-term contracts also impact the amount of savings generated through public-private partnerships. Private firms can better manage costs over a longer period of time and amortize up front investments in advanced technology and computerization. Some recent long-term water and wastewater O&M contracts have generated estimated savings ranging from 20 to 50 percent (Table 3.3).

3.4.2 Infrastructure Needs

Many water and wastewater systems were built with federal funds during the 1970s and need upgrades and replacement. Some systems include water and sewer infrastructures that are even older, some dating back to the early 1900s. In parts of St. Louis, for example, the wastewater system dates back to before the Civil War.[19] For other cities, such as Las Vegas, rapid economic growth is fueling the need for public-private partnerships. Congress has reduced grant funding for infrastructure replacement, and states offer only low-interest revolving loan funds.

The massive costs of replacing and maintaining water and wastewater infrastructure will necessitate an examination of public-private partnerships. The U.S. EPA recently estimated that the nation's 76,000 drinking water systems alone will require $150 billion in investments over the next 20 years.[20] The American Water Works Association (AWWA) examined water utilities in 20 large cities and estimated that $250 billion is needed over the next 30 years to replace aging drinking water mains, valves, and fittings.[21] Wastewater systems will require nearly $140 billion in the next 15 years to meet new water pollution rules, according to the EPA.[22]

TABLE 3.3 Estimated Cost Savings from Water and Wastewater Public-Private Partnerships

Municipality	System type	Contract term, years	Estimated savings, %
Atlanta, Ga.	Water	20	45
Franklin, Ohio	BOT wastewater	20	23
Franklin, Ohio	BOT water	20	30
Milwaukee, Wis.	Wastewater	10	30
New Haven, Conn.	Wastewater	15	30
Newport, R.I.	Wastewater	20	24
Plymouth, Mass.	DBO wastewater	20	19.7
Seattle, Wash.	DBO water	25	40
Tampa, Fla.	DBO water	15 + 5	21
Tampa, Fla.	DBOOT desalination	30	50

BOT = build, operate, transfer; DBOOT = design, build, own, operate, transfer.
Source: *Public Works Financing,* January 2001.

Cities facing financial limitations and citizen opposition to tax increases will find it difficult to finance water and wastewater improvements internally. Public-private partnerships, offering significant cost savings, operational improvements, and innovative financing schemes, will be an increasingly attractive option for urban officials.

3.4.3 Unfunded Mandates

While reducing its contributions to local water systems over the past 30 years, the federal government imposes strict water quality and effluent standards under the Clean Water Act and Safe Drinking Water Act. Unfunded mandates force municipal systems to meet federal regulations through local sources of revenues or state revolving loan funds. The EPA continues to toughen monitoring requirements, develop stricter contaminant removal standards, and more vigorously pursue compliance, resulting in higher costs for municipalities.

For example, AWWA estimates that the cost to local governments of meeting toughened arsenic standards is approximately $14 billion in capital investments nationwide and $1.5 billion in additional annual operating costs.[23] Nearly all the communities affected serve less than 10,000 people, putting further strain on limited local budgets.[24] Enforcement of a lower standard on radium levels in drinking water could impact more than 100 communities in Illinois.

In most public-private partnerships, the regulatory burden for meeting environmental mandates falls to the private firm. Compliance with EPA standards is often guaranteed by private firms through performance-based contracts. Cities such as Atlanta and New Orleans that faced possible EPA sanctions were able to come into full compliance through public-private partnerships. A major benefit cited by some city officials is the peace of mind of knowing that their private partners have responsibility for meeting compliance standards.[25]

3.4.4 Improved Performance

Along with meeting stricter regulatory standards, private firms can improve overall system performance and quality. Many cities are turning away from the traditional low-bid approach and seeking the "best value for the money" through advanced bidding procedures. Atlanta officials used a two-tiered bidding process that included a "best and final offer."

City officials are also using performance-based contracting to ensure optimum performance. Milwaukee's incentive-laden contract for wastewater treatment is a prime example—the contract set the permitted effluent discharge levels well below the levels permitted by state regulators. Performance exceeded even the more stringent level, earning the firm two $50,000 bonuses thus far.

In addition, private firms often invest in new technologies and computerization at water and wastewater facilities, expenditures that many local governments find difficult to make.[26] Private contractors also provide increased training opportunities for employees, another area where municipal budget-cutters look for reducing expenditures. Preventive maintenance plans that are common in many partnerships help to lengthen the useful life of existing assets and defer costs into the future.

3.4.5 Lack of Political Will

It can be difficult for local officials to make the necessary investments in community water systems. Water pipes and sewer mains are not visible and thus are easier for elected officials to ignore compared with expenditures for police and fire services. In addition, in many municipalities, water and sewer rates do not adequately cover the actual cost of providing services. Raising water and sewer rates to cover operations and maintenance, as well as capital replacement, is a risky move for elected officials.

Water and wastewater rates may increase under public-private partnerships but not at the rate they would under municipal control due to operational efficiencies and savings. In January 1997 Atlanta Mayor Bill Campbell estimated that his city would have been forced to raise water rates 81 percent under continued municipal management. As a result of the city's partnership for water treatment, the city's rates will rise on a blended scale of only 8 percent over 4 years.[27]

3.5 OBSTACLES TO PARTNERSHIPS

In spite of numerous successful public-private partnerships for water and wastewater services, many cities choose to keep operations in-house. It is often difficult to implement changes in municipal operations, especially when they affect people's livelihoods. As more cities gain positive experiences with public-private partnerships, however, many of these apprehensions may be overcome and residents, elected officials, and employees will become more comfortable with the concept.

3.5.1 Employee Opposition

A major concern facing municipal officials when contracting services is opposition from public employees who fear loss of employment, lower wages, or reduced safety standards. Given the labor intensity of many public services, some cost savings are likely to accrue from reduced personnel costs. Increased productivity in water and wastewater plants can arise from more efficient use of advanced technology and computerization that may necessitate fewer employees. While this ultimately means that more work is accomplished with fewer employees, it does not mean that current workers are readily dismissed.

Surveys of city officials confirm the existence of employee opposition to privatization. In a 1997 ICMA survey, opposition from line employees was the leading obstacle to privatization according to 63.5 percent of survey respondents.[28] If employees are unionized, the opposition to privatization is likely to be even more intense.

Much research exists to suggest that there is little or no impact on employees as a result of privatizing government services in general. In many cases, few layoffs occur and salaries and benefits are comparable to or better than previous salaries and benefits.[29] That information has not prevented misinformation from injecting itself into debates over privatization.

There is little data from specific industries on the effects of privatization on employees. For water and wastewater services, where technical capabilities are important, keeping some or all of existing staff is often in the best interests of the private firm. In most recent long-term contracts for water and wastewater services, the current workforce is retained by the private firm because of the benefits of having experienced employees. Also, massive reductions in the workforce can create hostility and opposition in the community.

If fewer employees are needed to efficiently operate a water or wastewater facility, many firms opt to reduce the workforce through attrition, which occurs when the private firm does not fill a vacancy created when an employee retires or voluntarily terminates employment. Recent long-term contracts for water and wastewater in Atlanta and Milwaukee, respectively, featured no-layoff clauses and reductions in the workforce only through attrition. After entering into a long-term partnership for wastewater services, Indianapolis reduced its workforce from 321 to 196 through attrition.[30]

City officials can take other approaches as well to avoid negative impacts on employees. Employees can be transferred to other departments within the city or be offered positions with the private contractor. In addition, cities can offer early retirement incentives to reduce the number of employees.

Another option for city officials is to enter into a management contract for water and/or wastewater services. Under such agreements, the private firm takes over responsibility for day-to-day management of operations while the existing workforce remains city employees. Buffalo is the largest city in the United States to operate under a 1997 management contract for water services. The water division's 160 employees remained in-house and were guaranteed jobs for 5 years. Employees are paid by the city, and the private firm reimburses the city for employee costs.

3.5.2 Loss of Control

In addition to opposition by public employees, elected officials and public administrators may oppose public-private partnerships because they fear a loss of control over daily water and wastewater operations. Elected officials may sense that private partners will take over many aspects of public policy decision making, especially rate making, and supervisory personnel perceive a threat to their livelihoods.

In practice, most of these concerns are overcome with effective performance monitoring and oversight policies included in the contract, but some officials remain unconvinced. For example, in most water and wastewater contracts, control over water and sewer rates remains with the municipality. In Camden, New Jersey, for example, the city maintained control over water and sewer rates, fees, and capital improvements. The current 20-year agreement with U.S. Water is in its third year and has increased the city's water and sewer system cash flow by nearly \$11 million. In addition, the firm increased water and sewer collections by nearly \$8 million over the amount the city had previously collected and brought the city back into regulatory compliance.[31]

More contracts are performance driven and guarantee service quality. Contracts often have penalties for nonperformance, and some provide financial incentives for superior performance. Private firms must meet environmental permit standards or face possible financial penalties.

With contractual guarantees on service quality and costs, most municipalities have benefited from public-private partnerships. The key is ensuring a contract that is written to include performance-based incentives and provides for effective monitoring and oversight by the municipality. With the proper contract provisions in place, local control can actually be strengthened through public-private partnerships.[32]

An emerging strategy used by opponents who fear a loss of control is to advocate prior approval by voters of any partnership agreements. Opponents of privatization in Omaha placed a measure on the ballot that would require public approval of any proposal to privatize the management of operation of the city's two wastewater treatment facilities.[33] The referendum passed with two-thirds of voters approving. In New Orleans, opponents floated a similar measure in an attempt to prevent privatization of the city's wastewater facilities.

3.6 THE SELECTION PROCESS

The decision to enter into a partnership agreement with a private firm for water services is a time-consuming and highly technical process. Many large cities first go through an internal evaluation to determine if reengineering and restructuring can achieve better performance and lower costs. Atlanta developed a two-step process for its water partnership that first assessed the city's water and wastewater systems and analyzed various options including continued reengineering, outsourcing, and privatization.

Once the decision is made to outsource operations, city officials begin a highly technical, and sometimes laborious, process to select a contractor. After 7 years of research and hundreds of thousands of dollars spent on studies, San Jose, California, officials voted in December 2000, to solicit bids for managing the city's water system.[34] One key element to engaging in such a process is to have in-house personnel with experience in contract development and oversight or securing such expertise from consultants.

3.6.1 Performance-based Selection

The process of selecting a private contractor for water and wastewater partnerships is evolving into a performance-based system. Municipalities are moving beyond price-based proposals to "best value" contracts that combine cost and quality factors. Low-bid alternatives are increasingly common for long-term contracts where municipalities rank potential private partners on a variety of measures instead of solely on the basis of lowest cost.

City officials are realizing that there are times when they will get more if they pay more—that best value is not necessarily always the least expensive. The concept of selecting firms to provide complex services or projects should be based on qualifications and technical merits, as long as the price is a value for what is promised, is becoming mainstream.[35] The Federal Acquisitions Regulations were amended in 1996 (FAR 2.101) to allow best-value source selections in government outsourcings. Federal Acquisitions Regulations define *best value* as "the expected outcome of an acquisition…providing the greatest overall benefit in response to the requirement." And the American Bar Association's revised Model Procurement Code incorporates best-value procurements as the standard.[36]

Traditionally, municipalities would develop specifications for their treatment facility's needs and draft a *request for proposals* (RFP). Requests for proposals define the scope of services to be provided by the private firm, performance requirements, roles and responsibilities of the partners, and other key elements. The municipality would then choose a firm that submitted the lowest bid. This process became outdated as it became increasingly difficult to write RFPs that cover all the possible variables involved in operating water and wastewater systems, especially for larger and more complex facilities.

3.7 RFQs AND RFPs

More municipalities are using *requests for qualifications* (RFQs) to identify the best private firm for the particular project. As with any public-private partnership, the key is selecting a capable and experienced contractor that has worked with utilities of similar size, scope, budget, and complexity. Cities may consider issuing an RFQ to ensure that only experienced, qualified contractors seek the project. A private firm is typically asked for information relating to its history, management structure, previous municipal projects, technical expertise, financial information, and quality assurance programs.

Requests for qualifications give private firms greater flexibility to tailor their proposals to the specific needs of the municipality. For example, a firm may offer funding for capital or infrastructure replacements that can improve efficiency and lower operating costs. Creative strategies to enhance performance and lower costs can produce a win-win outcome for both parties to the agreement.

City officials shape RFPs to reflect their priorities and to maintain control over the process. For example, cities that seek to avoid layoffs include provisions that ensure proper treatment of current employees. Atlanta officials inserted language stipulating that the winning firm must retain existing employees for at least 3 years.[37] Requests for proposals also include performance incentives and penalties based on water and wastewater service performance.

Municipalities can either choose a contractor at this stage or proceed to solicit a "best and final offer" from bidders. Atlanta chose the latter approach and developed a 70-point scoring system for analyzing and comparing the proposals. The value index balanced cost savings potential with technical and quality considerations. United Water reduced its price by nearly 21 percent in the best and final offer proposal stage and won the competition.

3.7.1 Contract Specifications

The next step in the process is developing a contract that clearly states the roles and responsibilities of both parties to the agreement. It should specify the scope of services to be provided and the facilities to be managed. The following is a list of major features that are critical to the success of a public-private partnership for water and wastewater services:

1. *Compliance and performance standards.* Most city officials use public-private partnerships as a way to reduce their liability for compliance with environmental regulations. Establishing clear responsibility for compliance and liability for nonperformance are crucial items to be negotiated. Indemnification provisions should also be included that define and cover any claims, liabilities, and losses that may arise. Other important compliance issues to be addressed include responsibility for permits, construction activities, operations, design and technology, economic and financial risks, and force majeure (casualty and business interruption).[38]

2. *Contract term and pricing.* The length of a public-private partnership depends on the needs of the municipality. More cities are entering into long-term water and wastewater partnerships as a result of the 1997 IRS regulations. Under those rules, pricing and contract length are tied together under a concept known as periodic fixed fee (PFF). The more compensation is based on a fixed fee, the longer is the allowable contract term. Most contracts have fixed price guarantees for the length of the contract and annual increases limited to changes in the Consumer Price Index (CPI). The contract should also include language on renewal options.

3. *Employees.* Contract language should carefully state how existing employees are to transition to private employment, if applicable. Many cities include language that mandates the hiring of the current workforce, ensures comparable wages and benefits, and requires the private firm to bargain in good faith if the employees are unionized.

4. *Equipment.* Contracts typically delineate the operator's maintenance responsibilities including all preventive and predictive maintenance functions. Municipalities often require the contract operator to submit an annual maintenance and capital improvements budget for review and approval. Milwaukee officials developed specifications for addressing capital repair and replacement that clearly define the responsibilities of the private firm (United Water) and the sewerage district. For example, the parties agreed that all maintenance would be done by United Water as part of its operating fee and that the firm would pay the first $5000 for each capital repair or replacement item as an incentive to do adequate maintenance.[39]

5. *Inspection, reporting, and review.* City officials should require monthly and annual operating reports in sufficient detail to enable an evaluation of the contractor's performance under the terms of the agreement. Other information that is often required for review includes maintenance reports; annual operating, capital, and maintenance budgets; and reporting procedures. City officials should also have the right to inspect the facilities and audit the contractors' records upon reasonable notice.

6. *Insurance.* The contract language should include a requirement that contractors carry general liability, automotive liability, and workers' compensation insurance. In addition, the firm should obtain insurance with the municipality named as an additional insured party.

7. *Termination.* Many contracts include termination clauses that require prior notice. This provision can serve to calm opponents of partnerships who fear being locked into a contract without recourse.

The development of a mutually acceptable contract, implementation of the agreement, and careful monitoring and oversight by municipalities are crucial aspects of the contracting process. Both sides must bargain in good faith and compromise to ensure a win-win outcome.

3.8 CASE STUDIES

Recently, some of the nation's largest cities and several small communities have conducted evaluations of water and wastewater privatizations. As the results of these case studies demonstrate, public-private partnerships have succeeded in many ways beyond just improving performance and lowering costs.

3.8.1 Indianapolis

A late-1999 report by the city of Indianapolis examined the success of the White River Environmental Partnership (WREP) in running the city's sewer collection system and wastewater treatment plants since 1994. The report measured performance in three crucial areas:

- *Employee treatment.* Employee wages and benefits have risen between 9 and 28 percent, accident rates have dropped 91 percent, and grievances are down 99 percent.
- *Environmental compliance.* WREP has improved on the city's record of environmental compliance in exceeded permits and effluent discharges.
- *Cost savings.* Over 5 years, privatization saved the city $78 million—surpassing the expected savings of $65 million.

In 1997, after 3 years of contract performance that exceeded expectations, the city decided to replace the existing 5-year contract with a new 10-year contract extending through 2007. Total savings from the contracts from 1994 to 2007 are expected to total $250 million. To date, the city has used most of the savings for capital improvements in the sewer system and treatment facilities and for rate reductions.

3.8.2 Milwaukee

In March 2000, Milwaukee released a second-year evaluation of its 10-year contract with United Water to operate the city's sewage collection system and wastewater treatment plants. For the second year, United Water exceeded the operating standards of the contract. Meanwhile, workplace injuries, sick days, and grievances remain at levels less than half those experienced under city management.

3.8.3 Atlanta

More recently, Atlanta released the results of an audit of the first 18 months of its 20-year contract with United Water to run the city's water utility. Among the key findings is the increasingly common approach taken by water firms to address employee issues. United Water employed 417 of the 535 employees at the utility when it assumed operations in January 1999. Since then, 49 quit, 14 were terminated for cause, and 4 were transferred to other cities in the region. The firm signed an agreement with the American Federation of State, County, and Municipal Employees (AFSCME) that unionized 95 percent of the workers—a first in "right-to-work" Georgia. Union members received benefits equal to or better than their former city packages and a 3 percent initial pay raise.

3.8.4 Small Communities

Privatization has succeeded not just in large cities, but in smaller communities as well. A public-private partnership in Mount Vernon, Illinois, not only saved money and improved performance, it led to expanded economic growth for the city of 17,000. In the mid-1980s, Mount Vernon was under a sewer connection ban because of compliance problems at its wastewater treatment plant, meaning the city could not accept any more sanitary sewer customers and was unable to attract or expand industry.

The city entered into a 20-year service partnership with Environmental Management Corporation to design, build, and operate an upgraded and expanded wastewater treatment facility. Sewer restrictions were lifted after the first phase of construction was completed. The agreement is guaranteed to meet EPA effluent standards and, in fact, led to the wastewater system operating significantly better than all EPA permit limitations. In addition, the agreement saved the city approximately $3 million in tax dollars and was completed in substantially less time than alternate proposals.

The impact on economic development was impressive. Within 18 months of the first phase of construction, the city attracted approximately $300 million in private investment.

Monmouth, Illinois, a city of 10,000, privatized its water and wastewater services as part of a contract with a firm to operate all public works services. The agreement saved the city approximately $300,000 (nearly 20 percent), improved the quality of services, and was a key factor in the city's recovery from severe financial problems. In addition, union employees endorsed the agreement before final city approval and have benefited from a better compensation package than what would have been available from the city.

3.9 SUMMARY

With water issues emerging as key elements of cities' long-term growth and economic competitiveness, officials are looking for cost-effective ways to provide clean drinking water and treated wastewater. Numerous cities have joined forces with the private sector not only to provide water services, but also to help rebuild a crucial element of urban infrastructure.

As more cities achieve success in partnering with the private sector, more will do so as they learn from the experiences of their peers. Already, major U.S. cities, such as Milwaukee, Houston, Atlanta, Seattle, and Indianapolis, have realized significant cost savings and improved performance through public-private partnerships for water services.

Public-private partnerships for water and wastewater stand as major examples of urban innovation at the dawn of the new century. As water issues increase in importance in the coming years, more cities will depend on partnerships as a way to strengthen their communities and grow in the future.

3.10 REFERENCES

1. Executive Order No. 12803 (May 4, 1992).
2. Executive Order No. 12893 (January 26, 1994), *http://www.nara.gov/fedreg/executive_orders_pdf/eo*12893*.pdf*.
3. United States Conference of Mayors Urban Water Council, *Public/Private Partnerships in Municipal Water and Wastewater Systems, Case Studies of Selected Cities,* September 25, 1997, *http://www.usmayors.org/uscm/urbanwater/ documents/casestud.htm.*

4. United States Environmental Protection Agency, *Community Water System Survey,* 1997, *www.epa.gov/safewater/cwsreport.pdf.*
5. Adrian T. Moore, "Clearing Muddy Waters: Private Water Utilities Lower Costs and Improve Services," *Reason Public Policy Institute Policy Brief No. 3,* July 1997, p. 1.
6. Author analysis of ICMA's *Profile of Local Government Service Delivery Choices* Survey 1997.
7. Robin A. Johnson and Norman Walzer, "Competition for City Services: Has the Time Arrived?" Springfield, Ill: Office of the Comptroller, 1996, p. 11.
8. W. A. Peterson, "Privatization at a Crossroads," *Water Environment and Technology,* November 1994, p. 57.
9. Author analysis of ICMA's *Profile of Local Government Service Delivery Choices* Survey 1997.
10. Robin A. Johnson and Norman Walzer, "Competition for City Services: Has the Time Arrived?" Springfield, Ill: Office of the Comptroller, 1996, p. 11.
11. William G. Reinhardt, "U.S. Water/Wastewater Contract Services Show 16% Growth to $1.7 Billion in 2000," *Public Works Financing,* March 2001, p. 1.
12. Steve Maxwell, "Where Are We Headed? Ten Key Trends and Developments in the Water Industry," *American Water Works Association Journal,* April 2001, p. 116.
13. Robin A. Johnson and Norman Walzer, "Privatization and Managed Competition: Management Fad or Long-term Systematic Change for Cities?" in Robin A. Johnson et al. (eds.), *Local Government Innovation,* Quorum Books, 2000, p. 172.
14. Robin A. Johnson and Norman Walzer, "Cities Continue to Increase the Extent and Scope of Contracting for Services," *Illinois Municipal Review,* p. 14 September 2000.
15. John Holusha, "In Latest Cost-cutting Trend, Cities Turn to Private Sector for Waste-Water Treatment," *New York Times,* May 5, 1996, p. 14.
16. Mark Crapeau, "Changing the Face of the Water Industry," *American Water Works Association Journal,* November 2000, p. 65.
17. Douglas Herbst and David Seader, "Providing Public Services Through Long-term Service Agreements," in Robin A. Johnson et al. (eds.) *Local Government Innovation,* Quorum Books, 2000, p. 114.
18. William G. Reinhardt, "Case Study: Seattle, Wash.," *Public Works Financing,* March 2001, p. 11.
19. Stephen J. Hoffman, "Cure for National Infrastructure Crisis Not Yet on the Horizon," *Water Investment Newsletter,* June 15, 2001, p. 1.
20. United States Environmental Protection Agency, *Drinking Water Infrastructure Needs Survey,* 2001, *www.epa.gov/safewater/needs.htm.*
21. American Water Works Association, *Dawn of the Replacement Era: Reinvesting in Drinking Water Infrastructure, www.awwa.org/govtaff/infrastructure.pdf.*
22. United States Environmental Protection Agency, Clean Water Needs Survey, 1996, *www.epa.gov/cwm/toc.htm.*
23. Tom Bell, "Opposition Mounts to Arsenic Rules," *Environment and Climate News,* The Heartland Institute, *www.Heartland.org,* January 2001.
24. Water Tech Online, "U.S. Arsenic Standard to Be Lowered," November 1, 2001, *www.watetechonline.com/news.asp?mode = 4&N_ID = 26866.*
25. Janet Ward, "The Pros and Cons of Long-term Privatization," *American City & County,* May 1998, p. 63.

26. James F. Eiden, "Privatization of Wastewater Treatment Systems: Headache or Cure?" *Keystone Water Quality Manager,* September–October 1994, p. 38.
27. Bill Campbell, "A Bold, Innovative Approach to Privatization," in Robin A. Johnson et al. (eds.), *Local Government Innovation,* Quorum Books, 2000.
28. Robin A. Johnson and Norman Walzer, "Privatization and Managed Competition: Management Fad or Long-term Systematic Change for Cities?" in Robin A. Johnson et al. (eds.), *Local Government Innovation,* Quorum Books, 2000.
29. Janet Ward, p. 66. James F. Eiden, p. 37. Robin A. Johnson and Norman Walzer, "Competition for City Services: Has the Time Arrived?", Springfield, Ill: Office of the Comptroller, 1996, pp. 28–29.
30. Stephen Goldsmith, *The Twenty-first Century City,* Washington, D.C.: Regnery Publishing Co., 1997.
31. William Reinhardt, "Case Study: Camden, N.J.," *Public Works Financing,* March 2001, p. 7.
32. Adrian Moore and Wade Hudson, "The Evolution of Privatization Practices and Strategies," in Robin A. Johnson et al. (eds.), *Local Government Innovation,* Quorum Books, 2000, p. 21.
33. Robin A. Johnson, "Election Produces Mixed Results for Government Reform," *Privatization Watch,* January 2001, p. 1.
34. Robin A. Johnson, "Cities Tread Cautiously on Water Privatization," *Privatization Watch,* February 2001, p. 3.
35. Associated Soil and Foundation Engineers (ASFE), *Establishing the Cost of Public-sector Design,* A.S.F.E. White Paper No. 2, Silver Spring, Md.: ASFE, 1999, p. 14.
36. *web.mit.edu/civenv/idr/MPCPage.htm.*
37. Bill Campbell, "A Bold, Innovative Approach to Privatization," in Robin A. Johnson et al. (eds.), *Local Government Innovation,* Quorum Books, 2000.
38. Janine Landow-Esser and Melissa Manuel, "Drafting the Best Possible Outsourcing Contract," *American City & County,* October 1997, p. 72.
39. Anne Spray Kinney, "Milwaukee's O&M Contract Protects Its Public Assets," *Public Works Financing,* June 2001, p. 2.

CHAPTER 4
WATER-USE MANAGEMENT: PERMIT AND WATER-TRANSFER SYSTEMS*

J. Wayland Eheart
Department of Civil Engineering
University of Illinois at Urbana-Champaign
Urbana, Illinois

Jay R. Lund
Department of Civil and Environmental Engineering
University of California, Davis
Davis, California

4.1 INTRODUCTION

This chapter reviews two approaches to managing the use of water, *water-use permit systems* and *water-transfer* or *water-marketing systems*. Permit systems typically imply a level of central administration and control, while transferable-use systems require a level of decentralized management. The two systems often operate in tandem, since transferable water-use systems usually require the firm establishment of property rights provided by a permit system.

The first sections of the chapter focus exclusively, if somewhat arbitrarily, on water-withdrawal permits programs. It is divided into nine sections following this introduction. Section 4.2 reviews the legal foundations upon which water regulations are or might be based. Section 4.3 reviews various types of permit systems. In Sec. 4.4 program objectives are introduced and discussed. Section 4.5 discusses options and decisions for management programs and associated

* This chapter is a slightly modified version of Chap. 32 of *Water Resources Handbook*, edited by W. Mays (McGraw-Hill, New York, 1996).

tradeoffs among the program objectives. Section 4.6 presents some technical aspects that must be addressed for any system of transferable or nontransferable permits. The review of water-transfer systems begins with Sec. 4.7. This discussion focuses primarily on the transfer of water to cities, but the principles remain applicable to other contexts of water transfers. This section contains some historical and economic background on the subject. Section 4.8 reviews various forms of water-transfer arrangements. Section 4.9 discusses various issues for implementing water transfers, including the role of government in supporting and regulating water transfers. Section 4.10 summarizes the discussion and implications of all the preceding sections and offers some concluding remarks on water-use management.

It should be noted that there are other forms of water-use management which are beyond the scope of this chapter. These include the use of plumbing codes, land-use regulation and incentives, and general water-conservation regulations and incentives now common in many cities in the western United States.

4.2 FOUNDATIONS OF WATER-WITHDRAWAL REGULATIONS

In the so-called humid regions of the world, rainfalls and streamflows have historically been sufficient to supply nearly all human needs. In those regions a set of common-law precedents known as the *riparian doctrine* has often evolved to govern surface-water use.

In the arid regions of the western United States, a different type of water-use doctrine evolved in recognition that in arid or semiarid environments the worth of land was intimately tied to the amount of water that could be used with it. The *appropriative doctrine* that was developed in that region grants each water user a clear right to a certain quantity of water. Associated with that right is a priority rank relative to all other users on the same stream that establishes the order of forfeiture of water in times of shortage.

The riparian doctrine is rather imprecise compared to the appropriative doctrine, and many riparian areas lack a strong, comprehensive set of water-use regulations (Linsley and Franzini, 1972; Dixon and Cox, 1985). In the days when the technical capacity to use water was limited and abstractions from streams were needed only for a few cattle or a small vegetable plot, the riparian doctrine functioned adequately. In recent years, however, problems of concentrated and massive use have become more common and severe in humid regions. Aquifer levels and aquatic stream habitats have been threatened by concentrated withdrawals as cities have expanded, and irrigation, which is highly consumptive, has increased (e.g., Goering and Cekay, 1988).

4.3 WATER-USE PERMIT SYSTEMS

With the recognition that the common-law riparian doctrine is ill-suited to manage water withdrawals in the current era, many riparian areas have begun to pass more detailed water-use laws (e.g., the 1972 Florida Water Resources Act, Florida Statutes, Chap. 373; Virginia Water Use Act, 1991).

While such water laws constitute great progress in setting the general principles under which regulatory programs must operate, they have generally not attained the specificity necessary for day-to-day regulation of water withdrawals. These tasks are usually implicitly or explicitly left to local regulatory agencies to implement as they best see fit. Saarinen and Lynn (1993), for example, discuss the problems of achieving economic efficiency under Florida's statute, even though such efficiency is a stated goal of the statute. While discussion of comprehensive management programs has occurred (e.g., Mack and Peralta, 1987; Walker et al., 1983), implementation has not been widespread.

There is thus a need for state, provincial, regional, or basin management agencies in traditionally riparian areas to develop day-to-day administrative systems or water-management programs. Ideally, such programs should routinely regulate the withdrawal of water efficiently (in both the bureaucratic and economic sense) and fairly and involve the courts only in exceptional cases such as interstate or international disputes.*

There is a corresponding need, albeit a lesser one, for a review of administrative systems in areas traditionally governed by the appropriative doctrine. In many such cases, ad hoc and even inconsistent procedures have become ensconced, so that revisions are appropriate.

In essence, such programs should have two overall purposes, although they may have several competing objectives. First, they should serve as water-allocation instruments in times of drought. We define drought, in this context, as an absence or shortage of water *when the availability of more is expected*. The lack of water in Death Valley, California is not a drought, as no one expects Death Valley to be a verdant paradise. The existence of water users, their water-using capacity, and their wants define the drought as much as the water shortage (WSTB, 1986). The second overall purpose of such programs is to prevent problems attending concentrated withdrawals. Such concentrated uses may cause problems with maintenance of minimum streamflows (i.e., aquatic habitats) and aquifer levels in times of normal rainfall as well as drought.

There are many options for such programs and a useful precursor to their development is an assessment and comparison of these options, with a view toward providing guidance to agencies engaged in their development. This chapter provides

* Work is ongoing to develop methods to effect water sharing among states in a smooth manner.

such a comparison. Alternative decisions involved in creating a water-withdrawal permit program are compared qualitatively in the context of six management objectives. Certain alternatives are identified as currently appearing more attractive than others, pending further research and public dialogue. No specific recommendations are made but an attempt is made to provide guidance to decision makers without imposing values on them.

4.3.1 Water-Management Programs

It is assumed here that a water-management agency (referred to as the *agency*) has been charged with developing and administering a water-withdrawal management program (although these two activities need not be undertaken by the same entity). The program is to consist of a set of uniform rules governing water withdrawal. A set of target minimum streamflows and aquifer levels (or allowable drawdown rates) is assumed already set. These external requirements may vary with time, monotonically changing or oscillating on an annual frequency. It is also assumed that the agency has been given statutory authority to impose any of the management programs discussed here on raw-water users. Obviously, if the statute precludes any of these options, they must be excluded from consideration.

There are four major philosophical approaches toward restricting water use, discussed as follows. They are not all mutually exclusive and may be combined in a program. Most programs will probably concentrate on one approach, possibly with a second filling in the gaps.

1. *Ad hoc restrictions.* This is essentially the absence of a management program. These restrictions simply disallow certain types of uses under certain conditions. Examples of such restrictions are the classical municipal prohibitions of lawn and garden watering and car washing during times of shortage. On a larger scale, one might envision restrictions on agricultural or industrial use in times of drought. Such emergency restrictions may be satisfactory in cases where supply shortfalls are brief and infrequent, but are ill-suited for long-term maintenance of water resources once the crisis has passed. They often are not economically efficient. Water restrictions may be imposed just when they can be accommodated least easily (e.g., restrictions on irrigation just when crops are in greatest need of water). WSTB (1986) presents the results of a workshop devoted to a thorough analysis of the problems and options of ad hoc restrictions.

2. *Water charges.* A second approach is the requirement that users pay a charge for using water. This is not simply a registration fee; it must be proportional to the amount of water used and high enough to induce conservation. Under such a program, installation or inspection of flow meters, monitoring of flows (to establish the required payments), and collection of payments will be necessary agency activities. A strength of such a program is that it tends to induce an economically efficient use of water. An important weakness, however, is that the charge must constantly be readjusted to keep the water-use rate within the limits required to maintain groundwater resources and streamflows. The program controls not the supply of water, but its unit cost to the users. Any user may take as much water as desired as long as the charge is paid. The charge program generates revenue for the government and forces the users to pay for the public resource they receive. It may be argued, as Lyon (1982) does for sales of transferable permits, that attempts to address the financial burden issue with refunds to users will render the program ineffective. If the refunds vary with the amount of water used, the users will incorporate that knowledge in their decision making and the outcome will not be economically efficient.

3. S*ubsidies.* A variation of the charge program is its opposite, i.e., a program of subsidies for water conservation. An analogous example is the federal construction grants program in controlling water pollution. These programs have the same weaknesses as the charge program, plus some of their own. First, they require money to be transferred to the users from the government and must be financed by the taxpayers. Second, users may receive subsidies for conservation measures they might have undertaken anyway.

4. *Water-use permits.* A fourth approach is a program under which each user receives a permit allowing a rate of withdrawal that depends, generally, on ambient conditions. Legally, the permit may be a deed to the water itself (implying a property right), or it may be a license (implying temporary permission to use the resource).* The agency must decide exactly what the permit should entitle its holder to do, and under what circumstances. It is customary to think of it as entitling its holder to withdraw a certain volumetric flow rate (e.g., L/s) of water, but an operational definition must be more complete and must say by whom, where, and for how long the water may be withdrawn and what happens when there is not enough water to supply all permits. Other agency decisions are on what basis permits should be distributed initially; how many should be distributed; whether or not they should be transferable and, if so, under what circumstances or restrictions; and so on.

Without offering a quantitatively buttressed argument, it is asserted here that of these four approaches, the latter, permits, is most useful as the mainstay of a water-management program. Some ad hoc processes may be necessary to address unforseen circumstances, and charges or subsidies may prove useful in certain cases. Permit systems, however, have the intuitive appeal of being revenue-neutral to the agency while, conceivably at least, offering a foundation for the goal of day-to-day operation of the management program stated previously.

* No decision in that regard is rendered here; throughout this document the word *permit* denotes an explicit legal tender to water, however limited. The word *right* denotes a more general claim, whether legally explicit or legally or culturally implicit.

4.4 PERMIT PROGRAM OBJECTIVES

The objectives in designing a water-management program can be grouped into six major categories. The categories are not mutually exclusive and some overlap may occur among them.

4.4.1 Ease of Implementation, Administration, and Enforcement

The first objective is to maximize the ease on the agency's part in setting up the program initially (implementation), operating it routinely (administration), and insuring compliance with it (enforcement). This objective could be included in the cost-efficiency objective following, but is separated because many agencies will consider their own costs separately from those to society at large.

Two important components of this objective are simplicity and comprehensibility. Regulations that are simple and easy to understand are usually easy to implement, administer, and enforce as well. Simple regulations are more likely to be followed because regulatees understand more clearly what is required of them and cannot as successfully use a confusion defense if caught in a violation.

4.4.2 Equity

The second objective is maintaining equity among water users. Equity is very much in the eye of the beholder, and the decision about what is equitable or inequitable is arbitrary to some extent. Moreover, each party affected by regulations can usually be counted on to view the most equitable rule as the one which benefits him or her the most. Consensus over equity issues may be difficult to achieve, but a compromise may be acceptable. Through negotiation, it should be possible at least to avoid a program that is agreed by all to be inequitable.

4.4.3 Effectiveness in Protecting Water Resources

The third objective is that for which the program was originally devised, viz., the maintenance of streamflows or groundwater resources. Some programs address these needs directly, others only indirectly. Better protection of resources may require greater effort of administration and enforcement or less economic efficiency.

4.4.4 Robustness and Flexibility

In devising any program to control water withdrawals, there will inevitably be errors in data collection and in predicting the outcome of the program. Hence, an objective in program design is that the social benefit of the program be robust, or insensitive to errors in the data upon which its design depends. Robustness is not necessarily a static ability to withstand unforseen changes but may instead be a dynamic correction in response to such changes.

4.4.5 Economic Efficiency

A fifth social objective, already touched upon, is economic efficiency of water use. Different management programs will generally result in different distributions of water among users, implying different aggregate levels of economic benefit (efficiency) from water use. Given the finite supply of water, only one distribution will maximize benefit, but some other programs may fare better with respect to other program objectives.

Generally, a lack of regulations results in an economically inefficient distribution because those to whom water is physically most available take a larger share than is efficient. On the other hand, regulations that restrict water use without consideration of economic efficiency may erect barriers to voluntary redistribution that would increase efficiency.

4.4.6 Political and Legal Feasibility

The final objective is political and legal feasibility. No matter how well designed a program is with respect to the other five objectives, it will fail if it is not accepted politically or if it cannot operate within prevailing laws. Most other objectives are fortunately coincident with, if not constituent of, this objective. However, a well-designed program may require small changes in prevailing law, and such programs may be rendered politically infeasible if misconceptions about them gain widespread acceptance through the actions of pressure groups or lobbyists.

4.5 PERMIT PROGRAMS FOR WATER-WITHDRAWAL CONTROL

The agency's choice of options determines how the water-permits program fares with respect to program objectives. It will not generally be possible to find a set of choices that maximizes all objectives, and the agency will be forced to trade one objective off against others. There is no optimal or "best" tradeoff point among those objectives, only one that is most acceptable in the agency's judgment. The following paragraphs discuss several options and the tradeoffs among the objectives for each.

Although it is important in the design of a program, no distinction will be drawn for the moment between water withdrawal and consumption; these terms will be used interchangeably, along with the term *water use*. It is assumed that all use is consumptive and that water, once withdrawn from a watercourse, disappears forever. The importance of return flows, their effects, and the options for incorporating them in the regulatory program are discussed near the end of this section.

4.5.1 Geographical and Temporal Configuration of Programs

The agency must decide the geographical and temporal extent of water-control programs and how the activities of smaller-scale programs should be delineated and coordinated to achieve overall agency goals at the larger scale (e.g., state or province). Conceivably, limits for surface-water programs would best follow basin boundaries, but those boundaries do not necessarily correspond to the natural lines of cleavage for groundwater. Moreover, there is a frequent reluctance of state or provincial and local authorities to delegate the necessary authority to multistate or multiprovince river-basin commissions, especially when nonwater interstate issues are at stake. Thus, the central coordination might most feasibly take place at the state or provincial level.

For both surface- and groundwater, there is the question of whether individual ad hoc programs should be developed for each local problem area, whether each new program should be extended to the smallest political boundary that contains the problem area (e.g., a city, township, county, or multicounty unit), or whether a more comprehensive program should be developed for a state, province, or river basin. A comprehensive uniform state- or province-wide program is more difficult to administer but has no greater effectiveness than individual programs designed for problem areas. Nevertheless, if the water problems of several individual areas, each with a local governing authority, expand to their jurisdictional boundaries, each local authority may act to protect its local interest. There may then be a need for a higher authority to intercede to distribute water among them. With local authorities already in place, political realities may require the central authority to administer through the local authorities. This may be less efficient than if more central regulation were in operation from the outset. Thus, local programs are likely to be effective only as long as an adequate plan for their future coordination and, if need be, their future subvention to central authority, is in place at the outset.

4.5.2 Permit Definition Basis

The manner of determining how much water the permit holder may take is referred to as the *definition basis* of the permit. The two major permit definition bases are discussed briefly here and in more detail by Eheart and Lyon (1983). Other alternatives are discussed by Eheart (1989); for brevity those are omitted here.

Let the amount of water (e.g., L/s) available for distribution to the permit holders be referred to as the *total allowable withdrawal* (TAW). For surface water the TAW equals the streamflow minus the minimum flow requirement for instream flow needs. For groundwater, the TAW is set considering target aquifer levels or drawdown rates. For now, assume that when the agency sets the TAW it also distributes it among users according to some "fair" formula. (The complexities of channel geometry, discussed later, will complicate the formula for rivers.) There are then three bases, discussed as follows, for defining permits, i.e., relating the allowable withdrawal to the TAW.

1. *Constant use basis.* Under this basis, each user is allowed a certain constant use rate. The rate may vary with time, possibly following an annual schedule, but does not depend on immediate stream conditions unless the streamflow is insufficient to satisfy all claims. The chief drawback of this basis is that it does not spell out what happens during drought.

2. *Prioritized permit basis.* Traditionally this basis has been used to allocate surface water in the western United States under the appropriative doctrine. A set of priorities among users is established. Any given user is allowed a certain constant rate of water withdrawal as long as the TAW is enough to satisfy him or her and all other users with a higher priority (including instream needs). When the TAW is insufficient to satisfy all users, they forego withdrawals according to their priorities.

3. *Flexible permit bases.* Under these bases, users' allowable withdrawals increase and decrease continuously with the immediate streamflow. The simplest of these bases is the fractional basis, under which each user is allotted a constant percentage of the TAW. Thus, as the TAW fluctuates, so does the amount of water allotted to each user and no user is ever entirely deprived of water. This type of permit has the advantage of homogeneity over prioritized permits; no assignment of priority need be made to individual users since no user's permit has priority over any other's. There may, however, be some perceived equity and administrative disadvantages to this basis. First, there is no easy way to issue free permits to newcomers, although they may easily buy their way in and may be accommodated through staggered limited-duration permits (see following). If entering users are to share available water with existing users who are already using it all, then the existing users must give up some. Forcing them to do so might be viewed as a confiscation of property without compensation. By contrast, the prioritized-permit basis may be structured to accommodate newcomers by assigning them lowest priority. The second disadvantage is that it is more difficult to account for the geometric complexity of real river systems under this type of permit-definition basis. Ways of addressing this problem also exist, however, and are discussed in greater detail following.

4.5.3 Allocation Basis

The agency must decide the basis for distributing permits among the users, i.e., deciding what size permit (and the priority, if appropriate) each user receives. The size of permit is the rate of allowable withdrawal for the prioritized-permit system; for the fractional-permit system, it is the fraction of the TAW the user is allowed to withdraw.

In the West, priority allocations were established by the rule of "first in time, first in right." Size allocations were based on the agency's judgment about the amount of water each user needed or deserved, based in part on past use or anticipated future use. These allocations were made on a case-by-case basis, and there was no guarantee of consistency from one time period or location to another.

It would be difficult to apply appropriative rights in riparian areas. First, since no water-rights system existed historically, it would be difficult to establish priorities among users. Second, the notion of priorities runs counter to the legal precedent of the riparian doctrine and has been legally rejected in some riparian areas (see, e.g., for Illinois, *Bliss v. Kennedy*, 1867).

One alternative approach is a formula that considers the historical use of water by each prospective recipient as a measure of need. Presumably, the size allocations would be roughly proportional to some measure of past-use rate. Provisions could be incorporated to avoid rewarding those who had used water wastefully in the past, and certainly to dissuade users from profligate use in the present for the sole purpose of receiving a higher allocation in the future. This poses the problem of estimating past withdrawal rates that may have been unmeasured. Most municipal and industrial withdrawals are gauged, but many agricultural withdrawals are not.

An alternative basis rests on the theory that the right to a certain amount of water is internalized in the worth of riparian land. In this context, the issuance of a permit by the agency may be viewed as an attempt to separate and grant to the riparian landowner the water right that was historically bound up with the land right. Under this approach, the size of the fractional permit issued to a given surface-water user might be proportional to such a parameter as the length of riparian streamfront or the area of riparian land. (An equivalent rule for the distribution of groundwater permits for irrigation might be in proportion to the area of overlying land.) Such a basis of allocation might be considered equitable for agricultural users, but the agency may deem it desirable to set aside a portion of the water for other uses, to be distributed among them on a different basis (e.g., proportional to population equivalent for municipalities and earnings, taxes, or employees for industries).

Another alternative is simply an ad hoc allocation basis that seeks to use no formula in distributing the available water among users but, rather, leaves such decisions to the judgment of the agency. While such an approach has the benefit of flexibility, it risks potential challenges of arbitrary or capricious behavior on the part of the agency.

While priority allocation to individuals is inconsistent with the riparian doctrine, as noted, some states in the United States have embraced a priority allocation among use types, placing so-called *natural needs* above *artificial wants* in priority (e.g., Clark, 1985). Even in appropriative regions, such priorities often take precedence over the first-in-time rule. Thus, a regulatory system might serve all municipal requirements before satisfying the requests of industrial, power, or agricultural uses. It may be, however, that only the portion of municipal withdrawal that constitutes natural needs is felt to be eligible to be set aside first, so that the remaining portion of the municipal claim should have to compete with other claims.

4.6 TECHNICAL DETAILS OF PERMIT SYSTEMS

Several detailed aspects must be addressed for establishing any permit system.

4.6.1 Duration of Permits and Accommodation of Newcomers

The agency faces a dilemma in deciding the length of time a permit is valid. As noted by Eheart and Lyon (1983) and Young (this volume, Chap. 3), permits that are valid for only a short time may not allow the users sufficient time to pay off capital equipment and may thus result in economically inefficient decisions. (For example, the user may purchase less-expensive equipment that uses water inefficiently.) A long permit validity will present difficulty in accommodating newcomers, and there is a risk of overcommitting the resource and being unable to reverse the allocation process except by repurchase. Eheart and Lyon (1983) note that one way of addressing this dilemma is a system of staggered permits of n-year duration under which the agency may lower the total number of permits by as much as $1/n$ per year simply by not reissuing expired permits.

There is a potential problem in initiating such a system or accommodating new users, since users prefer long-term to short-term permits. Nevertheless, it may be possible to initiate the system if the agency acts at an early stage, before demands become significant in comparison to supplies. If $1/n$ or less of the target number of permits are currently held and less than $1/n$ additional will be claimed in aggregate by all new users each year, the agency may issue n-year permits to all new applicants.

A staggered system of limited-duration permits could be structured to include newcomers in the allocation, granting them the same status as existing users and enabling them to acquire an increasingly large number of permits each year. Thus, for example, under a 5-year staggered system, a newcomer who, as an existing user, would have claim to 15 percent of the permits would, like existing users, be allocated 15 percent of the 20 percent of new permits (3 percent) that are reissued each year. Requiring a user to wait 5 years for the full allocation might not be practical; depending on the water-use application, operating at reduced production capacity may not be economical. In such cases the staggered system might be effective when combined with a spot market, so that the user could rent permits for the interim.

4.6.2 Averaging Periods

It is not possible for water users to restrict their withdrawals to a certain flow rate at all times, nor is it always desirable from the agency's perspective for them to do so. Crops need only be irrigated when they undergo moisture deficit, and municipal demands fluctuate according to the weather and the incidence of fire. It is therefore appropriate for the agency to grant users some flexibility by restricting their time-averaged, rather than their instantaneous, withdrawals. The question then becomes one of choosing the averaging period. The larger the averaging period, the more flexibility the users have but the greater the opportunity that exists for them to overexploit the water resource, at least temporarily.

If dewatering of an unimpounded stream is to be prevented, the averaging period will usually have to be from less than a day to a few days, depending on the size of the stream. Since most aquifers recharge only at certain times of the year, a 1-year averaging period might be appropriate for groundwater withdrawals.

In its monitoring activities, the agency will calculate the average periodically, since it is likely to be impossible for it to keep a running tab on the time average of every user's withdrawal continuously. The necessity of performing these calculations at certain times presents incentives for perverse behavior, however. Toward the beginning of a period (i.e., just after averaging) there is an incentive for users who put less stock in future worth to increase withdrawals, perhaps hoping that it will rain more toward the end of the averaging period. Toward the end of the averaging period (i.e., just before averaging) there is an incentive for users who originally husbanded their water to increase withdrawals because they will lose what they don't take. (If they don't lose what they don't take, the averaging period is effectively longer.) Thus, around the time that the calculations are performed, there may be increases in total withdrawals. To address this problem, it may be worthwhile to stagger the times of calculating averages for different users so that the increased withdrawals are spread evenly over time. Under prioritized permits, carryover from the previous period could be assigned a low priority, thus making it available to the user, but not with the same value.

4.6.3 Large-Scale Groundwater Restrictions

There are several issues related to selecting an aggregate allowable rate of removal from groundwater aquifers. One is how much water should be saved for the future. When aggregate withdrawal exceeds aggregate recharge, aquifer mining is said to be taking place. The word *mining*, although commonly used, may be a misnomer;

unlike coal or oil, groundwater is a renewable resource and will eventually return to higher levels when aggregate pumping stops or is reduced. Nevertheless, many years' worth of recharge may be removed in a very short time, and future generations may have to wait some time before enjoying the aquifer levels of their forebears. Sometimes it is difficult to document mining because aquifer levels usually undergo annual fluctuations that may mask long-term decreases.

Apart from this issue, aquifer mining implies a continual depression of the piezometric head and affects most users drawing from a groundwater unit. It requires periodic redrilling of wells or lowering of pumps as a matter of course, as well as a continuous increase in pumping costs. Thus, even if aquifer mining has the endorsement of the agency, it is reasonable to require that a substantial number of users in a region be similarly disposed toward mining and be willing to cope with its cost and inconvenience in order to reap the benefit of (temporarily) increased allowable pumping rates.

4.6.4 Complexity of Surface-Water Programs Using the Flexible-Permit Basis

There is a problem with adopting the flexible-permit basis for surface water in humid regions. If it happens that all water flows from an area where none is used, through a central point from which it is physically available to all users, fractional permits could have the clear meaning of restricting each user to its fraction of the flow rate at that point. But where sources and sinks of water are geographically dispersed, and the amounts that are physically available to be shared are not the same, such fractional assignments among users lose their meaning.

Two methods are proposed here of indexing the allowable withdrawal to the observed streamflow. Under one, a gauge upstream of each user determines its allowable withdrawal; under the other, the stream-gauge reading controls users' withdrawals upstream of it. The former, termed the upstream-gauged system is administratively easier, but requires a large number of stream gauges to be used effectively. The latter, termed the downstream-gauged system, is administratively difficult but uses a smaller number of gauges. They are discussed in greater detail as follows.

Upstream-Gauged Systems. Under these systems, a group of users located in a stream reach will share a TAW for that reach equal to the streamflow at the nearest upstream gauge minus minimum streamflow and an amount set aside for users downstream of their reach. There are two variations of this system, distinguished according to how the amount of water a user must pass through to downstream users is determined. Only the fixed pass-through system is practicable; a discussion of the variable pass-through system is important to aid conceptual understanding of the issue.

Variable Pass-Through. Under the variable pass-through system, a set of contingency arrangements based on readings at upstream gauges determines the withdrawal rate to which any permit holder is entitled. Assuming fractional permits as the implementation of flexible permits, each user is assigned a number proportional to its share of all the water available in any basin that contains it, minimized over all basins that contain it. As the basin becomes larger, more water is coming into it, but there are more users with whom the user has to share. For example, a user located on a small tributary to the Mississippi River in Iowa might be entitled to the minimum of: half of the water shared by it and its nearest two neighbors; 5 percent of the water shared by those users and those in the basin upstream of the next downstream gauge; and so on, to 0.00015 percent of all the available water in the Mississippi basin. Unfortunately, this system is impracticable as it requires an infeasible amount of information processing and assumes a zero residence time in the channels.

Fixed Pass-Through. An alternative that avoids this problem is to impose fixed obligations to downstream users, so that curtailment orders to a given group of users are in no way dependent on stream-gauge readings downstream of them, regardless of what actually happens to the downstream users' water supply. This will allow indexing a user's allowable withdrawal to the nearest upstream gauge. Under this system, users are divided into groups residing between adjacent stream gauges and each group of users is required to forego a fixed fraction of its incoming TAW for downstream users. An important question is how that fraction should be determined. Conceivably, it should be chosen in consideration of equity and economic efficiency, and should ideally take account of the number and type of downstream users as well as their potential for supplies from other tributaries. One possibility is to use an estimate of the average value of the aggregate downstream claim (as a fraction of incoming TAW) under the variable pass-through system. Harrison (1991) has studied this method for a limited data set and has concluded that it has high economic efficiency.

Permit transfers would be administratively convenient for the fixed pass-through system; the size of the transferred permit would simply be assigned to the new reach. As an example, consider an upstream user that currently is allowed to take 20 percent of the TAW at its upstream gauge, 45 percent being allocated to other users in the group and 35 percent passed to downstream users. If that user wishes to sell its permit to a downstream user, the remaining users in the upstream group will be required to pass $20 + 35 = 55$ percent to the downstream users. The downstream neighbors of the buyer will each have a smaller percentage of their incoming TAW, but that TAW will be commensurately larger. For a transfer of a permit in the upstream direction, the system operates the same, except that in the example, no more than 55 percent could be moved to the upstream group.

Downstream-Gauged Systems. Under these systems, the permit is actually a set of different permits, each contingent on a different gauge constraining or limiting withdrawals. In essence, the system operates like the upstream-gauged system except that an attempt is made to infer the streamflow at an ungauged point upstream of the users from the observed streamflow downstream of them. The allowable withdrawal by a user is its fraction of the TAW at a stream gauge, as shared with all other users upstream of the stream gauge. A correction must be made to the TAW to account for current use of water; thus, assuming 100 percent consumption, the sum of current uses is added to the observed flow rate at the gauge, the minimum flow requirement is subtracted from that quantity, and the resulting TAW is distributed among the users between that gauge and the nearest upstream gauge. While this system is practicable, it may not be practical, since it essentially uses a feedback control mechanism in that a user's withdrawal influences the streamflow at a gauge, which in turn influences the user's allowable withdrawal. This may lead to problems of stability if return flows are not adequately accounted for, but some alteration of the curtailment order process based on experience and trial-and-error by the agency may address this problem in some cases, especially if the basin is small.

4.6.5 Complexity of Surface-Water Programs Using Prioritized Permits

For prioritized permits there is little difference between upstream variable pass-through and downstream-gauged systems. Curtailment orders are issued to upstream user's in decreasing order of seniority whenever a senior user's withdrawal is unavailable. The gauges are used primarily to determine where curtailment occurs and to insure that minimum streamflows are met.

The fixed pass-through system could be used for such permits if a fixed estimate is used to represent downstream senior claims. For example, if the claim of a senior user is 15 m^3/s, and the average flow available to that user from tributaries with no users is 10 m^3/s, one possible fixed pass-through system would require an upstream junior user to pass the first 5 m^3/s of its incoming TAW to the downstream senior user.

This system does not use a feedback mechanism, but it could result in a reversal of allocation compared to the variable pass-through system. If the upstream user's stream has a high flow and the other tributaries of the downstream senior user have low flows, the upstream junior user might be completely satisfied while only the first 5 m^3/s of the downstream senior user's claim would be fulfilled.

4.6.6 Withdrawal versus Consumption and Accounting for Return Flows

As noted later, if all withdrawal permits are defined in terms of consumption, some, but not all, of the problems of third-party impacts may be avoided. The avoided problems are the middle-user impairment problem (Anderson, 1983; 1983b; Johnson et al., 1981; Tregarthen, 1983) and exacerbation of the feedback problem for downstream-gauged systems. Definition in terms of consumption also serves as a greater incentive for water conservation since users would be required to hold permits for only the water they consume.

There are some drawbacks to consumptive-use permit definition, however. It requires assuming that there is always adequate additional flow in the stream for each user's pass-through use, that such use is returned near the point of abstraction, and that the return flow rate is accurately known. Unfortunately, return flow rates or consumptive fractions are not always constant, may change seasonally and unpredictably, and may be difficult to determine (especially for irrigation, whose return flows are geographically dispersed). Thus, while it may be more fair to base permit definition on consumption, it is administratively easier to base it on withdrawal.

One approach to this dilemma is to operate the program as though all withdrawal is 100 percent consumptive unless the user can document his or her return flow accurately, and to choose a program structure that is robust to the operational uncertainty caused by this assumption. Assuming 100 percent consumption, the return flows could be thought of as tributaries of uncertain and unpredictable magnitude. Downstream-gauged systems would not be robust to this uncertainty because they require accurate predictions of return flow, as do upstream-gauged, variable pass-through systems. For fixed pass-through upstream-gauged systems, however, uncertain return flows show up at nearby gauges to add to the flow at the next reach. Under that system, therefore, while the assumption of 100 percent consumption might be viewed as inequitable by users with significant but difficult-to-measure return flows, it would not cause operational problems.

4.6.7 Interactions between Ground and Surface Waters

If an aquifer is hydraulically connected to a stream, it may be possible for withdrawals from one medium to deplete the other, interfering with proper accounting of both. The ideal accounting method is to determine for a given user what portion of his or her withdrawal comes ultimately from each source and to require a permit of that quantity to be held. Unfortunately, it is difficult to determine a priori what each portion is, because it depends on a complex set of natural parameters that vary with time (e.g., the relative fractions of rainfall that infiltrate and run off), as well as the actions of the user and other users.

4.7 VOLUNTARY WATER-TRANSFER SYSTEMS

One feature that the agency may wish to consider is to allow permits to be transferred voluntarily among users, either on a permanent or temporary basis. Water transfers are a common component of many regional water systems and are being increasingly considered for meeting growing water demands and for managing the impacts of drought. Water transfers can take many forms and can serve a number of different purposes in the planning and operation of water-resource systems. However, to be successful, water transfers must be carefully integrated with traditional water-supply augmentation and demand-management measures as well as with the institutional systems which regulate water use. This integration requires increased cooperation among different water-use sectors and resolution of numerous technical and institutional issues, including impacts to third parties. This section identifies the many forms that water transfers can take, some of the benefits they can generate, and the difficulties and constraints which must be overcome in their implementation.

The most frequently cited argument in favor of this approach is an economic one (see, e.g., Anderson, 1983a; 1983b; Eheart and Lyon, 1983; Wong and Eheart, 1983; Enright and Lund, 1989). Greater economic efficiency will accrue if permit trading is allowed; the approach is also flexible, robust, and does not require strong intervention by the agency. For example, when newcomers buy permits, there is an automatic redistribution of water use toward greater efficiency.

Historically, advances in water-system management have been motivated by socioeconomic and environmental considerations. Since the 1970s, the increasing expense and environmental impact of developing traditional water supplies (e.g., reservoirs) have encouraged innovative use of existing facilities (e.g., conjunctive use and pumped-storage schemes) and have led to expanded demand-management efforts. In recent years, growth in water demands and environmental concerns have caused even these innovations to yield diminishing marginal returns. These economic and environmental conditions, combined with recent droughts, have spurred further efforts to improve traditional supply-augmentation and demand-management measures and have motivated the recent consideration and use of water transfers. The use of water transfers in many parts of the country, especially in the West, can be seen as a natural development of the water resources profession seeking to explore and implement new approaches in water management. We begin with brief reviews of recent water-transfer activity in California (Lund and Israel, 1995a; 1995b; Lund et al., 1992). Following that, some of the more relevant issues for water managers and planners contemplating the use of water transfers are reviewed.

4.7.1 Existing Examples of Water Transfers

Water transfers and water marketing have existed in one form or another in many parts of the United States since early in this century. Many metropolitan areas have some form of water market in operation, usually involving a single large seller, typically a large central city or utility company, selling water to numerous large and small suburban cities and water districts. These sales arise from the economies of scale of urban water-supply acquisition, conveyance, and treatment and the historical legacy of central cities being the first to acquire most of the better, larger, and least expensive water supplies in many regions. Both central city and suburban parties to these transfers and sales accrue significant advantages from this arrangement, in the form of lower water-supply costs, higher supply reliabilities, and greater capability and certainty in regional water-supply planning. Still, there is often some degree of controversy and conflict between parties to these transfers (Lund, 1988).

Water marketing and transfers within agricultural regions is a still more ancient practice. Maass and Anderson (1978) describe a very effective water-marketing arrangement that has been in effect in one area of Spain since the fifteenth century. In addition, there are almost countless water trades and sales between farmers throughout much of the western United States. The majority of these transfers occur within mutual irrigation companies. These companies are typically informally constituted cooperatives of farmers, without governmental status. Each farmer has a share of the total amount of water available to the company. Water is then transferred by rental or sale of these shares to other farmers within the venture (Hartman and Seastone, 1970). It has been estimated that there are roughly 9200 such mutual water companies in the western United States (Revesz and Marks, 1981).

Other examples of existing water transfers are presented by MacDonnell (1990). This review found that almost 6000 water-right change applications were filed in six western states between 1975 and 1984, primarily in Colorado, New Mexico, and Utah. The vast majority of these applications were approved by state authorities. There are untold additional cases where transfers have been effected without legal need for state approval. For example, water transfers within the Bureau of Reclamation's Central Valley Project (CVP) generally do not require state review, since the Bureau is the holder of very general and flexible water rights. Between 1981 and 1988, CVP contractors were involved in over 1200 short-term transfers involving over 3700 Mm^3 (3 million acre · ft; Gray, 1990).

Most of the transfers described above are confined to specific water sectors and within individual metropolitan areas or irrigation systems. However, contemporary interest in water transfers has broadened the scope of traditional transfers to include transfers between different water-use sectors, e.g., agriculture to urban, often over larger geographical distances. These transfers often involve many parties with diverse views, facilities, and water demands which are more geographically separated. They may also require the use of conveyance and storage systems controlled by parties who are neither selling nor purchasing water. The controversies and complexities of effecting water transfers under these conditions may have initially deterred water managers from pursuing this option. However, with the changing economic and policy environment of water management and the absence of other attractive choices, water transfers can offer engineers a cost-effective alternative for enhancing the performance and flexibility of their systems (Lund et al., 1992).

4.7.2 Economic Theory of Water Transfers

There is vast literature on the merits of water markets and voluntary water transfers (Milliman, 1959; Hartman and Seastone, 1970; Howe et al., 1986; Brajer et al., 1989; Eheart and Lyon, 1983). One important question addressed by some of this literature is the magnitude of the potential efficiency gains from trading. In several studies, it was estimated by computer simulation of markets in irrigation water to be significant. Wong and Eheart (1983) report an improvement of about 13 percent over nontransferable permits for surface-water permits from the Little Wabash River in Illinois. Enright and Lund (1991) report only around 1 percent for a simple demand system exposed to different hydrologies, the Mad River in California and Tionesta Creek in Pennsylvania. Eheart and Barclay (1990) report an improvement ranging between 3 and 86 percent, depending on the amount of water available and the predictability of weather and crop-yield response. Other investigators (e.g., Tregarthen, 1983) have indirectly confirmed these findings.

Improved economic efficiency is not the only important advantage of such a program. Transfers also provide an incentive to develop and adopt ways of using water more efficiently through recycling and waste reduction. It is widely believed that many of the historically documented cases of wasteful use of water under the appropriative system might have been avoided by allowing transfers (Anderson, 1983b).

With regard to political feasibility, transfers of both water and pollution permits have received endorsements from policy analysis organizations of various political persuasions (Tietenberg, 1985; Bandow, 1986; Stavins, 1989). Permit transfers are a part of the most recent version of the Clean Air Act (1990), following a report commissioned by the U.S. Congress that endorsed a host of market-based incentives for environmental protection (U.S. Congress, 1988).

The additional administrative costs imposed by transfers are expected to be modest. The agency must maintain a registry of permits. Registration of trades must be sufficiently formal, and enforcement adequate, to prevent users from simply transferring permits from one to the other just before the enforcement agent arrives to check for violations. Transfer restrictions must be decided upon and administered. No cost data need be collected, however, and no cost optimization need be done by the agency. The agency may opt to set up a brokering operation or may let a private concern do so, but transfers are generally voluntary and need not be brokered by anyone.

In addition to these potential strengths of permit-transfer systems, there are some issues that must be addressed before an agency would wish to embark on the development of a system of such transfers.

4.7.3 Imperfect Markets

While water-market transfers are often desirable, the economic efficiency of water markets is usually imperfect when compared to ideal market performance. The conditions required for a perfect market are difficult to attain for a commodity such as water. Some problems include (Howe et al., 1986; Brajer et al., 1989):

- Water rights are often poorly defined.
- Water transfers can have high transaction costs.
- Water markets will often consist of relatively few buyers and/or sellers.
- Water is often costly to convey between willing buyers and sellers.
- Communication between buyers and sellers may be difficult.
- In humid regions, the dispersed nature of water sources and sinks may make definition of water rights difficult.

These difficulties commonly exist for other goods and services provided with great success through market mechanisms, and are not barriers to the use of water markets. The political appeal of market in offering trading opportunities and incentives for innovation is undiminished by these concerns. However, in appraising water transfers, planners, engineers, and policymakers should consider that trading activity may not be as lively as originally anticipated.

4.7.4 Third-Party and Environmental Impacts

The transfer of water can significantly affect third parties not directly involved in the transfer. The neighbors of the buyer may be impaired and the neighbors of the seller may receive a windfall benefit under permit transfers. For example, a freely transferable permit to pump a given amount of groundwater could impair the neighbors of the buyer with additional drawdowns and pumping costs in their wells. Furthermore, permits for irrigation will tend to be transferred toward farms whose soils have low moisture-retention capacities and these farms are often close together.

Another type of *third-party impairment* associated with prioritized water permits defined as withdrawals (Anderson, 1983a; 1983b; Johnson et al., 1981) is the forfeiture of third-party rights that may attend a permit transfer. Consider, for example, a low-priority, third-party user situated downstream of the seller of a high-priority permit and upstream of its buyer. If the seller historically has a large return flow, the middle user has become dependent on that return flow, and the transfer requires the middle user to curtail or forego withdrawal to preserve a streamflow adequate to satisfy the downstream user.

The greatest challenge for implementing water transfers in the future may lie in properly identifying the affected parties and adequately mitigating these impacts. Many interests can be affected by water transfers, as noted in Table 4.1. Impacts can be direct, as with reduced instream flows below the diversion point for a transfer; or secondary, as represented by the loss of farm-related jobs in an agricultural region when farmers choose to transfer their water supplies. More detailed discussions of the third-party impacts of water transfers appear in Eheart and Lyon

TABLE 4.1 Some Potential Third Parties to Water Transfers

Urban	
	Downstream urban uses
	Landscaping firms and employees
	Retailers of lawn and garden supplies
Rural	
	Farm workers
	Farm service companies and employees
	Rural retailers and service providers
	Downstream farmers
	Local governments
Environmental	
	Fish and wildlife habitat
	Those affected by potential land subsidence, overdraft, and well interference
	Those affected by potential groundwater-quality deterioration
General	
	Taxpayers

(1983); National Research Council (1992); Howe et al. (1990); and Little and Greider (1983).

Paradoxically, water transfers might aid members of a group in one region while harming other members of the same group in another region. Water transfers from one farming region to another will lower farm employment and demand for farming services in the selling region and increase them in the purchasing region. Similarly, transfers of surface water from farms to cities can both help and harm fish and wildlife. By reducing application of water to farms, water quality downstream of the farm might improve, to the benefit of fish and perhaps other downstream water users. Also, there is a likely reduction in fish kills at the farm intake pumps because of the decreased withdrawals. Yet, where the on-farm application of water serves as habitat for migrating waterfowl, the removal of this water could harm bird populations.

Several mechanisms have been suggested to ameliorate the impact of or compensate groups harmed by water transfers. These mechanisms include (National Research Council, 1992; California Action Network, 1992):

- Taxing transfers to compensate harmed third parties
- Requiring transferors to provide additional water for environmental purposes
- State compensation to help economic transitions in water-selling regions
- Requiring public review and regulatory and third-party approval of transfers
- Requiring prior evaluation of third-party impacts of transfers, similar to an environmental impact report

- Requiring formal monitoring of third-party impacts
- Restricting transfers to "surplus" waters
- In certain cases, redefining water rights to prevent third-party effects

Trading restrictions may undermine efficiency gains and may not be effective in protecting third parties anyway. Eheart and Barclay (1990) found in simulations of water-permit markets for irrigation in Kankakee County, Illinois that the return of water was sufficiently homogenous among users that no user bought more than 40 percent of his or her original allotment even without the imposition of transfer restrictions. At the same time, Cravens et al. (1989) found that seasonal drawdown associated with irrigation is very significant for the same aquifer, even without transfers.

Third-party compensation has been endorsed by some researchers (e.g., Coase, 1960). Others (e.g., Baumol and Oates, 1988) note potential problems of strategic behavior if third parties are given final authority on whether a transfer is allowed.

Direct impacts on third parties can be reduced through legislation. Potential third-party effects from changes in return-flow quantity are commonly eliminated by state regulation allowing the transfer of only consumptive water use (Gray, 1989). Nevertheless, difficulties in assessing consumptive use may cause impacts to users of return flows (Ellis and DuMars, 1978). Likewise, the relative magnitude of secondary impacts is often difficult to determine accurately, but their presence is undeniable. Under ideal economic conditions of full employment and perfect labor and materials markets, such secondary impacts should be self-canceling in the aggregate. However, the common presence of significant unemployment, imperfect labor and capital mobility, and potentially important equity impacts raise these secondary economic impacts of water transfers to prominence.

4.7.5 Nonuser Market Players

Fourth, if permits are initially given away rather than sold and are also transferable, there is a potential for certain other kinds of inequities, if only perceived ones. A user who never intended to exercise a permit could, upon receiving it free from the agency, sell it and reap a windfall reward. There thus may be a need, if only a political one, for the agency to scrutinize requests to insure that claimants are bona fide potential users of water and not just speculators attempting to acquire windfall permits to sell later. On the other hand, some would argue that anyone who meets the established requirement to receive a permit is entitled to do so whether or not his or her doing so is for speculative reasons.

4.7.6 Market Thinness

A fifth issue related to water transfers is that because of the usually small number of participants, one or a few parties may be able to manipulate the permits market to their advantage. Many other markets are vulnerable to such manipulation, and while manipulation may pay off to an individual, trading will always improve the overall economic efficiency of the outcome. No exchange, no matter how it is manipulated, will take place unless both parties have an interest in it. Several researchers (e.g., Hahn, 1984; Saleth et al., 1991) have studied this problem and have estimated that the potential for market dominance is small, but exceptions are possible, and would occur where one user is much larger in size than any other and can exercise monopoly or monopsony strength. Such cases pose problems of both efficiency and equity.

4.7.7 Multiple-Forum Origins of Transfers

A sixth issue is that water transfers can emerge from various forums: bipartisan or multilateral negotiations, several forms of brokerage and bidding, and other means (Hartman and Seastone, 1970; Saleth et al., 1989). There is of course the potential to mix the use of different forums in the water-transfer process, using one forum to set a price and quantity, with other forums performing technical and legal review of transfer proposals. The forum or institutional mechanism under which water transfers are developed, reviewed, and approved can substantially affect the number, type, and details of transfers that actually take place, and is particularly important for the consideration of third-party impacts (Nunn and Ingram, 1988; Little and Greider, 1983; Eheart and Lyon, 1983).

4.8 TYPES OF WATER-TRANSFER ARRANGEMENTS

Water transfers can take many forms, as presented in Table 4.2. The specific needs of the purchasing and selling parties may dictate the type of transfer sought and the forum through which transfer arrangements are made. However, existing legislation and recent transfer experiences will also be important in selecting the most appropriate form of transfer. Each transfer form can have a different use in system operation and has different advantages and disadvantages for water buyers, water sellers, and other groups (Lund et al., 1992). The various uses and associated benefits of water transfers are summarized in Table 4.3. Additionally, water transfers, like many forms of water-source diversification, increase the flexibility of a water system's operation, particularly in responding to drought. This flexibility allows new forms of operation that could not be accomplished without transfers and in many cases allows modification of system operations on a rapid time-scale. The following discussion on transfer types focuses on the possible uses and associated benefits of each.

TABLE 4.2 Major Types of Water Transfers

Permanent transfers
Contingent transfers/dry-year options
Long-term
Intermediate-term
Short-term
Spot market transfers
Water banks
Transfer of reclaimed, conserved, and surplus water
Water wheeling or water exchanges
Operational wheeling
Wheeling to store water
Trading waters of different qualities
Seasonal wheeling
Wheeling to meet environmental constraints

TABLE 4.3 Major Benefits and Uses of Transferred Water

Directly meet demand and reduce costs
Use transferred water to meet demand, either permanently or during drought
Use purchased water to avoid higher cost of developing new sources
Use purchased water to avoid increasingly costly demand-management measures
Seasonal storage of transferred water to reduce need for peaking capacity
Use drought-contingent transfers to reduce need for overyear storage facilities
Wheeling low-quality water for high-quality water to reduce treatment costs
Improve system reliability
Direct use of transferred water to avoid depletion of storage
Overyear storage of transferred water to maintain storage reserves
Drought-contingent contracts to make water available during dry years
Wheeling water to make water available during dry years
Improve water quality
Trading low-quality water for higher-quality water to reduce water-quality concerns
Purchase water to reduce agricultural runoff
Satisfy environmental constraints
Purchasing water to meet environmental constraints
Wheeling water to meet environmental constraints
Using transferred water to avoid environmental impacts of new supply capacity

4.8.1 Permanent Transfers

A permanent transfer of water involves the acquisition of water rights and a change in ownership of the right. Permanent transfers are a form of supply augmentation and serve many of the same needs as capacity-expansion projects, including direct use to meet demands and improved system reliability. In some instances, the direct use of permanently transferred water can delay the implementation of increasingly costly demand-management measures or the need for system expansion, which in turn has the advantage of avoiding or at least delaying potential environmental impacts associated with construction (Table 4.3).

The majority of permanent transfers involve the purchase of agricultural water rights by urban interests. These transfers can involve reversion of the farmland to dryland agriculture, the immediate or gradual fallowing of farmland, the replacement of the farm's water supplies with an alternate supply source of possibly lower quality (from an urban-use perspective), or the lease of the transferred water back to the farmer in wet years when urban supplies are plentiful. Another form of permanent water transfer, common in Arizona, is for the developer to acquire groundwater rights associated with recently developed, formerly agricultural suburban lands. Some Arizona cities have made the provision of such rights to the urban water supplier a prerequisite for annexation of new suburban developments to urban water systems (MacDonnell, 1990). This ties permanent changes in water use to changes in land use and does not require water rights to be severed from the land, a political and legal difficulty in some cases.

4.8.2 Contingent Transfers/Dry-Year Options

In many cases, potential buyers of water are less interested in acquiring permanent supplies than in increasing the reliability of their water-supply system during drought, supply interruptions due to earthquake, flooding, contamination, or mechanical failure, or during periods of unusually great demand. For these cases temporary transfers contingent on water shortages may be desirable. The appropriate time horizon and conditions for a contingent transfer agreement will depend somewhat on the particular source of unreliability that the buyer would like to eliminate. For example, the timing of the call mechanism for earthquake supply interruptions would likely be very different from the call mechanism for responding to drought. Regardless, drought-contingent contracts for water are probably best made with holders of senior water rights, since they are the least likely to be shorted during drought. However, the increased reliability of water from senior rights tends to raise its market value (Lund et al., 1992; National Research Council, 1992). An important benefit of contingent transfers is that longer-term arrangements allow for a more thorough analysis and mitigation of potential third-party impacts.

The time horizon of contingent transfers is important. Contingent-transfer agreements can be established to cover a period of several decades. This provides each party long-term assurance of the terms and conditions of water availability. Such long-term agreements can help an urban water utility modify release rules for reservoir storage to maintain less drought storage than would otherwise be desired or reduce the need for new source development. Long-term arrangements also can provide flexibility where future water demands may not meet expectations. However, long-term leasing of water does entail risk for water buyers if water demands meet or exceed current forecasts. Long-term leasing or contingent contracts allow water-right owners to retain long-term investment flexibility in anticipation of potentially greater future values for water leasing or sale of a water right.

Intermediate-term (3 to 10 year) contingent-transfer contracts might be employed to help reduce the susceptibility of the buyer's system to drought during periods prior to the construction or acquisition of new supplies. Short-term (1 to 2 year) contingent-transfer contracts might be utilized in the midst of a drought by a water agency with depleted storage, preparing for the possibility that the drought might last a year or two longer. This type of short-term contingent-transfer contract would enable the buyer to have committed water supplies when their system might be extremely vulnerable.

Advantages of contingent transfers for the seller, typically agricultural interests, are the immediate infusion of cash when the contract is made, the infusion of additional revenues if the contingent-transfer option is called, and an increased ability to predict the conditions and timing of transfers, rather than relying on the vagaries of timing, price, and quantity of a water spot market.

The potential sale of water by farmers during drought affects the need for groundwater management (if available as an alternate supply of water) and the special operation of conveyance and storage facilities. The ability of farmers to sell water might also affect the operation-rule curves used by agricultural water suppliers for allocating water from storage to farmers over multiyear droughts. Perhaps additional hedging or overyear storage by agricultural water suppliers will increase farm incomes more than adherence to current reservoir operating rules, by creating a greater scarcity of water and higher water-transfer incomes during drought years. Similar issues relate to the overyear use of groundwater storage.

4.8.3 Spot Market Transfers

It may be desirable for the agency to allow short-term transfers in a water-rental or water-futures market. Spot market transfers are short-term transfers or leases. Typically spot market transfers are agreed to and completed within a single water year. However, for large systems, there is a possibility to establish spot futures markets for water for seasonal or overyear periods.

Spot transfers may garner improvements in economic efficiency beyond those from long-term permits. Eheart and Barclay (1990) and Wong and Eheart (1983) estimated that for irrigation such improvements are small (less than 2 percent), as long as trading of long-term permits is allowed. For other types of uses, they could conceivably be more significant, but the seasonal variation of the value of water is usually greater for irrigation than other types of water use.

Spot market transfers are typically established by some sort of bidding process, often with some of the conditions for transfer being fixed (e.g., price and quantity). However, spot market transfers can arise from negotiations between individuals or groups of buyers and sellers. A wide variety of bargaining rules for the operation of spot markets have been examined on a theoretical basis and through the use of simulation (Saleth et al., 1991). These results illustrate the importance of bargaining rules when the numbers of buyers and sellers are small, less than about a dozen participants. For large spot markets, the effects of particular bargaining rules are quickly overshadowed by competition among buyers and sellers.

Spot market purchases can be advantageous in both dry or wet years. During periods of drought, short-term transfers may be sought to directly meet demands, especially demands still not met after implementation of drought water-conservation and traditional supply-augmentation measures. As with permanent transfers, temporary transfers used to meet demands directly can have the advantage of delaying or avoiding the costs of developing new supply sources or implementing more stringent demand-management measures.

In wet years, water purchased through a spot market can be stored in reservoirs or aquifers as overyear storage. This enhances the yield of the system during drought years by increasing the amount of stored water available upon entering a drought. Overyear storage of transferred water is particularly well suited to acquiring water from junior water-rights holders. Junior water rights are typically less expensive than senior water rights, although they may only be available during relatively wet years. However, storage of transferred water during wet years may require additional surface or groundwater-storage capacity, and is subject to evaporative and seepage losses and any costs associated with storage. This approach may also work for within-year storage.

4.8.4 Water Banks

Water banks are a relatively constrained form of spot market operated by a central banker. Here, users sell water to the bank for a fixed price and buy water from the bank at a higher fixed price. The difference in prices typically goes to covering the bank's administrative and technical costs. Each user's response to the bank and involvement in the market is largely restricted to the *quantity* of water it is willing to buy or sell at the fixed price.

The California Drought Emergency Water Banks, beginning in 1991, are examples of water banks or spot markets where the terms and price of transfer were relatively fixed, with the state acting as a banker (California DWR, 1992; Howitt et al., 1992). A similar, but smaller water bank was established in Solano County, California (Lund et al., 1992). In agricultural regions, it is common for water banks or pools to exist within large irrigation systems. For many existing water pools, sellers avoid only the cost of purchasing unneeded water from the system. Water buyers in these pools pay the system normal wholesale water prices, plus some administrative cost (National Research Council, 1992; Gray, 1990).

Where spot market or water bank transfers have become established, as in California, agencies of all types are likely to plan on these markets being available for either buying or selling water (Lund et al., 1992; Israel and Lund, 1995). The existence of spot markets and water banks during droughts provides incentives for urban water suppliers to rely somewhat less on more expensive forms of conventional water-supply capacity expansion and urban water conservation in planning, and also may encourage different designs for new facilities and modified operation of existing facilities. For agricultural water districts, the existence of water banks and spot markets during drought has implications for the wording of water-supply contracts and the management of water and cropland during a drought.

4.8.5 Wheeling and Exchanges

In the electric-power industry, power is often wheeled through the transmission system between power companies and electric generation plants to make power less expensive and more reliable. Water can similarly be wheeled or exchanged through water-conveyance and storage facilities to improve water-system performance. Again, such movements of water involve the institutional transfer of water among water users and agencies. There are a number of forms of wheeling water or water exchanges (Lund et al., 1992).

Sometimes the cost of conveying water or the losses inherent in water conveyance can be reduced by wheeling water through conveyance and storage systems controlled by others. An example would be the use of excess capacity in a parallel lined canal owned by another agency, rather than use an agency's own unlined canal to convey water. Differences in pumping efficiencies might also motivate operational wheeling between conveyance facilities. Similar considerations might apply to decisions on where to store water during a drought when different reservoirs have different seepage or evaporation rates (Kelly, 1986) or if the distribution of hydropower heads is considerable for different storage options.

Seasonal wheeling of water is common in agricultural regions where different subareas have complementary demands for water over time. This can provide opportunities for one water user to exchange water to another user during the low-demand season, with repayment coming in the form of additional water during the user's high-demand season.

Also, by paying farmers not to use their rights to water, the consumptive use foregone becomes available for in-stream demands downstream. This mechanism is particularly applicable to riparian rights which cannot be legally transferred for use away from the riparian lands (Lund et al., 1992). Another application of wheeling to meet environmental constraints could involve the use of storage facilities to release water when desired for in-stream flows while meeting demands before this time from other reservoirs or groundwater.

In many cases, historical happenstance has left agricultural users with rights to high-quality water for irrigation while new urban development is left with remaining water sources of lesser quality. In such cases the additional costs of treating low-quality water for urban use is usually much greater than the costs from slightly lower crop yields from use of the lower-quality water. Given reasonable conveyance costs, it therefore becomes desirable for water-quality-based trades between agricultural and urban users. Urban users can often afford to make these trades on an uneven basis, trading more low-quality water for less high-quality water or providing a monetary inducement for a volumetrically even trade of water. Lesser-quality waters might also be traded for environmental uses of aquifer recharge or habitat maintenance (Lund et al., 1992).

4.8.6 Transfer of Reclaimed, Conserved, and Surplus Water

Although not always recognized as such, the purchase of water made available by reclamation or reductions in water demand is a form of water transfer. Numerous urban water utilities have become involved in purchasing water back from their retail customers. Such schemes usually involve rebates to customers for installing low-flow toilets or removing relatively water-intensive forms of landscaping (California DWR, 1988). Some cities have developed clever schemes where water transfers are made within their customer base. For instance, Morro Bay, California has a program whereby developers can receive water utility hook-up permits if they cause a more than equivalent reduction in existing water demand through plumbing retrofits, landscaping, or other measures (Laurent, 1992).

Urban areas have taken an interest in financing the conservation of irrigation water to make additional water available for urban supplies. This has primarily been accomplished through the lining of irrigation canals. For example, the transfer of water between the Imperial Irrigation District (IID) and the Metropolitan Water District of Southern California (MWD) involves a 35-year contract for MWD payments for canal lining and other system improvements in IID's irrigation infrastructure in exchange for the water saved by these improvements. The savings are estimated at 123.3 Mm^3/y (100,000 acre $\cdot$ ft/y) from IID's Colorado River water supplies (Gray, 1990; Sergent, 1990). This approach can have additional benefits where agricultural seepage and drainage water has led to water-quality problems or high water tables, but can create additional problems where canal seepage is used to recharge groundwater.

4.9 IMPLEMENTING WATER TRANSFERS

Perhaps the most important implication of water-transfer planning is the need to increase integration and cooperation among diverse water users. Since for economic reasons most water for water transfers will probably come from agricultural users and much of this water will go to urban and perhaps environmental users, any planning for water transfers implicitly integrates urban, agricultural, and environmental water supplies. As the tendency to seek and implement water transfers continues, it will become less possible, and less desirable, for individual urban or agricultural water districts or regions to plan and operate their water supplies independently. This necessary coordination of planning and operations between functionally diverse water agencies will imply potentially protracted and probably controversial negotiations, at least for long-term transfer arrangements.

Additionally, if intersectoral and interregional water transfers are to become significant long-term components of water-resources planning, they must be integrated with traditional water-supply augmentation and demand-management measures. Given the complex nature of many water-resource systems and the wide variety of different possible water-transfer designs, it seems apparent that some form of water-supply-system computer modeling will be required to achieve this integration of water transfers with other water-management measures.

Most major water-supply agencies already possess significant conventional water-modeling capability. However, most models are specific to individual water systems, in accordance with the needs of traditional water-supply and water-conservation measures which can be implemented by a single system. The integration of water transfers will likely require significant modifications to these single-system models to allow explicit examination of long- and short-term water transfers and exchanges. Water transfers also encourage more explicit consideration of the economic nature of water-supply operations in system modeling. System models for examining water-transfer options together with supply source and water-conservation expansions and modifications might usefully provide economic measures of performance (component and net costs) in addition to traditional technical measures of performance (e.g., yields and shortages). Various agencies and academic researchers have already begun such efforts (Lund and Israel, 1995a; Smith and Marin, 1993).

The economic nature of the design of water transfers and their integration with other water-supply-management measures encourages the use of optimization models, where the model itself suggests promising combinations of water transfers, construction, and water conservation. While technically more difficult and still somewhat inexact, optimization modeling can aid in identifying promising solutions, which can then be examined in more detail with simulation models. Performance of economically based optimization (or simulation) of water-resource systems with water transfers requires technical studies estimating the value of and the willingness to pay for different water uses and different water quantities.

During California's recent drought in which water transfers were actively pursued and implemented, both traditional supply infrastructure and demand-management strategies continued to have important, albeit modified, roles in water management (Israel and Lund, 1995). Some hints of how the integration should take place and several specific areas of concern for implementing water transfers are discussed below.

4.9.1 Legal Transferability of Water

The legal transferability of water is a major consideration in designing water transfers. Legislation pertaining to the transferability of water will vary between states and can vary within a given state over time as a state's water law evolves. California has strong statutory directives to promote water transfers (Gray, 1989; Sergent, 1990), yet legal constraints still pose a significant threat to water-transfer activity. Legal considerations are particularly important when a proposed transfer involves changes in conditions stipulated by the original water right, such as changes in type of use, place of use, or timing of withdrawals. The type of water right to be transferred is also an important consideration. Riparian rights, for instance, are generally nontransferable from their initial location of use, and the transferability of groundwater rights varies substantially by state.

Also, different types of water contracts impose different transferability requirements. In California, many water contracts stipulate that any water not used by the contractor reverts to the contractee, while others may stipulate that water cannot be transferred outside of a district and can only be transferred within a district at cost. These types of provisions reduce the ability and incentive of contractors to sell surplus or conserved water (Sergent, 1990; Gray, 1989). Short-term, emergency water transfers may be able to gain relatively easy approval and rapid implementation, given sufficient flexibility in the conveyance and storage system and sufficient professional flexibility and readiness on the part of water managers. Legislation often exists which reduces or eliminates barriers to transfers during drought or other emergency conditions. This was certainly the case for the 1991 and 1992 California Drought Emergency Water Banks (California DWR, 1992). On the other hand, long-term, planned transfers, such as dry-year option contracts and permanent water transfers typically face more difficult legal and economic constraints. Many of the longer-term transfers that require the storage of surplus water during wet years also involve complex legal issues (Getches, 1990), particularly for groundwater storage (Kletzing, 1988). The costs, delays, and risks involved in overcoming these constraints can induce agencies not to consider or participate in water transfers.

4.9.2 Real versus Paper Water

Where water transfers are motivated by real water shortages, the transfer of water by contract must correspond closely with the transfer of water in the field. This is sometimes known as the distinction between real and paper water. Associating quantities of paper water to real water is a difficult technical problem. In the case of transfers from farms, farmers typically do not know with certainty how much water they use or how much real water would become available if land were to be fallowed or cropping patterns altered (Ellis and DuMars, 1978). Even where such flow measurements are made, they are often inexact.

As water moves through a complex conveyance and storage system, there are seepage and evaporation losses, withdrawals by or return flows from other users, and natural accretions downstream. All these factors complicate the estimation of how much water is physically available to the receiver of a water transfer, given that the sender has relinquished use of a given amount. Another problem with linking paper water to real water is establishing the hydrologic independence or interdependence of water sources. This is a common problem where pumped groundwater may induce recharge from nearby surface water.

Particularly where there are many potential buyers and sellers of water, there would seem to be some need for standards or governmental involvement in tying real water to paper water transfers (Blomquist, 1992). Without such standard accounting, amounts of paper water are likely to exceed amounts of wet water available, leading to excessive withdrawals by water users to the detriment of downstream users and those not party to transfers. This will be true for transfers of water for both consumptive and in-stream uses. Litigation and calls for greater regulation of water transfers would be the likely result.

4.9.3 Conveyance, Storage, and Treatment

The mere purchase of water is usually insufficient to effect a water transfer. Transferred water must typically be conveyed and pumped to a new location, often stored, and commonly treated. Since both emergency short-term transfers and long-term transfers may require modifying the operation of existing water infrastructure, considerable work may be required to coordinate the use of conveyance, storage, and treatment systems. This can be particularly challenging because these facilities are often designed for very different operations. Occasionally, canals must be run backwards, water must flow backwards through pumps, and treatment plants must treat waters of a quality different from their design specifications. Construction of additional conveyance interties or other facilities may be required in some cases.

The difficulties encountered by San Francisco illustrate well the traditional engineering limitations and concerns with the use of water transfers in system operations and planning (Lougee, 1991; Lund et al., 1992). San Francisco purchased 62 Mm^3 (50,000 acre·ft) from the 1991 California Drought Water Bank, but their water-treatment plant was unable to accept more than a limited rate of transferred water from the Sacramento–San Joaquin Delta. Delta water is of lower quality than San Francisco's normal Sierra supply and the mixing of waters in the treatment plant beyond certain ratios increased the likelihood of trihalomethane formation. This limitation forced much of the transferred water to be stored in state-owned facilities and slowly released into San Francisco's treatment plant. California's East Bay Municipal Utility District (EBMUD) faced similar quality limitations on the treatability of transferred water which, combined with other difficulties in effecting transfers, led EBMUD not to use transferred water and to rely more on urban water-conservation measures.

Water transfers are likely to be more successful in regions with an extensive system of conveyance and storage facilities and well-coordinated operations. Locations with restricted conveyance and storage infrastructure are likely to have limited potential for effecting water transfers unless creative operations or new conveyance and storage facilities can be developed. The coordination and physical completion of water transfers will be more difficult, and perhaps impossible, if agencies controlling major components of a region's water conveyance and storage system choose not to participate in transfers, are legally restrained from participating, or participate only in a limited way.

4.9.4 Contracts and Agreements

The legal transfer of water is typically effected by contracts which must specify a number of logistical and financial conditions of the transaction. Among the logistical and fiscal details that must be specified are: the location and timing of water pick-up from the seller; the fixed or variable price of the water; the fixed or variable quantity of water; and potentially the quality of the water. The responsibilities for contract execution and liabilities for failure to completely execute the contract might also be included.

Where transferred water cannot be conveyed directly between the buyer and seller, agreements are often required with other entities, either to make use of their conveyance facilities (pumps or aqueducts) or to coordinate the conveyance of transferred water through natural waterways, within environmental limitations (Lougee, 1991). Similarly, facilities owned or operated by entities not directly involved in the transfer may be necessary to store transferred water until it can be used. This will often require agreements or contracts for the storage of water with agencies which oversee storage facilities. When water is stored in aquifers, recharge and pumping facilities will be required, and legal arrangements with overlying landowners are common (Kletzing, 1988). Likewise, contractual arrangement may be required for the treatment of transferred water.

4.9.5 Price, Transaction Costs, and Risks

As demonstrated by the 1991 and 1992 California Drought Water Banks, both sellers and buyers can be quite sensitive to the price established for water (Lund et al., 1992). At lower prices, there are fewer willing sellers and greater demand for water from agricultural users. Higher prices encourage sellers but tend to exclude most potential agricultural buyers. The price set by the market, through negotiations, or by a governmental water bank has important implications for the character and number of resulting transfers.

The cost of water to a user includes more than its purchase price. As noted above, much of the work in establishing successful transfers of water lies in arranging for the conveyance, storage, and perhaps treatment of the transferred water. In some cases, the costs of these activities may exceed the cost of the water itself. For example, in 1991 San Francisco purchased 18.5 Mm^3 (15,000 acre · ft) from Placer County at a price of \$36,500/Mm^3 (\$45/acre · ft). However, total costs including wheeling charges through state and federal facilities and storage costs were between \$200,000 and \$300,000/Mm^3 (\$250 to \$350/acre · ft). Also, the final delivery cost of water purchased by San Francisco from the Water Bank was nearly double the purchase price of \$142,000/Mm^3 (\$175/acre · ft). Water transfers are also subject to numerous other transaction costs, including legal fees, costs of public agency review, costs of required technical studies, and costs involved in settling claims from third parties. MacDonnell's survey (1990) found that transaction costs averaged several hundred dollars per acre · ft of transferred perpetual water right, with averages of \$309,000/Mm^3 (\$380/acre · ft) of perpetual right in Colorado and \$150,000/Mm^3 (\$184/acre · ft) in New Mexico. These transaction costs can add substantially to the purchase price of water, which in Colorado and New Mexico ranges from \$200,000 to \$1.2 million per Mm^3 (\$243 to \$1,500 per acre · ft). The unit costs for transactions commonly decrease for larger transfers and increase with the controversy of a transfer. Still, transaction costs are highly variable between transfers.

The risks of a transfer not being completed may also dissuade potential partners in transfers. The risk of a proposed transfer being stopped entirely is particularly palpable where a substantial part of the transaction costs must be expended before a transfer agreement is finally approved, or if there are high costs to delaying implementation of other water-supply alternatives while transfers are being negotiated. This would be the case where large expenditures for technical and legal work must be made before final approval of a transfer is in place (Lund, 1993).

4.9.6 Evaluation of Impacts to Third Parties

Evaluating the third-party impacts of water transfers can be formidable and inexact, involving difficult ecological and economic studies (National Research Council, 1992). There is currently little technical work quantifying physical, environmental, economic, and social impacts from water transfers (Howe et al., 1990;

Agricultural Issues Center, 1993). Less is known about how these impacts would vary with different specific transfer cases and mechanisms and how effective different approaches to mitigating third-party impacts might be.

Some of the technical issues in managing third-party impacts are illustrated by the case of Yolo County, California. Farms in Yolo County contributed about 185 Mm^3 (150,000 acre · ft) of water to the 1991 California Drought Water Bank. Some of this water came from fallowing farmland and transferring the surface-water rights. However, most of the surface water was replaced by increased groundwater pumping. Yet the county does not employ a water engineer or groundwater specialist dedicated to countywide water-supply problems who could assess and manage the long-term impacts of these transfers (Jenkins, 1992). There is also little legal authority for counties to assume this role. Furthermore, rural county governments may lack expertise to estimate the economic impacts of different types of transfers. Without an understanding of the economic and physical effects of water transfers, water-exporting regions are likely to be suspicious of and somewhat resistant to water transfers.

This same lack of a technical basis for assessing and managing impacts of water transfers takes on a more important role at the statewide level where water-transfer policies are made. Technical studies are needed to support policies and perhaps specific cases should be investigated of when and how water transfers are made and how any third-party impacts should be managed (Howitt et al. 1992). Of course, as noted earlier, there is a possibility of avoiding certain kinds of third-party impacts by defining water rights in terms of consumption.

4.9.7 Roles for Government in Water Transfers

The role of state and federal government is so important in many cases that it must be considered part of the system engineering. In California, for instance, a significant part of state and federal involvement in water transfers is due to the technical role required by their ownership and operation of major conveyance and storage facilities and their requirements and responsibilities under various environmental regulations.

A number of roles for federal, state, and local governments can be identified for facilitating water transfers, some of which may require modification of existing regulations, legislation, and local agency enabling legislation. Perhaps the most appropriate role for government in water transfers is that of an arbiter of technical and third-party disputes and a regulator of the market. This role is needed to ensure a close tie between trades of paper water and real water and the coordination of the movement of transferred water with environmental regulations (Blomquist, 1992). State or regional governments would also seem to have a useful role in the collection and analysis of data for monitoring and resolving external and third-party impacts. Regional governments can also act as bankers in the formation of regional water markets, taking advantage of the regional hierarchy of governmental water jurisdictions commonly found in water management.

Government involvement can improve the prospects for water transfers by:

1. Improving information regarding transfers and transfer impacts
2. Establishing a process for managing third-party impacts
3. Reducing the transaction costs of arranging and implementing water transfers
4. Increasing the probability that efforts between parties to arrange a water transfer will be successful and reducing the risks to parties from involvement with transfers

Of course, as noted previously, government must maintain the registry of permits and may act as broker (but need not necessarily). Individual agencies have their own agendas and will continue to pursue short- and long-term contracts regardless of the existence of government-sponsored water banks. However, government involvement can greatly accelerate the development of water-transfer agreements by initial sponsorship of transfers through the establishment of water banks or by other means. The development of transfers as part of a larger water-resource system is likely to continue after government sponsorship of water banks has ended.

4.10 SUMMARY AND IMPLICATIONS FOR REGULATORY PROGRAMS AND WATER TRANSFERS

Against the backdrop of the sections on water regulatory programs, it is possible to single out several options that currently appear more attractive than others and that should be considered for a water-permits program in a humid region. This is not to be construed as a recommendation, but a suggestion that these options are appropriate to consider first, pending further research and public dialogue for each specific situation.

The flexible-permit basis seems to be more consistent, on an equity basis, with the riparian doctrine than prioritized permits or nonpermit approaches such as ad hoc restrictions, subsidies, or charges. The western first-in-time rule for prioritization may not be legally feasible in riparian areas. Fractional permits are also more consistent with legal precedents for groundwater regulation. The constant use basis for permits is simply a physically intractable means of allocating water from a watercourse whose flow rate fluctuates with time.

One way or another, the agency must grapple with the equity issue of deciding how the resource must be distributed among the participants. The historical precedent or riparian equivalent bases are administratively easier and will probably be regarded as more equitable than either the historical-use or ad hoc bases, even though the latter may be more flexible.

It should be possible to set up a system of regulation that requires little administrative effort until supplies become constraining and that automatically invokes controls at that point. In areas where demand is currently light compared to supply, users could be issued permits which entitle them to their current normal withdrawal. By issuing the same number of new permits of limited duration every year, the agency could eventually achieve a staggered permit system.

Given the host of potential problems associated with alternative approaches, the upstream-gauged, fixed pass-through system seems to be the most administratively tractable method for addressing the complexity of river systems. It enables the use of fractional permits without ambiguity of individual users' allowable withdrawal. The disadvantage of this approach is that it requires a larger number of gauges and therefore entails a higher administrative cost.

Averaging periods could be set at any level initially, as long as there is a proviso that they may be lowered later. To prevent substantial streamflow depletion, averaging periods for free-flowing streams should usually be from hours to a week. Lakes and reservoirs could use a longer averaging period, perhaps months, depending on the size of the water body. Aquifer averaging periods could be as large as a year or more. The agency should consider staggering the time of calculation of averages over users so that the incentives for increased pumping near that time do not occur simultaneously.

Allowing transfers of water rights among users has advantages and disadvantages. The principal advantage is that economic efficiency may be improved by voluntary trading and the allowance of such trading is considered equitable by the users. This improvement in efficiency is robust to data errors and does not require data collection or planning by the agency. The principal disadvantage is that the agency relinquishes some control over where withdrawals occur. This may undermine its effectiveness in protecting streamflows and aquifers and may lead to some third-party impairment problems. Most third-party problems are solved when permits are defined in terms of consumption, however.

Water transfers have far-reaching implications for water-resource planning and management. In addition to contributing to the bag of tricks available to water managers, transfers require a broader conceptualization of water-management problems. Unlike traditional supply-augmentation and demand-management measures, which can typically be accomplished by a single water agency, water transfers require coordinated planning and operations between both groups party to the transfer. Also, water transfers often require the use of storage and conveyance facilities belonging to or operated by entities not directly involved in the buying or selling of water. The evaluation of transfers demands a more explicit economic perspective on the purposes of water-resource systems and more detailed economic measures of operation performance. The water acquired by transfers can serve a variety of operational, environmental, and economic purposes. Overall, the multiple forms of water transfers and their flexibility, combined with legal, third-party, and technical issues in implementing transfers, make water transfers one of the more promising, yet complex techniques for improving water management.

As traditional forms of water-resource development become more difficult and expensive, the profession must turn to the management of water use, including the management of water allocations and water demands. This chapter has reviewed the use of permit systems and water-transfer systems for managing water allocations. Notwithstanding significant difficulties, both approaches have been increasingly employed in recent years and show promise for the future.

4.11 Acknowledgments

The authors gratefully acknowledge the support of the Energy and Environmental Affairs Division of the Illinois Department of Energy and Natural Resources under Project SENR WR-26, "Analysis of Options for Water Use Regulation for the State of Illinois," for the study of permit systems and funding provided by the U.S. Army Corps of Engineers Hydrologic Engineering Center and University of California Water Resources Center for the study of water-transfer systems. While the authors assume full responsibility for any errors or omissions, appreciation is extended to the following students and colleagues who have helped conduct this research or whose viewpoints have contributed to its planning and the interpretation of its results: John Barclay, Jean Bowman, John Braden, Gary Clark, Christopher Enright, Inês Ferreira, Paul Hutton, Morris Israel, Mimi Jenkins, Jon Liebman, Karen Miller, Rezaur Rahman, Robert Reed, Maria Saleth, Maureen Sergent, other current and former students, and the many professionals interviewed for this research for their comments on work on these subjects. The findings reported in this paper are those of the authors and do not necessarily reflect the views of the Illinois Department of Energy and Natural Resources or the U.S. Army Corps of Engineers.

4.12 REFERENCES

Agricultural Issues Center, "California Water Transfers: Gainers and Losers in Two Northern Counties," University of California, Davis, 1993.

Anderson, T. L., ed., *Water Rights*, Pacific Institute for Public Policy Research, San Francisco, 1983a.

Anderson, T. L., *Water Crisis*: *Ending the Policy Drought*, Cato Institute and Johns Hopkins University Press, Baltimore, Md., 1983b.

Bandow, D., *Critical Issues—Protecting the Environment: A Free-Market Strategy*, The Heritage Foundation, Washington, D.C., 1986.

Baumol, W. J., and W. E. Oates, *The Theory of Environmental Policy*, Prentice-Hall, Englewood Cliffs, N.J., 1988.

Bliss v. Kennedy (43 Ill. 67) 1867.

Blomquist, W., *Dividing the Waters: Governing Groundwater in Southern California*, ICS Press, San Francisco, 1992.

Brajer, V., A. L. Church, R. Cummings, and P. Farah, "The Strengths and Weaknesses of Water Markets as They Affect Water Scarcity and Sovereignty Interests in the West," *Natural Resources Journal*, 29(Spring): 489–509, 1989.

California Action Network, "Open Letter from California Action Network Regarding Water Marketing and Public Policy," Davis, Calif., August 11, 1992.

California Department of Water Resources, *Landscape Water Conservation Guidebook No. 8*, Sacramento, Calif., 1988.

California Department of Water Resources, *The 1991 Drought Water Bank*, Sacramento, Calif., 1992.

Clark, G. R., "Illinois Groundwater Law[:] The Rule of 'Reasonable Use,'" paper prepared by the State of Illinois, Department of Transportation, Division of Water Resources for presentation to the Illinois Groundwater Association semiannual meeting, Joliet, Illinois, October 8, 1985.

Cravens, S. J., S. D. Wilson, and R. C. Barry, "Summary Report: Regional Assessment of the Ground-Water Resources in Eastern Kankakee and Northern Iroquois Counties," SWS Contract Report 456, by the Illinois State Water Survey to the Illinois Department of Energy and Natural Resources and the Illinois Department of Transportation Division of Water Resources, April 1989.

Coase, R. H., "The Problem of Social Costs," *Journal of Law and Economics,* 3:1–44, 1960.

Dixon, W. D., and W. E. Cox, "Minimum Flow Protection in Riparian States," *Journal of Water Resources Planning and Management,* ASCE, 111 (2):149–156, 1985.

Eheart, J. W., "Programs for Management of Water Withdrawals in Illinois," Final Report, Project SENR WR-26 Illinois Department of Energy and Natural Resources, Energy and Environmental Affairs Division, Springfield, Ill., 1989.

Eheart, J. W., and J. P. Barclay, "Economic Aspects of Groundwater Withdrawal Permit Transfers, *Journal of Water Resources Planning and Management,* ASCE, 116(2):282–303, Mar/Apr 1990.

Eheart, J. W., and R. M. Lyon., "Alternative Structures for Water Rights Markets." *Water Resources Research* 19(4):887–894, 1983.

Ellis, W. H., and C. T. DuMars, "The Two-Tiered Market in Western Water," *Nebraska Law Review,* 57(2):333–367, 1978.

Enright, C., and J. R. Lund, "Alternative Water District Organization: A Screening Level Analysis," *Journal of Water Resources Planning and Management,* ASCE, 117(1):86–107, 1991.

Florida Water Resources Act, Florida Statutes, Chap. 373, 1972.

Getches, D. H., *Water Law,* West Publishing, St. Paul, Minn., 1990.

Goering, Laurie, and Thomas Cekay, "Irrigation Pulls Plug on Residents' Water," Chicago *Tribune,* July 25, 1988.

Gray, Brian E., "A Primer on California Water Transfer Law," *Arizona Law Review,* 31:745–781, 1989.

Gray, Brian E., "Water Transfers in California: 1981–1989," in MacDonnell, Lawrence J. (principal investigator), "The Water Transfer Process as a Management Option for Meeting Changing Water Demands," vol. II, USGS Grant Award No. 14-08-0001-G1538, Natural Resources Law Center, University of Colorado, Boulder, 1990.

Hahn, R. W., "Market Power and Transferable Property Rights," *The Quarterly Journal of Economics,* 99(4):754–765, 1984.

Harrison, K. W., "The Allocation of Flowing Water in Humid Regions," M.S. thesis, University of Illinois, Urbana, Ill., 1991.

Hartman, L. M., and D. Seastone, *Water Transfers: Economic Efficiency and Alternative Institutions,* Johns Hopkins Press, Baltimore, Md., 1970.

Howe, C. W., J. K. Lazo, and K. R. Weber, "The Economic Impacts of Agriculture-to-Urban Water Transfers on the Area of Origin: A Case Study of the Arkansas River Valley in Colorado," *American Journal of Agricultural Economics,* Dec.:1200–1204, 1990.

Howe, C. W., D. R. Schurmeier, and W. D. Shaw, Jr., "Innovative Approaches to Water Allocation: The Potential for Water Markets," *Water Resources Research,* 22(4):439–445, 1986.

Howitt, R., N. Moore, and R. T. Smith, "A Retrospective on California's 1991 Emergency Drought Water Bank," March, California Department of Water Resources, Sacramento, Calif., 1992.

Israel, M. and J. R. Lund, "Recent California Water Transfers: Implications for Water Management," *Natural Resources Journal,* 35:1–32, Winter 1995.

Jenkins, Mimi, "Yolo County California's Water Supply System: Conjunctive Use without Management," Masters report, Department of Civil and Environmental Engineering, University of California, Davis, Calif., 1992.

Johnson, R. N., M. Gisser, and M. Werner, "The Definition of a Surface Water Right and Transferability," *Journal of Law and Economics,* 24:273–288, Oct. 1981.

Kelly, K. F., "Reservoir Operation During Drought: Case Studies," Hydrologic Engineering Center, U.S. Army Corps of Engineers, Research Document no. 25, Davis, Calif., 1986.

Kletzing, R., "Imported Groundwater Banking: The Kern County Water Bank—A Case Study," *Pacific Law Journal,* 19(4):1225–1266, 1988.

Laurent, M. L., "Overview New Development Process/Water Allocations/Conservation," City of Morro Bay, Calif., 1992.

Linsley, R. K., and J. B. Franzini, *Water-Resources Engineering,* Chap. 6, McGraw-Hill, New York, 1972, pp. 148–160.

Little, R. L., and T. R. Greider, *Water Transfers from Agriculture to Industry: Two Utah Examples,* Research Monograph no. 10, Institute for Social Science Research on Natural Resources, Utah State University, Logan, Utah, 1983.

Lougee, N. H., "Uncertainties in Planning Inter-Agency Water Supply Transfers," in J. L. Anderson, ed., *Water Resources Planning and Management and Urban Water Resources,* American Society of Civil Engineers, New York, 1991, pp. 601–604.

Lund, J. R., "Metropolitan Water Market Development: Seattle, Washington, 1887–1987," *Journal of Water Resources Planning and Management,* ASCE, 114(2):223–240, 1988.

Lund, J. R., "Transaction Risk versus Transaction Costs in Water Transfers," *Water Resources Research,* 29(9):3103–3107, 1993.

Lund, J. R., and M. Israel, "Optimization of Water Transfers in Urban Water Supply Planning," *Journal of Water Resources Planning and Management,* ASCE, 121(1):41–48, 1995a.

Lund, J. R., and M. Israel, "Water Transfers in Water Resource Systems," *Journal of Water Resources Planning and Management,* ASCE, 121(2):193–204, 1995b.

Lund, J. R., M. Israel, and R. Kanazawa, "Recent California Water Transfers: Emerging Options in Water Management," Center for Environmental and Water Resources Eng. Report 92-1, Dept. of Civil and Env. Eng., University of California, Davis, 1992.

Lyon, R. M., "Auctions and Alternative Procedures for Allocating Pollution Rights," *Land Economics,* 58(1):16–32, 1982.

Maass, A., and R. Anderson, . . . *And the Desert Shall Rejoice: Conflict, Growth, and Justice in Arid Environments,* MIT Press, Cambridge, Mass., 1978.

MacDonnell, L. J. (principal investigator), "The Water Transfer Process as a Management Option for Meeting Changing Water Demands," vol. I, USGS Grant Award No. 14-08-0001-G1538, Natural Resources Law Center, University of Colorado, Boulder, 1990.

Mack, L. E., and A. W. Peralta, "Water Allocation: Benevolent Czar or Crystal Pitcher Approach," in *Proceedings of the Seventeenth Mississippi Resources Conference, March 25–27, 1987,* Arkansas Water Research Center, Fayetteville, Ark., 1987.

Milliman, J. W., "Water Law and Private Decision-Making: A Critique," *Journal of Law and Economics,* 2(Oct.):41–63, 1959.

National Research Council, *Water Transfers in the West: Efficiency, Equity, and the Environment,* National Academy Press, Washington, D.C., 1992.

Nunn, S. C., and H. M. Ingram, "Information, the Decision Forum, and Third Party Effects in Water Transfers," *Water Resources Research,* 24(4):473–480, 1988.

Revesz, R. L., and D. H. Marks, "Local Irrigation Agencies," *Journal of the Water Resources Planning and Management Division,* ASCE, 107(WR2):329–338, 1981.

Saarinen, P. P., and G. D. Lynn, "Getting the Most Valuable Water Supply Pie: Economic Efficiency in Florida's Reasonable-Beneficial Use Standard," *Journal of Land Use and Environmental Law,* 8(2):491–520, 1993.

Saleth, R. M., J. B. Braden, and J. W. Eheart, "Bargaining Rules for a Thin Spot Water Market," *Land Economics,* 67(3):326–339, 1991.

Sergent, M. E., "Water Transfers: The Potential for Managing California's Limited Water Resources," Masters thesis, Civil Engineering Department, University of California, Davis, 1990.

Smith, M. G., and C. M. Marin, "Analysis of Short-Run Domestic Water Supply Transfers Under Uncertainty," *Water Resources Research,* 29(8):2909–2916, 1993.

Stavins, R. N., "Harnessing Market Forces to Protect the Environment," *Environment,* 31(1):5, 1989.

Tietenberg, T. H., *Emissions Trading: An Exercise in Reforming Pollution Policy,* Resources for the Future, Washington, D.C., 1985.

Tregarthen, T. D., "Water in Colorado: Fear and Loathing of the Marketplace," in T. L. Anderson, ed., *Water Rights,* Pacific Institute for Public Policy Research, San Francisco, 1983.

U.S. Congress, "Project 88—Harnessing Market Forces to Protect our Environment: Initiatives for the New President," a public policy study sponsored by Senators Timothy E. Wirth of Colorado and John Heinz of Pennsylvania, December, 1988.

Virginia Water Use Act of 1991.

Walker, W. R., W. E. Cox, and M. S. Hrezo, eds., "Legal and Administrative Systems for Water Allocation and Management: Planning in the Southeastern States," OWRT B-123-VA(1), NTIS # PB187120, March 1983.

Water Science and Technology Board, National Research Council, Commission on Engineering and Technical Systems, Commission on Physical Sciences, Mathematics, and Resources, "Drought Management and Its Impact on Public Water Systems," Report on a Colloquium Sponsored by The Water Science and Technology Board, September 5, 1985, National Academy Press, Washington D.C., 1986.

Wong, B. D. C., and J. W. Eheart, "Market Simulations for Irrigation Water Rights." *Water Resources Research* 19(5):1127–1138, 1983.

P • A • R • T • 2

DEMAND AND MANAGEMENT MODELS

CHAPTER 5
WATER DEMAND ANALYSIS

Benedykt Dziegielewski
Southern Illinois University
Carbondale, Illinois

Eva Opitz
PMCL
Carbondale, Illinois

5.1 DEFINITION AND MEASUREMENT OF WATER USE

From the hydrologic perspective, *water use* can be defined as all water flows that are a result of human intervention within the hydrologic cycle. A more restrictive definition of water use refers to water that is actually used for a specific purpose. Table 5.1 contains definitions of nine such uses (Solley et al., 1998). Urban water supply systems deliver water to most of these categories of use with domestic and commercial uses being almost entirely dependent on public deliveries. Several other categories such as industrial, irrigation, and public uses are also present in urban areas. Some categories are found primarily outside of urban areas or require large quantities of untreated water, and they tend to be self-supplied.

Measurements of water use are reported as water volumes per unit of time. The volumetric units include cubic meters, cubic feet, gallons, and liters, and their decimal multiples. In some cases, composite volumetric units such as acre·foot or units of water depth such as inches of rain may be used. The time periods used include second, minute, day, and year. Because the annual volumes of water use usually involve large numbers, annual water-use totals are often reported as the average daily usage rates. Two popular units for measuring total urban demands are thousand cubic meters per day (Km^3/d) and million gallons per day (mgd).

TABLE 5.1 Major Purposes of Water Use

Water-use purpose	Definition
Commercial use	Water for motels, hotels, restaurants, office buildings, and other commercial facilities and institutions
Domestic use	Water for household needs such as drinking, food preparation, bathing, washing clothes and dishes, flushing toilets, and watering lawns and gardens (also called residential water use)
Industrial use	Water for industrial purposes such as fabrication, processing, washing, and cooling
Irrigation use	Artificial application of water on lands to assist in the growing of crops and pastures or to maintain vegetative growth in recreational lands such as parks and golf courses
Livestock use	Water for livestock watering, feed lots, dairy operations, fish farming, and other on-farm needs
Mining use	Water for the extraction of minerals occurring naturally and associated with quarrying, well operations, milling, and other preparations customarily done at the mine site or as part of a mining activity
Public use	Water supplied from a public water supply and used for such purposes as firefighting, street washing, municipal parks, and swimming pools
Rural use	Water for suburban or farm areas for domestic and livestock needs which is generally self-supplied
Thermoelectric power use	Water for the process of the generation of thermoelectric power.

Source: Adapted from Solley et al., 1998.

In order to make the estimates of water use easy to comprehend and to make meaningful comparisons of water use for various purposes (and various users), the annual or daily quantities are divided by some measures of size for each purpose of use. The result is an average rate of water use such as gallons per capita per day (gcd), gallons per employee per day (ged), or other unit-use coefficients.

The reported quantities of water use can be in the form of direct measurements obtained from water meters that register the volume of flow (such as displacement meters or venturi meters), or they may be estimates. Estimates of water use derived from the measurements of water levels in storages or from pumping logs are generally more accurate than those derived from related data on the volume of water-use activity. For example, the estimates of water use for industrial purposes may be obtained by multiplying the number of manufacturing employees or the value added (in dollars) by a water-use coefficient. For example, in 1995 the ratio of water use to value added by manufacture in the United States was 5783 gallons (gal) per $1000 (Dziegielewski et al., 2002a).

5.2 PUBLIC SUPPLY WATER USE

National data on water withdrawals for public supply purposes include public and private water systems that furnish water, year-round, to at least 25 people or that have a minimum of 15 hookups (Solley et al., 1998). Nearly 55,000 community water systems serve more than 263 million people. Transient community water systems and nontransient noncommunity water systems serve another 20 million people.

Table 5.2 shows the distribution of the community systems by the size of population served. Approximately 81 percent of the population is served by 3769 systems, which deliver water to communities with more than 10,000 persons. While nearly 80 percent of public water supply systems rely on groundwater, more than one-half (58 percent) of the larger systems use surface water as their principal source of supply (calculated from table 2.1, USEPA, 1997).

In any public water supply system, water-use records can be characterized with respect to the relative needs of various customer groups (e.g., single-family residential, hotels, food processing plants), the purposes for which water is used (e.g., end uses such as sanitary needs, lawn watering, cooling), and the seasonal variation in water use. The analysis of water use can be expanded to also include the development of information on water users in the service area. Data on housing stock, household characteristics, business establishments, and other demographic and economic statistics are important because such characteristics are major determinants of water use.

The data and techniques for analyzing water demand in a public water supply system are the subject of the following sections. Such analysis must necessarily begin with a determination of quantities of water used. While total water deliveries to urban areas can be measured at one (or several) points on the water supply system, the volumes of water used for specific purposes can be obtained through an inventory or sampling of individual users.

5.3 SAMPLING OF WATER USERS

Water supply agencies and regional or state regulatory bodies usually have the ability to monitor the water use of all users or entire classes of users. In statistical terms, studies involving all users would represent the use of entire populations. However, a complete enumeration or inventory of all users may not always capture the entire population because in addition to *populations* defined in terms of users, which can be viewed as finite (or delimited), some studies may require expanded definitions of populations. For example, a study population can be defined as the monthly water-use quantities of all users over a time horizon. Because such a definition includes future water use, historical records of withdrawals constitute only a part (or a sample) of the total population. Also, in

TABLE 5.2 Community Water Systems in the United States

System description	Number of systems	Percent of systems	Population served	Percent of population
		By system size		
500 or less	31,688	59	5,148,700	2
501–3,300	14,149	26	19,931,400	8
3,301–10,000	4,458	8	25,854,100	10
10,001–100,000	3,416	6	96,709,100	37
>100,000	353	1	116,282,800	44
Total	54,064	100	263,926,100	100
		By water sources		
Groundwater	42,661	79	85,868,500	33
Surface water	11,403	21	178,057,600	67
Total	54,064	100	263,926,100	100

Note: Groundwater systems include groundwater and purchased groundwater. Surface-water systems include surface water, purchased surface water, and groundwater under the influence of surface water.

Source: USEPA, 2000.

many cases, a study of an entire population must limit the number of measurements of each unit (due to cost constraints) and may not be capable of producing answers to some research questions. Because of these considerations, knowledge of water use is almost invariably based on samples or fragments of total populations. Sampling has many advantages over a complete enumeration (or inventory) of the population under study. These advantages include reduced cost, greater speed of obtaining information, and a greater scope of information that can be obtained. In addition, a greater precision of measurements can be secured by employing trained personnel to take the necessary measurements and analyze the data.

Scientific sampling designs specify methods for sample selection and estimation of sample statistics that follow the *principle of specified precision at the minimum cost,* i.e., they provide, at the lowest possible cost, estimates that are precise enough for the study objectives. *Probability sampling* refers to any sampling procedure that relies on random selection and is amenable to the application of sampling theory to validate the measurements obtained through sampling. This requires that, within the sampled population, one is able to define a set of distinct samples (where each sample consists of sampling units) with known and equal probabilities of being selected. One of these samples is then selected through a random process. In practice, the sample is most commonly constructed by specifying probabilities of inclusion for the individual units, one by one or in groups, and then selecting a sample of desired size and type. *Nonprobability sampling* refers to sampling procedures that do not include the element of random selection because samples are restricted only to a part of the population that is readily accessible, selected haphazardly without prior planning, or they consist of typical units or volunteers. The only way of examining how good the nonprobability sample may be is to know parameters for the entire population or to compare it with the probability sample statistics taken from the same population. For more information on sampling procedures, see Dziegielewski et al. (1993), Cochran (1963), Kish (1965), or Fowler (1988).

5.3.1 Types of Sampling Plans

There are many ways of constructing a probability sample of water users. *Simple random sampling* refers to a method of selecting n sampling units out of a population of size N, such that every one of the distinct samples (where each sample consists of n sampling units) has an equal chance of being drawn. In *stratified sampling,* the sampled population of N units is first divided into several nonoverlapping subpopulations. These subpopulations are called *strata* because they divide a heterogeneous population into homogeneous subpopulations. If a simple random sample is taken from each stratum, then the sampling procedure is described as *stratified random sampling.* In order to design a stratified random sampling plan, it is necessary to determine (1) which population characteristic (i.e., variable) should be used in stratification, (2) how to construct the strata (i.e., how many strata to use and where to set the stratification boundaries), and (3) what sample sizes should be obtained from each stratum. The statistical theory of stratified sampling offers some methods for selecting the optimal number of strata, strata boundaries, and sample sizes in advance (Cochran, 1963). However, it is usually necessary to collect and examine some data before designing a good sampling plan.

Systematic sampling is often the most expeditious way of obtaining the sample and may be used in situations where time is critically constrained. The units in the population sampled are first numbered from 1 to N in some order. To select a sample of n units, one should take the first unit at random from the first k units and every kth unit thereafter. The selection of the first unit determines the whole sample, which is often called an *every* k*th systematic sample.*

If individual sampling units are arranged according to some characteristic or variable (e.g., water use), then the systematic sample is equivalent to a stratified sample in which one sampling unit is taken from each stratum. Constructing a list of sampling units can be avoided by dividing a geographic area into subunit areas; for example, a subarea could be a county or river basin. This sampling plan, called *cluster sampling,* can result in significant cost savings. For example, a simple random sample of 600 industrial users may cover a state more evenly than 20 counties containing a sample of 30 plants each, but it will cost more because of the time devoted to travel and finding individual establishments. However, cluster sampling creates a greater risk of obtaining a nonrepresentative sample.

The *sample size* depends on the precision of measurement that is required and the variance in the parameters to be estimated. The *precision* of an estimate refers to the size of the deviations from the mean of all sample measurements obtained by repeated application of the sampling procedure. In contrast, the term *accuracy* is usually applied to indicate the deviations of the sample measurements from the true values in the population. For example, a simple random sample can be used to estimate average daily water use and variance in water use of all single-family houses in an urban area during a given year. According to sampling theory, the mean water use $\bar{y}$ obtained from the simple random sample is an unbiased estimate of the average water use $\bar{Y}$ for all houses (i.e., the population mean). Also, $\bar{Y} = N\bar{y}$ is an unbiased estimate of total water use of the population (N customers).

The *standard error* of $\bar{y}$, which describes the precision of the estimated mean value, is

$$\sigma_{\bar{y}} = \sqrt{\frac{N-n}{N}}\ \frac{S}{\sqrt{n}} \tag{5.1}$$

where S is obtained from population variance S^2 (by taking its square root). Because in practice S^2 may not be known, it must be estimated from the sample data using the following formula:

$$S^2 = \frac{\sum_{i}^{n} (y_i - \bar{y})^2}{n-1} \tag{5.2}$$

which provides an unbiased estimate of S^2 and where n is the sample size. Usually, with a population having a mean $\bar{Y}$ and a simple random sample having mean $\bar{y}$, control of the following probability condition is desired:

$$\Pr\left(\left|\frac{\bar{y}-\bar{Y}}{\bar{Y}}\right| \geq r\right) = \alpha \tag{5.3}$$

where α is a small probability (e.g., 0.05) and r is the relative error expressed as a fraction of the true population mean. By multiplying both sides of the parenthetical expression in Eq. (5.3) by $\bar{Y}$, the same condition can be restated as

$$\Pr\left(\left|\bar{y}-\bar{Y}\right| \geq r\bar{Y}\right) = \alpha \tag{5.4}$$

If, instead of the relative error r, control of the absolute error d (i.e., the absolute value of the difference between the sample mean and the population mean) in Y is desired, the formula can be written as

$$\Pr\left(\left|\bar{y} - \bar{Y}\right| \geq d\right) = \alpha \tag{5.5}$$

It is usually assumed that $\bar{y}$ is normally distributed about the population mean $\bar{Y}$, and given its standard error from Eq. (5.1), the product $r\bar{Y}$ is

$$r\bar{Y} = t\sigma_{\bar{y}} = t\sqrt{\frac{N-n}{N}} \cdot \frac{S}{\sqrt{n}} \tag{5.6}$$

where t is the value of the normal deviate corresponding to the desired confidence probability. This value is 1.64, 1.96, and 2.58 for confidence probabilities 90, 95, and 99 percent, respectively. Solving Eq. (5.6) for n gives

$$n = \frac{(tS/r\bar{Y})^2}{1 + (1/N)\,(tS/r\bar{Y})^2} \tag{5.7}$$

The expression in the denominator represents a finite population correction, and it should be used when n/N is appreciable. Without this correction, we can take the first approximation of the desired sample size n_o as

$$n_o = \left(\frac{tS}{r\bar{Y}}\right)^2 \tag{5.8}$$

According to this equation, n_o depends on the coefficient of variation (the ratio $S/\bar{Y}$) of the population that is often more stable and easier to guess in advance than S itself. It also depends on the error r that can be tolerated and the confidence level that is needed as captured by the value of t. For the absolute error specification as in Eq. (5.5), Eq. (5.8) is changed into

$$n_o = \left(\frac{tS}{d}\right)^2 \tag{5.9}$$

The preceding sample size relationships are illustrated in the following example.

5.3.2 Example of Sample Size Determination for Continuous and Proportional Data

A water supply agency serves 80,000 customers. The analysis of billing frequencies for the entire fiscal year indicates that average daily use per customer is 250 gal and the standard deviation is 180 gal. Using simple random sampling, how many customers must be sampled to be 95 percent confident of estimating average daily use within 2 percent of the true value?

Solution: $N = 80{,}000$, $S = 180$ gal, $\overline{Y} = 250$ gal, $\alpha = 0.05$, $t = 1.96$, and $r = 0.02$. Substituting these values into Eq. (5.8):

$$n_o = \frac{t^2 S^2}{r^2\, \overline{Y}^2} = \frac{(1.96)^2\,(180)^2}{(0.02)^2\,(250)^2} = 4979$$

where n = sample size (n_o is the first approximation)
N = population size
S = population standard deviation
$\overline{Y}$ = population mean
t = confidence probability (t statistic)
r = relative error

Because n_o/N is not negligible, we need to take the finite population correction, found from Eqs. (5.7) and (5.8):

$$n = \frac{n_o}{1 + n_o/N} = \frac{4979}{1 + 4979/80{,}000} = 4687$$

The results indicate that if the average water use is unknown, to be 95 percent confident of estimating it by sampling billing records with an error of 2 percent (or 5 gal), a sample of 4687 single-family homes would be required.

In some cases, it may be necessary to obtain estimates of the percent of water users who possess a certain characteristic (e.g., use groundwater as their sources of supply) by surveying a sample of users. The sampling problem in this case is referred to as *sampling for proportions,* where the respondents are classified into two classes: groundwater users and users of other sources. In order to determine the required sample size, we must decide the margin of error d in the estimated proportion p of users who rely on groundwater and the risk α that the actual error will be larger than d. Therefore, control of the following probability condition is desired:

$$\Pr\left(\left|p - P\right| \geq d\right) = \sigma \tag{5.10}$$

where P is the true proportion of users of groundwater. Assuming simple random sampling and a normal distribution of p, the standard error of p, σ_p, is given by:

$$\sigma_p = \sqrt{\frac{N-n}{N-1}}\sqrt{\frac{PQ}{n}} \tag{5.11}$$

where N = population size
n = number of respondents in the sample
P = proportion of groundwater users in the population
Q = proportion of users of other sources in the population (i.e., $Q = 1 - P$)

The formula for the desired *degree of precision* is

$$d = t\sqrt{\frac{N-n}{N-1}}\sqrt{\frac{PQ}{n}} \tag{5.12}$$

where t is the critical value of the t distribution corresponding to the desired confidence probability (i.e., the abscissa of the normal curve that cuts off an area of α at the tails). Solving Eq. (5.12) for n gives:

$$n = \frac{t^2\, PQ/d^2}{1+\,(1/N)\,(t^2PQ/d^2 - 1)} \tag{5.13}$$

If n/N is negligible because N is large, we can take the first approximation of n_o by using an advanced estimate p for P (and q for Q) from the formula

$$n_o = \frac{t^2\, pq}{d^2} \tag{5.14}$$

After obtaining n_o, we can introduce the finite population correction to the sample size from the following formula:

$$n = \frac{n_o}{1+n_o/N} \tag{5.15}$$

The sampling plans and sample size determinations are important elements of the process of collecting water-use data. Because water users are unlikely to form a homogeneous group, stratified random sampling is the most useful procedure for obtaining representative samples of water users.

5.4 DEVELOPMENT OF DATA SETS

Water-use data are usually collected for the purpose of monitoring water use. Many states require that water users submit annual (and/or monthly) records of their water withdrawals or discharges as a part of their permitting process. These data can be used for statistical analysis of water-use trends as well as for the development of water-use models. The latter purpose would also require data on variables which influence water use such as weather, price, employment, and land use.

In economics, data on economic activities (or variables) are collected at micro- or macrolevels. Observations on individual households, families, or firms are referred to as *microdata.* National-level accounts and observations on entire industries are called *macrodata.* In water-use modeling and analysis, the corresponding types of data are sometimes referred to as *disaggregate* and *aggregate* data. Levels of aggregation may vary from the end-use level (e.g., toilet flushing, lawn watering) or the municipal level (e.g., total production or total metered use) to total withdrawals in a region or state.

In developing water-use models, it is necessary to distinguish among different types of variables and their levels of measurement. The latter are also referred to as scales. In a mathematical sense, a variable is a quantity or function that may assume any given value or set of values, as opposed to a constant that does not or cannot change or vary. Depending on the character (or type) of values that can be assumed by a variable, the following can be distinguished: (1) continuous variables, (2) discrete variables, and (3) random variables. A *continuous variable* can assume any value within the interval where it exists (e.g., monthly water use in a demand area). A *discrete variable* can assume only discrete values such as the number of customers in a demand area. Finally, a *random* (or stochastic) *variable* can take any set of values (positive or negative) with a given probability. Random variables can be discrete or continuous.

5.4.1 Data Scales

The data on a variable can be measured on different kinds of scales. Such scales are used to describe the data and variables used in the analysis. The data can be nominal (either ordinal or nonordinal) or interral (including ratio data).

Nominal, or *categorical,* data are measurements that contain sufficient information to classify and count objects. Nominal data can be classified according to ordinal or nonordinal scales. *Ordinal* scales rank data according to the *value* of the variable that is being analyzed. Objects in an ordinal scale are characterized by relative rank, so a typical relationship is expressed in terms such as *higher, greater,* or *preferred to.* Ordinal ranking of data commonly occurs in the use of surveys. For example, survey data of household income are usually ranked (or classified ordinally) by income range, such as 1 ≤ \$25,000; 2 = \$25,000 to \$49,999; 3 = \$50,000 to \$74,999; etc. Notice then that ordinal rankings are hierarchical. *Nonordinal* data are ranked by variable type and therefore cannot be ranked hierarchically on a numerical scale. An example of nonordinal data comes from a survey of residential landscapes in southern California. Survey teams were asked to classify turf landscapes as 1 = Bermuda grass, 2 = other warm-season grasses, 3 = tall fescue, and 4 = cool-season grasses. Unlike the example of ordinal ranking given above, category 2 is not in any sense greater than category 1, because the categorization is based on landscape type without reference to numerical measurement of the distance between ranked objects. Finally, there are *interval* data that contain numerical values from a continuous scale. Variables with interval data are therefore most frequently called continuous variables. Because a continuous variable can assume an infinite number of values, continuous variables can theoretically be measured only over an interval, hence the name interval data. If the continuous measurement scale contains a true theoretical zero (e.g., water use), then it is called a ratio scale.

Another classification of variables is used in statistical or econometric modeling. In constructing statistical relationships, an attempt is usually made to predict or explain the effects of one variable by examining changes in one or more other variables that are known or expected to influence the former variable. In mathematics, the variable to be predicted is called the *dependent* variable, while the variables that influence it are called *independent.* However, in the analytical literature from various disciplines, a number of alternative terms are often used to describe and classify variables. Table 5.3 contains a list of such terms.

5.4.2 Data Arrangements

For modeling purposes, data (or observations) on water-use and related variables can be obtained and arranged in several ways. Depending on which type of arrangement is used, four types of data configurations can be distinguished: (1) time series data, (2) cross-sectional data, (3) pooled time series and cross-sectional data, and (4) panel (or longitudinal) data. In *time series* data, observations on all variables in the data set are taken at regular time intervals (e.g., daily, weekly, monthly, annually). In *cross-sectional* data, observations are taken at one time (either point in time or time interval) but for different units, such as individuals,

TABLE 5.3 Alternative Terms for Dependent and Independent Variables

Dependent	Independent
Explained	Explanatory
Predicted	Predictor
Regressed	Regressor
Response	Causal
Endogenous	Exogenous
Target	Control

households, sectors of water users, cities, or counties. *Pooled* data sets combine both time series and cross-sectional observations to form a single data matrix. Finally, *panel* data represent repeated surveys of the same cross-sectional sample at different periods of time. Databases can be built from records of water use by supplementing them with data from other sources such as random samples of water users. The possible data configurations that can be constructed from water-use histories of individual water users include

1. Time series of measurement period water-use data for individual water users
2. Time series of measurement period water-use data for all users in the sample
3. Cross-sectional water-use data for the same measurement period extending across all customers in the sample
4. Cross-sectional water-use data aggregated for two or more measurement periods representing seasonal or annual use extending across all users in the sample
5. Pooled time series cross-sectional data for all measurement periods and all water users
6. Pooled time series cross-sectional data aggregated over two or more measurement periods for all water users

In mathematical terms, we can describe each data configuration by designating the water use of user i during measurement period t as q_{it}. If n and m represent, respectively, the number of users in the sample and the number of measurement periods, we can describe the six data configurations as

1. *Customer time series data.* Represents total water use of user i in each time period t:

$$q_{it}$$

where i is a constant and $t = 1, \ldots, m$.

2. *Aggregate time series data.* Represents total water use of all users in the sample in each time period t:

$$\sum_{i=1}^{n} q_{it}$$

where $t = 1, \ldots, m$.

3. Cross-sectional billing period data. Represents total water use of *each* user i during billing period t:

$$q_{it}$$

where $i = 1, \ldots, n$ and t is a constant.

4. Seasonal (or annual) cross-sectional data. Represents total water use of each user i during seasonal or annual period k:

$$\sum_{t=1}^{k} q_{it}$$

where $i = 1, \ldots, n$ and k is the number of billing periods in each season.

5. *Pooled time series cross-sectional data.* Represents water use of each user i in each time period t:

$$q_{it}$$

where $i = 1, \ldots, n$ and $t = 1, \ldots, m$.

6. *Pooled time series cross-sectional data with seasonal (or annual) aggregations.* Represents water use of each user i in each season comprising k billing periods:

$$\sum_{t=1}^{k} q_{it}$$

where $i = 1, \ldots, n$ and k is the number of billing periods in each season.

If observations on water use, q_{it}, are supplemented with data on variables that are believed to be predictors of water demand, then regression analysis can be applied to any of the above data configurations. Section 5.7.2 describes various water-use modeling techniques.

5.5 WATER-USE AND SERVICE AREA DATA

Water-use data from water supply agency (i.e., water utility) records can be used for examining historical trends in water use and disaggregating total use into seasons, sectors, and specific end uses within each sector.

5.5.1 Water-Use Data

Water production records are a good source of data on total water demands in the area served by a public system. Water utilities usually have one or more production meters that are generally read at least daily. These production meters are typically maintained for accuracy and therefore usually produce highly reliable measurements of water flows into the distribution system. The water treatment plants or pumping stations usually employ continuous metering of the flow of finished water to the distribution system. The data may be recorded on paper recording charts which can be used to generate a time series of total production (or production from various supply sources) at daily or hourly time intervals. These data can be used for deriving temporal characteristics of aggregate water use (e.g., peak-day, peak-hour, day-of-week). The usefulness of the production data for water-demand analysis may include, but is not limited to, (1) the analysis of unaccounted water use (comparing production with water sales data), (2) the measurement of the aggregate effect of emergency conservation campaigns on water use, or (3) the analysis of the relationship between water production and weather variability.

Customer billing data can be used for disaggregating water use into customer sectors. Typically, retail water agencies maintain individual computer records of monthly, bimonthly, quarterly, or, less often, semiannual or annual water-consumption records for all metered customers. Active computer files usually retain up to 12 or 15 past meter readings for each customer. Depending on the length of the billing cycle, the active file records can contain a 12- to 36-month history of water use. This is a valuable source of water-use data that is often not exploited to its full potential. However, billing data can suffer from the following problems: (1) unequal billing periods, (2) lack of correspondence between billing periods and calendar months, (3) estimated meter readings or incorrect meter readings due to meter misregistration, (4) unusual usage levels, (5) meter replacements and manual adjustments to meters, and (6) changes in customer occupancy. Because of these problems, sole reliance on customer billing records necessarily limits the usefulness of these data for various measurements. Although billing data are becoming more readily available on electronic media, the use of electronic media is still not routine with many utilities.

Special metering is sometimes undertaken in order to obtain water-use data for research purposes. In recent years, water utilities have begun to utilize new meter-reading technologies that greatly reduce the cost of monitoring the water use of individual customers. However, the initial investment costs of adopting these new technologies can be prohibitively high. The new technologies include automatic meter-reading (AMR) devices and electronic remote meter-reading (ERMR) devices (see Schlenger, 1991). Automatic meter-reading devices are carried by meter readers on their routes and plugged into the site meters for the automatic meter reading. The data from the automatic meter readings are stored in the AMR device and then can be downloaded to central computer systems. The ERMR devices can be used to read site meters without actually visiting the meter location. Various methods of remote meter reading include (1) telephone dial-outbound, (2) telephone dial-inbound, (3) telephone scanning, (4) cable television, (5) radio frequency, and (6) power-line carrier.

These new technologies may permit water agencies to obtain daily or weekly meter readings from individual accounts. More frequent readings can improve the precision of water-use measurements. However, the most useful measurements can be obtained by devices that monitor individual pulses of water use on the service line. These pulses can be correlated with water flows through individual fixtures and appliances on customer premises (toilet flushes, shower flows, etc.), thus permitting accurate measurements of all end uses of water (DeOreo et al., 1996). For example, the Research Foundation of the American Water Works Association and 12 cities in the United States and Canada sponsored a unique study of water use in 1200 households. The study utilized the latest technology in micrometering of residential flows to measure precisely the amounts of water used for individual domestic purposes such as showering, bathing, toilet flushing, clothes washing, dish washing, yard watering, and other end uses of urban residents. Water meter readings were recorded in 10-second intervals using electronic data loggers. More detailed technical information about this study can be found in the final technical report authored by DeOreo et al. (1998).

Currently, these technologies are used only by a few water agencies. However, as the technology becomes more frequently adopted, there is great potential for new information sources for water-demand analysis. Unobtrusive metering technologies that permit accurate measurements of end uses of water are particularly useful for measurement of efficiency-in-use and water conservation planning.

5.5.2 Service Area Data

Because water use is a function of demographic, economic, and climatic factors, an accurate description of the characteristics of the resident population, housing stock, economic activities, and weather patterns in the service area can serve as a basis for the analysis of water demands. This section discusses the types of information that can be used to characterize the service area and identify potential data sources.

Accurate *service area maps* are indispensable. Often water service areas do not follow political (i.e., city or county) boundaries. However, demographic and socioeconomic data are most readily available by political boundaries or by census-designated boundaries (i.e., census tracts). Therefore, in order to relate water service area data to demographic and socioeconomic characteristics for planning purposes, it is often necessary to determine the relationship between service area boundaries and political or census-designated boundaries. A service area map with overlays for census tract, zip code, or other political boundaries is most useful for this purpose. Service area maps with geographic or land-use overlays have usefulness in many other planning activities. For example, only parts of the service area may be targeted for specific activities (e.g., the service area might be disaggregated spatially into pressure zones or rate zones for the purpose of water-use forecasting and/or facility planning).

Usually, a set of maps can be obtained from the facility planning or engineering department of the water agency. These maps should indicate the historical growth of the service territory and potential future additions. Also, the maps should note areas within a given community that are partially served or are served by other water supply agencies. Maps denoting political boundaries and demographic characteristics can often be obtained from local and regional planning agencies. Mapping work is greatly simplified by geographic information system (GIS) technology, which is a computerized mapping technique. The GIS may have the service area divided into separate units (e.g., census tracts, pressure zones) and have several information bases about each separate unit (e.g., water-use characteristics, land use, socioeconomic characteristics).

Population and housing characteristics (i.e., household income, lot size, persons per household, home value) are determinants of residential water use. Therefore, it is important to obtain information on these characteristics as well as to understand their impact on water use. The conventional method of estimating *population served* by multiplying *total service connections* by *persons per connection* is not very accurate. Although this method may be acceptable to estimate the population served in the residential sector (assuming an accurate measurement of persons per connection), this is not an accurate method for total service area population served because of the confounding effect of commercial, institutional, governmental, and industrial accounts.

Knowledge of the number and type of housing units in the service area is very useful in water-demand analysis because water-use patterns differ among housing types. On both a per housing unit and a per capita basis, water use in the multifamily sector tends to be lower than in the single-family sector. This is the result of different household compositions and the fact that residents in multifamily housing, on a per unit basis, have less opportunity for outdoor water-use practices. The total number of residential *accounts* served by a water supply system is not a good indication of the number of housing *units* in the service area because of the varying number of units served by multifamily accounts. However, the number of single-family customer accounts is typically a good indication of the number of single-family housing units served by the water system. Unless the water agency maintains records on the number of multifamily living units per multifamily account, housing count data must be obtained from the demographic data developed by (1) the U.S. Bureau of the Census, (2) state departments of finance or economic development, (3) regional associations of governments, and (4) local county and city planning agencies.

Again, deviations of the water service boundaries from political boundaries must be considered when using these data. In addition to population and housing counts, some of the agencies listed above can also provide data on population characteristics (e.g., family size, age, income) and data on local housing (e.g., number of homes by type, new construction permits, vacancy rates). Table 5.4 gives examples of demographic and housing data that are included in the 2000 census files. Two primary questionnaires were used in the collection of census data—the short form which included questions that were asked of all persons and

TABLE 5.4 Selected Examples of U.S. 2000 Census Data

Population	Housing
100% component	
Household relationship	Number of units in structure
Sex	Number of rooms in unit
Race	Tenure (owned or rented)
Age	Value of home or monthly rent paid
Ethnic/racial origin	Congregate housing (meals included in rent)
	Vacancy characteristics
20% sample components	
Social characteristics:	
Education	Year moved into residence
Ancestry	Number of bedrooms
Migration	Plumbing and kitchen facilities
Language spoken at home	Telephone in unit
	Heating fuel
Economic characteristics:	Source of water and method of sewage disposal
Labor force	Year structure built
Place of work and journey to work	Condominium status
Year last worked	Farm residence
Occupation, industry, and class of worker	Shelter costs, including utilities
Work experience in 1999	
Income in 1999	

housing units (i.e., the 100 percent component) and the long form which targeted only about 20 percent of the population on additional subject items (i.e., the sample component). These data are presented by geographic and political subdivisions (i.e., states, counties, cities). For major urban areas, census data are further disaggregated into census tracts, city blocks, and block groups (but not for individual dwellings).

Information on *commercial, institutional, and industrial activities* in the service area is helpful in analyzing nonresidential water demands. There is a great diversity of purposes for which water is used in the commercial, institutional, and industrial (manufacturing) sectors. The uses of water may include sanitary, cooling and condensing, boiler feed, and landscape irrigation. The type of business activity conducted in a commercial and industrial establishment can provide useful information regarding the purposes for which water is used and therefore the types of conservation measures that might be applicable. Furthermore, data on square feet of floor space, land acreage, number of employees, number of rooms (for hotels and schools), and financial performance can also be useful information in predicting commercial and industrial water use. However, some of this information is not readily available for individual establishments or aggregated into political or census-designated boundaries. Some of the previously listed agencies maintain data on local economic activities including the number of establishments, employment, and financial performance of businesses (e.g., sales) disaggregated by industry type as denoted by the Department of Commerce Standard Industrial Classification (SIC) and the newly devised North American Industry Classification System (NAICS) codes.

Additional establishment or employment data can be obtained from local and regional planning commissions and local chambers of commerce or purchased from private firms. Some private vendors can provide customized computer databases containing information on large samples of businesses in designated geographic areas. Establishment and employment data can be analyzed to determine types of business establishments that represent a major portion of nonresidential water use either because of large employment or because of large water requirements for processing or other needs. Business types can be cross-checked with agency billing records and used for disaggregating water use into specific groups of nonresidential users.

Water used in parks, cemeteries, school playgrounds, and highway medians can account for a significant portion of total use and offer a potential for conservation. Data on *public and government facilities* can be obtained from city and regional planning departments, city park districts, street departments, and the department of transportation (for highway medians). Information that might be obtained includes land-use data for various purposes (in square feet or acres) as well as the number of employees in various facilities.

5.6 COMPONENTS OF WATER DEMAND

The quantities of water delivered by a public water supply system can be disaggregated by user sector and season. Table 5.5 gives an example of such decomposition of water demands in the urban area of southern California. This section describes analytical methods for disaggregating total urban water use.

5.6.1 Sectors of Water Users

Customer billing records can be used to obtain estimates of total metered use of water and to determine the distribution of total water use among several homoge-

TABLE 5.5 Sectoral and Seasonal Disaggregation of Urban Water Use in Southern California (Most Likely Ranges Given in Parentheses)

User sector/ subsector	Disaggregation of urban sectors	Percent of total annual use						
		Seasonal disaggregation				Components of outdoor use		
		Nonseasonal (base use)	Seasonal (peak use)	Indoor use	Outdoor use	Irrigation	Cooling (AC)	Other
Residential sector								
Single-family	34.4	68.7	31.3	65.4	34.6	30.8	0.0	3.8
(65–75)	(25–35)	(60–70)	(30–40)					
Multifamily	25.0	83.9	16.1	82.2	17.8	16.1	0.4	1.3
(80–90)	(10–20)	(75–85)	(15–25)					
Total residential	59.4	72.1	27.9	69.6	30.4	27.2	0.2	3.0
(70–80)	(20–30)	(65–75)	(25–35)					
Nonresidential sector								
Commercial	18.8	74.9	25.1	71.3	28.7	21.8	6.9	0.0*
(70–80)	(20–30)	(70–75)	(25–30)					
Industrial	6.0	79.5	20.5	79.5	20.5	12.3	8.2	0.0*
(75–90)	(10–25)	(75–90)	(10–25)					
Public5.1	46.2	53.8	46.2	53.8	53.8	0.0†	0.0†	
(30–50)	(50–70)	(30–50)	(50–70)					
Other 1.1	58.0	42.0	58.0	42.0	42.0	0.0†	0.0†	
	(30–60)	(40–70)	(30–60)	(40–70)				

TABLE 5.5 *(Continued)*

User sector/ subsector	Disaggregation of urban sectors	Percent of total annual use						
		Seasonal disaggregation				Components of outdoor use		
		Nonseasonal (base use)	Seasonal (peak use)	Indoor use	Outdoor use	Irrigation	Cooling (AC)	Other
Irrigation (10–50)	0.4 (50–90)	34.0	66.0	0.0	100.0	100.0	0.0	0.0
Total nonresidential	31.4	70.0	30.0	67.4	32.6	26.9	5.7	0.0
Unaccounted use	9.2	100.0	0.0	100.0	0.0			
Total urban use	100.0	74.0	26.0	71.7	28.3	24.6	1.9	1.8

*Other uses in these sectors are included under landscape irrigation.
†Cooling and other uses are included under landscape irrigation.
Source: Dziegielewski et al. (1990).

neous classes of water users. A disaggregation of total metered use into major user sectors (such as residential, commercial, and industrial) can be developed from customer billing records by using one of the following four methods: (1) analysis of available premise (user type) categories, (2) distribution of meter sizes, (3) sampling of billing files, and (4) development of premise code data.

If individual customer files contain *customer premise categories* identifying the type of customers, then a simple computer program may be used to produce annual billing summaries by customer type. Each premise code can be assigned to one of the following homogeneous sectors of water users: (1) single-family residential, (2) multifamily complexes and apartment buildings, (3) commercial sector, (4) government and public sector, (5) manufacturing (industrial) sector, and (6) unaccounted uses. Water use in these sectors can be disaggregated into seasons and end uses if necessary.

In cases where the customer billing file does not contain premise code (or customer-type) information, an approximate separation of users by residential, commercial, and industrial categories can be performed based on the distribution of *meter sizes* with a manual classification of the largest users. For example, single-family homes and small businesses are usually serviced by $^5/_8$-inch (in) or $^3/_4$-in meters. The problem with this method is that some meter sizes, particularly the larger meters (e.g., 1-in) may overlap several customer types thus decreasing the precision of the disaggregation of water use into customer classes.

In cases where water-use data by customer class are not available and the distribution of water use by meter sizes produces unreliable estimates of water use by customer class, disaggregation can be accomplished by taking a *random sample of customer accounts.* The number of accounts in the sample (i.e., sample size) will depend on the desired accuracy of water use in different customer classes. Sampling efficiency (or precision) can be improved by taking a stratified random sample of users with complete enumeration of the large water-using customers. Possible stratified sampling procedures include (1) taking a random sample of customer accounts from each meter-size category (sample size within each stratum can be proportional to the total water use within each stratum); (2) using the same approach as in procedure 1 except excluding the upper strata (e.g., meter sizes greater than 2 in) from the sampled population and performing complete enumeration of the largest users; or (3) separating all accounts into two categories based on meter size, with the smallest meter sizes representing the residential sector and the remaining meters representing the nonresidential sector, and then taking a random sample from each category.

The samples of the customer accounts can then be assigned manually into customer classes by visual inspection of customer record (account name) and/or by telephone verification of customer accounts. Depending upon the sample sizes, this can be a time- and resource-intensive exercise.

Regardless of which sampling approach is selected, analysis of the sample should produce the following two estimates: (1) proportion (or percent) of total customers by customer class, and (2) proportion (or percent) of total metered water use by customer class. It is also desirable to calculate the precision of the estimates based on the sample variance.

If sufficient time and resources are available, the classification of all customer accounts into appropriate user types is a worthwhile undertaking. The classification would require: (1) adding a data field (or using existing unassigned fields) to the customer computer file; (2) developing a set of nonoverlapping customer classes and precisely defining each class; (3) determining the customer class for each existing customer by classifying all customers during meter reading or surveying (independently) all customers and requesting them to classify their premises on water bills; and (4) adding customer classification categories to the application forms for new connections.

5.6.2 Seasonal and Nonseasonal Components

Within each user sector, water use can be separated into its seasonal and nonseasonal components. *Seasonal use* can be defined as an aggregate of end uses of water, such as lawn watering or cooling, that varies from month to month in response to changing weather conditions (or due to other influences that are seasonal in nature). *Nonseasonal use,* on the other hand, can be defined as an aggregate of end uses of water, such as toilet flushing or dishwasher use, that remain relatively constant from month to month because these uses are not sensitive to weather conditions or other seasonal influences. Often, seasonal and nonseasonal components of water use are taken to represent the outdoor (or exterior) and indoor (or interior) water uses, respectively. Such an assumption is imprecise because some uses that occur inside the buildings can be seasonal (e.g., humidifier use or evaporative cooler), and some outdoor uses can be nonseasonal (e.g., car washing in warmer climates). The difficulties in classifying various end uses into outdoor and indoor categories must be kept in mind when water use is divided into seasonal and nonseasonal components.

Monthly water use data can be used to derive estimates of seasonal and nonseasonal water use. The terms *seasonal* and *nonseasonal* relate to the method of characterizing a monthly time series of water-use records. This method is sometimes referred to as the *minimum-month* method because it uses the month of lowest use to represent the nonseasonal component of water use. With the minimum-month method, the percent of annual use in a given year that is considered seasonal is calculated from the formula

$$S_P = 100 - (M_P \cdot 12) \tag{5.16}$$

where S_P is the percent of annual use that is seasonal and M_P is the percent of annual use during the minimum month.

The best representation of how much water is used during a given calendar month is the aggregation of daily pumpage information which records how much water enters the distribution system every day. Monthly water use information for customer groups is more difficult to obtain because of the effects of monthly and bimonthly billing cycles. When water utilities summarize the amount of water sold in a given month (both in aggregate or by customer group), this information typically represents the amount of water billed in a given month rather than the amount of water used. Bimonthly billing cycles (which indicate the amount of water used over a 2-month period) further confounds calendar month water use. Differences between the amount of water billed in a given month and the actual water consumed occurs because (1) accounts are read in different months (e.g., some accounts are read in January, March, May, etc., and others are read in February, April, June, etc.), and (2) meter readings are typically recorded on any given day within a month.

In order to get a better representation of monthly water-use patterns, it is necessary to allocate water use from monthly and bimonthly billing cycle records into water use during specific calendar months. Thus, the primary purpose of the allocation (or smoothing) techniques is to adjust water consumption billing records which are read on either a monthly or bimonthly basis (e.g., even accounts read on even months and odd accounts read on odd months in the bimonthly case) into calendar month consumption for purposes of further analysis. Data smoothing procedures are performed on two levels. First, the information provided by water utilities on the amount of water billed to a customer group in a given month is smoothed to represent the amount of water actually consumed by a customer group in a given month. The smoothing procedure varies depending on whether a bimonthly or monthly billing cycle is in effect. Second, account-level water-use records are smoothed so that estimates of calendar month water use can be determined. Both types of procedures are described below.

Figure 5.1 presents a graphical representation of the procedure used to smooth aggregate sales data which utilize a *monthly billing cycle.* A monthly billing practice involves reading water meters of individual customers in approximately 1-month-long time intervals. Meters are read every working day, and all meters read during, for example, the month of February, are billed and recorded as the February water use ($Q_{\text{Feb4}}{}^{b}$). In reality, only a portion of the billed water use, $Q_{\text{Feb}}{}^{b}$, actually occurred during the calendar month of February. Theoretically, $Q_{\text{Feb}}{}^{b}$ represents water use of individual customers during n monthly periods (where n is the number of billed customers) ending between the first and last meter-reading date in February. Therefore, for individual customers, 1-month-long periods of water use would fall between January 1 and February 28.

Assuming that all users in a given customer group (e.g., single-family, commercial) are relatively homogeneous with respect to water use and that the effects of weather on water use during the two consecutive calendar months are not

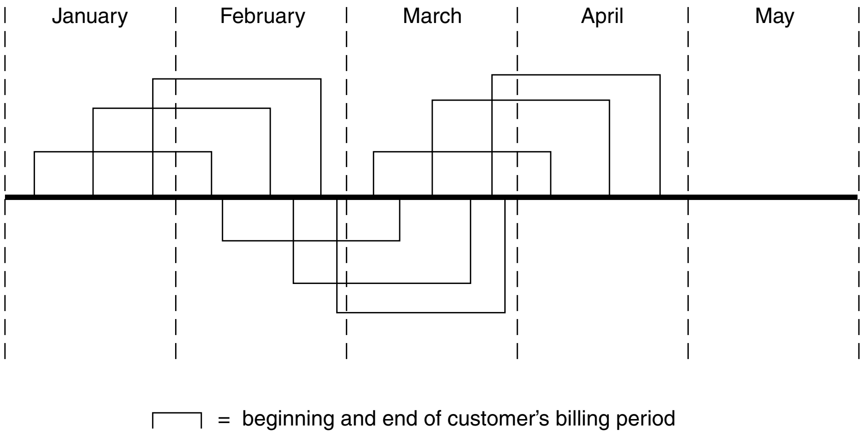


FIGURE 5.1 Allocation of monthly water-use billing records into calendar month consumption. Estimate of calendar month water-use ($Q_N{}^c$): February, $Q_{\text{Feb}}{}^c = x\,Q_{\text{Feb}}{}^b + y\,Q_{\text{Mar}}{}^b$; March, $Q_{\text{Mar}}{}^c = x\,Q_{\text{Mar}}{}^b + y\,Q_{\text{Apr}}{}^b$; Nth month, $Q_N{}^c = x\,Q_N{}^b + y\,Q_{N+1}{}^b$. $Q_{\text{Mar}}{}^b$ = water use billed to customer during the month of March; x and y = proportions of $Q_N{}^b$ and $Q_{N=1}{}^b$ billed water use allocated to $Q_N{}^c$ and $Q_{N=1}{}^c$.

substantially different (i.e., that water use of the individual customers is evenly distributed throughout the period between meter-reading dates), the calendar month water use during month N can be estimated as

$$Q_N{}^c = 0.5Q_N{}^b + 0.5Q_{N+1}{}^b \tag{5.17}$$

where Q^c is the amount of water used during the calendar month and Q^b is the amount of water billed during the calendar month. This equation indicates that water actually used during the calendar month of, for example, March ($Q_{\text{Mar}}{}^c$) would comprise one-half of the consumption billed in March ($Q_{\text{Mar}}{}^b$) and one-half of the consumption billed in April ($Q_{\text{Apr}}{}^b$).

Whereas Eq. (5.17) allows the calculation of calendar month water use, a variant of this equation can be used to calculate the average water use per account in a given month when the number of billed accounts varies between the two consecutive months:

$$q_N = \frac{Q_N}{A_N}\,\frac{A_N}{A_N + A_{N+1}} + \frac{Q_{N+1}}{A_{N+1}}\,\frac{A_{N+1}}{A_N + A_{N+1}} \tag{5.18}$$

where q_N = average water use per account in any given month
Q_N = amount of water billed in any given month
A_N = number of accounts billed in any given month

Figure 5.2 presents the procedure that was used to allocate aggregate monthly sales data produced as a result of a bimonthly meter reading cycle. In this procedure, although individual meters are read every 2 months, customer billing is performed during each calendar month. As a result, all customers within a given sector are divided into two groups, referred to as group A and group B. Meters of customers in group A are read and billed during odd months (January, March, May, etc.), while in group B, meters are read and billed during even months (February, April, June, etc.).

Because some accounts are billed in any given month, aggregate monthly sales data will reflect, in any given month, only approximately one-half of the true num-

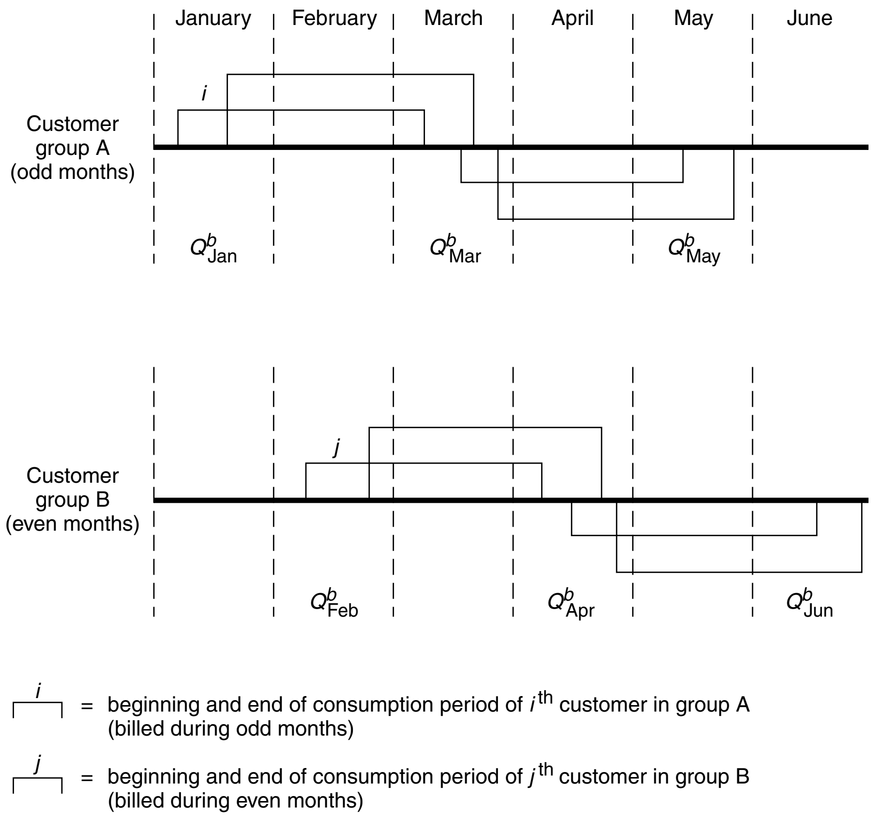


FIGURE 5.2 Allocation of bimonthly water-use billing records into calendar month consumption. Estimate of calendar month water use (Q_N^c): March, $Q_{\text{Mar}}^c = 0.25Q_{\text{Mar}(A)}^b = 0.50Q_{\text{Apr}(B)}^b + 0.25Q_{\text{May}(A)}^b$; April, $Q_{\text{Apr}}^c = 0.25Q_{\text{Apr}(A)}^b + 0.50Q_{\text{May}(B)}^b + 0.25Q_{\text{Jun}(a)}^b$; Nth month, $Q_N^c = 0.25Q_N^b + 0.50Q_{N+1}^b + 0.25Q_{N+2}^b = [(Q_N + Q_{N+2})/2 + Q_{N+1}]/2$. Q_{Mar}^b = water use billed during the month of March (includes only customers in group A); Q_{Apr}^b = water use billed during the month of April (includes only customers in group B).

ber of accounts. For example, water use billed and recorded as the March consumption ($Q_{\mathrm{Mar}}{}^{b}$) includes only customers of group A whose individual 2-month consumption period falls between January 1 and March 31. By analogy, water use billed and recorded as the April consumption ($Q_{\mathrm{Apr}}{}^{b}$) includes only customers of group B whose individual 2-month consumption period falls between February 1 and April 30.

Again, assuming that water users in both groups are homogeneous with respect to water use (because they represent the same user sector) and that weather effects during the consecutive calendar months are not drastically different (i.e., that water use of the individual customer is evenly distributed throughout his or her 2-month consumption period), we can estimate total water use during any calendar month using the following formula:

$$Q_N{}^c = 0.25Q_N{}^b + 0.5Q_{N+1}{}^b + 0.25Q_{N+2}{}^b \tag{5.19}$$

This relationship is derived from Fig. 5.2. It indicates that water actually used during the calendar month of March would comprise one-fourth of consumption billed in March, one-half of that billed in April, and one-fourth of that billed in May.

Whereas Eq. (5.19) allows the calculation of calendar month water use given bimonthly billing records, a variant of this equation can be used to calculate *average water use per account* in a given month:

$$q_N = \left(0.25\,\frac{Q_N}{A_N} + 0.25\,\frac{Q_{N+2}}{A_{N+2}}\right)\frac{0.5A_N + 0.5A_{N+2}}{0.5A_N + A_{N+1} + 0.5A_{N+2}} + 0.5\,\frac{Q_{N+1}}{A_{N+1}}\,\frac{A_{N+1}}{0.5A_N + A_{N+1} + 0.5A_{N+2}} \tag{5.20}$$

$$= 0.25\left(\frac{Q_N}{A_N} + \frac{Q_{N+2}}{A_{N+2}}\right)\frac{A_N + A_{N+2}}{A_N + 2A_{N+1} + A_{N+2}} + 0.5\,\frac{Q_{N+1}}{A_{N+1}}\,\frac{2A_{N+1}}{A_N + 2A_{N+1} + A_{N+2}} \tag{5.21}$$

where q_N is the average water use per account in any given month.

In the case where both *monthly* and *bimonthly* billing cycles exist within a water utility, total monthly water use in a given month can be obtained by adding smoothed calendar water use from the monthly smoothing procedure to smoothed calendar water use from the bimonthly smoothing procedure. In the case of determining the average water use per account in any given month given the existence of both monthly and bimonthly billing cycles, the following weighting procedure can be used:

$$q_{N,avg} = q_{N,b}\,\mathrm{WF}_b + q_{N,m}\,\mathrm{WF}_m \tag{5.22}$$

where $q_{N,\mathrm{avg}}$ = average water use per account in any given month

$q_{N,b}$ = average water use per account in any given month from bimonthly billing cycle

$q_{N,m}$ = average water use per account in any given month from monthly billing cycle

WF_b = weight factor for water use per account from bimonthly billing cycle [see Eq. (5.23)]

WF_m = weight factor for water use per account from monthly billing cycle [see Eq. (5.24)]

$$\mathrm{WF}_b = \frac{0.5A_{N,b} + A_{N+1,b} + 0.5A_{N+2,b}}{(0.5A_{N,b} + A_{N+1,b} + 0.5A_{N+2,b}) + (0.5A_{N,m} + 0.5_{N+1,m})}$$

$$= \frac{A_{N,b} + 2A_{N+1,b} + A_{N+2,b}}{(A_{N,b} + 2A_{N+1,b} + A_{N+2,b}) + (A_{N,m} + A_{N+1,m})} \tag{5.23}$$

$$\mathrm{WF}_m = \frac{0.5A_{N,m} + 0.5A_{N+1,m}}{(0.5A_{N,b} + A_{N+1,b} + 0.5A_{N+2,b}) + (0.5A_{N,m} + 0.5A_{N+1,m})}$$

$$= \frac{A_{N,m} + A_{N+1,m}}{(A_{N,b} + 2A_{N+1,b} + A_{N+2,b}) + (A_{N,m} + A_{N+1,m})} \tag{5.24}$$

In addition to the allocation of aggregate billing data into calendar month, water consumption records for individual customers can also be allocated to calendar months. For example, if a single account's meter is read on March 15 and then again on May 15, a smoothing procedure can be used to standardize individual account billing cycle data into calendar month use. The meter-reading cycle can be monthly, bimonthly, or even trimonthly. The composition of the equation applied to this type of smoothing is as follows:

$$Q_n^{\,c} = \sum_{i=(n-x)}^{i=(n+y)} N_i\,\overline{Q_i^{\,b}} \tag{5.25}$$

where n = nth calendar month

x = number of months prior to month n that fall within the billing period

y = number of months beyond month n that fall within the billing period

i = summation index (ith month)
Q_n^c = quantity of water allocated to calendar month n
N_i = number of days in ith month that are proportioned to consumption in month n
Q_i^b = average daily water consumption for billing period which represents ith month

This smoothing procedure is applied to the water billing histories for each account. In this procedure, each account will be smoothed in accordance with the current read date and the prior read date. That is, the smoothing procedure must first look at the current and prior read dates. Next, consumption for each record is allocated into calendar months by the fraction of water use that belongs to each month encountered in the billing period [see Eq. (5.25)]. Finally, consumption is summed for each account by month.

5.6.3 End Uses of Water

A meaningful assessment of the efficiency of water use cannot be made without breaking down the seasonal and nonseasonal uses of water into specific end uses. Precise measurements of the quantities of water used for showering, toilet flushing, and other purposes require installation of flow-recording devices on each water outlet found on customer premises. Because such measurements are very costly, engineering estimates are often used. However, such estimates are of limited validity because they tend to rely on many assumptions and often ignore the physical and behavioral settings in which water use takes place. Analytical methods for quantifying the significant end uses of water are described in Sec. 5.7.

5.7 WATER-USE RELATIONSHIPS

Reasonably precise estimates of water use can be obtained by disaggregating the total delivery of water to urban areas into two or more classes of water use and determining separate average rates of water use for each class. The *disaggregate estimation of water* use can be represented as the product of the number of users (or the demand driver count) and a constant average rate of usage.

$$Q_t = \sum_c N_{t,c} q_c \tag{5.26}$$

where $N_{t,c}$ represents the number of customers in a homogeneous user sector c at time t, and q_c is the unit-use coefficient (or average rate of water use per customer) in that sector.

Gains in accuracy of disaggregated estimates are possible because the historical records often show that the variance of average rate of water use within some homogeneous sectors of water users is smaller than the variance of the aggregate use. Although in some nonresidential sectors the variance of q_c is large, it is often more than offset by the smaller variance in the residential sectors with a disproportionately greater number of customers.

The *sectoral disaggregation* of total urban demands may also be extended spatially and temporally. With the added dimensions of disaggregation, Eq. (5.26) would be expanded to the form:

$$Q_t = \sum_c \sum_s \sum_g N_{t,c,s,g}\, q_{c,s,g} \tag{5.27}$$

where s denotes the disaggregation of water use according to its seasonal variations (e.g., annual, seasonal, monthly) and g represents the spatial disaggregation of water use into various geographic areas, such as pressure districts or land-use units, which are relevant for planning purposes. An example of a unit-use coefficient $q_{c,s,g}$ could be average use in single-family residential sector during the summer season in a pressure district.

The level of disaggregation is limited by the availability of data on average rates of water use in various sectors and the ability to obtain accurate estimates of driver counts (N) for each disaggregate sector. The latter are typically obtained from planning agencies who maintain data on population, housing, and employment for local areas.

The last three decades have produced numerous studies of the *determinants of urban water demand*. The advancements in theory were followed by the development of water use models which recognized that both the level of average daily water-use and its seasonal variation can be explained adequately by selected demographic, economic, and climatic characteristics of the study area. Such advanced models retain a high level of disaggregation; however, they allow the average rates of water use and the number of users (i.e., drivers) to change in response to changes in their determining factors:

$$Q_t = \sum_c \sum_s \sum_g N_{t,c,s,g}\, q_{t,c,s,g} \tag{5.28}$$

where

$$N_{t,c,s,g} = f(Z_i) \tag{5.29}$$

and

$$q_{t,c,s,g} = f(X_j) \tag{5.30}$$

where Z_i and X_j are, respectively, determinants of the number of water users (e.g., residents, housing units, or employees) and determinants of the average rates of water use (such as average water use per person, per household, or per employee). Table 5.6 gives examples of determinants of the number of users and their average rates of water use.

The remainder of this section describes the available data on the average rates of water use and reviews a number of empirical water-use models.

5.7.1 Average Rates of Water Use

The first step in the analysis of water demands in an area served by a public water supply system is to determine average annual rates of water use. The simplest rate is the gross per capita water use which is determined by dividing the total annual amount of water delivered to the distribution system by the estimated population served. Other rates are obtained by dividing the metered water use by various urban sectors by the number of users or customers in each sector. In addition to the average annual use, average rates during high-use and low-use seasons can be estimated.

The average rates of use are sometimes compared among service areas in order to assess the relative efficiency of water use. However, the aggregate nature of these rates precludes any meaningful comparisons among various service areas. These rates vary from city to city as a result of differences in local conditions that are unrelated to the efficiency in water use. For example, per capita use may range from 50 to 500 gal per person per day [189 to 1893 liters per day (L/day).] Table 5.7 shows the distribution of per capita rates among 392 water supply systems serving approximately 95 million people in the United States. The mean per capita use in

TABLE 5.6 Determinants of Urban Water Demand

Determinants of demand drivers, Z_i	Determinants of average rates of use, X_j
Natural birth rates	Air temperature and precipitation
Net migration	Type of urban landscapes
Family formation rates	Housing density (average parcel size)
Availability of affordable housing	Water-use efficiency
Economic growth and output	Household size and composition
Labor participation rates	Median household income
Urban growth policies	Price of water service
	Price of wastewater disposal
	Industrial productivity

TABLE 5.7 Average per Capita Rates of Water Use

Range of per capita use, gcd	Number of systems	Percent of systems
50–99	30	7.7
100–149	132	33.7
150–199	133	33.9
200–249	51	13.0
250–299	19	4.8
300+	27	6.9
Total	392	100.0

Source: Source: AWWA (1986).

this sample was 175 gcd (662 L/day) with a standard deviation of 72 gcd (273 L/day) (AWWA, 1986).

Generally, high per capita rates are found in water supply systems servicing large industrial or commercial sectors. Therefore, more meaningful comparisons would require the disaggregation of total use into homogeneous sectors of water users. Table 5.8 shows average use rates per housing unit in the residential sector for selected cities. Nonresidential water-use rates in gallons per employee per day (ged) are shown in Table 5.9.

5.7.2 Modeling of Water Use

Water-use models are usually obtained by fitting theoretical functions to the data sets described in Sec. 5.1. The selection of appropriate data and estimation techniques depends on the desired characteristics of the final model. For example, time series models of aggregate data can be used for developing models of aggregate use for near-term forecasting. Generally, time series aggregate models can be expected to provide predictions of water use that are less reliable than those obtained from pooled time series cross-sectional observations on water use of individual customers from homogeneous groups of users.

The most commonly used technique for developing water-use models is *regression analysis.* A simple regression model that captures the relationship between two variables can be written as

$$y_i = \alpha + \beta X_i + \varepsilon_i \tag{5.31}$$

TABLE 5.8 Average Rates of Water Use in Single-Family and Multiple-Unit Buildings

		Average use per dwelling unit, gpd	
Area/State	Year	Multiple-unit buildings	Single-family homes
A. Metered sales data			
National City, Calif.	1987	226	309
San Diego, Calif.	1990	261	
Santa Monica, Calif.	1988	231	376
Torrance, Calif.	1987	183	383
Anaheim, Calif.	1987	284	526
Beverly Hills, Calif.	1987	241	917
Camarillo, Calif.	1987	246	437
Fullerton, Calif.	1988	246	566
Los Angeles, Calif.	1987	264	408
Cape Coral, Fla.	1990	123	
Boston, Mass.	1990	167	
Framingham, Mass.	1992	152	
Newton, Mass.	1992	199	
Las Vegas, Nev.	1989	290	
B. Samples of buildings			
New York City, N.Y.	1990	248	
Seattle, Wash.	1988	107	
Baltimore, Md.	1965	218	
Washington, D.C.	1981	197	
Springfield, Ill.	1985	117	
Los Angeles, Calif.	1976	165	
San Francisco Bay, Calif.	1976	183	
Central Valley, Calif.	1975	144	
San Diego, Calif.	1991	137	
Pasadena, Calif.	1991	187	
Newton, Mass.	1992	155	
Bailment, Mass.	1990	121	
Framingham, Mass.	1992	164	

TABLE 5.9 Average Rates of Nonresidential Water Use from Establishment-Level Data

Category	SIC code	Use rate, ged	Sample size
Construction	—	31	246
General building contractors	15	118	66
Heavy construction	16	20	30
Special trade contractors	17	25	150
Manufacturing	—	164	2790
Food and kindred products	20	469	252
Textile mill products	22	784	20
Apparel and other textile products	23	26	91
Lumber and wood products	24	49	62
Furniture and fixtures	25	36	83
Paper and allied products	26	2614	93
Printing and publishing	27	37	174
Chemicals and allied products	28	267	211
Petroleum and coal products	29	1045	23
Rubber and misc. plastics products	30	119	116
Leather and leather products	31	148	10
Stone, clay, and glass products	32	202	83
Primary metal industries	33	178	80
Fabricated metal products	34	194	395
Industrial machinery and equipment	35	68	304
Electronic and other electrical equipment	36	95	409
Transportation equipment	37	84	182
Instruments and related products	38	66	147
Misc. manufacturing industries	39	36	55
Transportation and public utilities	—	50	226
Railroad transportation	40	68	3
Local and interurban passenger transit	41	26	32
Trucking and warehousing	42	85	100
U.S. Postal Service	43	5	1
Water transportation	44	353	10
Transportation by air	45	171	17
Transportation services	47	40	13
Communications	48	55	31
Electric, gas, and sanitary services	49	51	19
Wholesale trade	—	53	751
Wholesale trade—durable goods	50	46	518
Wholesale trade—nondurable goods	51	87	233
Retail trade	—	93	1044
Building materials and garden supplies	52	35	56
General merchandise stores	53	45	50
Food stores	54	100	90
Automotive dealers and service stations	55	49	198
Apparel and accessory stores	56	68	48

TABLE 5.9 (*Continued*)

Category	SIC code	Use rate, ged	Sample size
Retail trade			
Furniture and home furnishings stores	57	42	100
Eating and drinking places	58	156	341
Miscellaneous retail	59	132	161
Finance, insurance, and real estate	—	192	238
Depository institutions	60	62	77
Nondepository institutions	61	361	36
Security and commodity brokers	62	1240	2
Insurance carriers	63	136	9
Insurance agents, brokers, and service	64	89	24
Real estate	65	609	84
Holding and other investment offices	67	290	5
Services	—	137	1878
Hotels and other lodging places	70	230	197
Personal services	72	462	300
Business services	73	73	243
Auto repair, services, and parking	75	217	108
Miscellaneous repair services	76	69	42
Motion pictures	78	110	40
Amusement and recreation services	79	429	105
Health services	80	91	353
Legal services	81	821	15
Educational services	82	117	300
Social services	83	106	55
Museums, botanical, zoological gardens	84	208	9
Membership organizations	86	212	45
Engineering and management services	87	58	5
Services, NEC	89	73	60
Public administration	—	106	25
Executive, legislative, and general	91	155	2
Justice, public order, and safety	92	18	4
Administration of human resources	94	87	6
Environmental quality and housing	95	101	6
Administration of economic programs	96	274	5
National security and international affairs	97	112	2

Source: Planning and Management Consultants, Ltd., 1994, unpublished data.

where y_i = water use of customer i

α = intercept term of the equation and the component of the effect of X upon y that is constant regardless of the value of X

β = slope coefficient of the equation and the component of the effect of X upon y that changes depending upon the value of X

ε_i = error term for the ith customer, and $i = 1, 2, \ldots, n$, which measures the difference between the estimated value of y and the true observed value of y

This model also assumes that X, the independent variable, influences y, the dependent variable, while the dependent variable does not influence the independent variable in any way. Equation (5.31) decomposes water use y_i into *explained* and *unexplained* components, where the explained component is expressed as a function of a systematic force X. The unexplained component is expressed as random noise. In other words, in Eq. (5.31), $\alpha + \beta X$ is the deterministic component of y, and ε is the stochastic or random component.

In *ordinary least-squares* (OLS) regression analysis, the parameters α and β are estimated by fitting a regression line to water-use data so that the sum of squared residuals of y ($\Sigma\varepsilon_i^2$) away from the line is minimized. The method of least squares dictates that one choose the regression line where the sum of the squared deviations of the points from the line is a *minimum,* resulting in a line that "fits" the data as well as possible.

In order for OLS to yield valid results, the residuals (ε) must meet the five assumptions of simple linear regression:

1. *Zero mean.* $E(\varepsilon_i) = 0$ for all i, or the expected value of mean error is 0. In other words, the errors are expected to fluctuate randomly about 0 and, in a sense, cancel each other out.
2. *Common (or constant) variance.* Var $(\varepsilon_i) = \sigma^2$ for all i, which states that each error term has the same variance for each customer.
3. *Independence.* ε_i and ε_j are independent for all $i \neq j$.
4. *Independence of* X_j. ε_i and X_j are independent for all i and j, which says that the distribution of ε does not depend on the value of X.
5. *Normality.* ε_i are normally distributed for all i. This also implies that ε_i are independently and normally distributed with mean 0 and a common variance σ^2. The concept of normality is needed for inferences on parameters, but is not required to find least-square estimates.

When the five basic assumptions of the regression model are satisfied, OLS provides unbiased estimates of the regression coefficients α and β, which have minimum variance among all unbiased estimates. In other words, the least-squares estimators indeed yield the estimated straight line that has a smaller residual sum of squares than any other straight line. For this reason, OLS estimates are referred to as *best linear unbiased estimates* (BLUE). Any violation of these assumptions can reduce the validity of the OLS method. The greater the departure of the model from this set of assumptions, the less reliable is OLS. In such situations, one must use alternative estimation procedures depending on the type of violation of the above assumptions. One alternative estimation technique called generalized least squares (GLS) is described later in this section.

Multiple-regression techniques should be used to model a dependent variable instead of simple regression because (1) the dependent variable can be predicted more accurately if more than one independent variable is used and (2) if the dependent variable depends on more than one independent variable, a simple regression on a single independent variable may result in a biased estimate of the effect of this independent variable on the dependent variable. The theoretical model of multiple regression is basically the same as in simple regression. The only difference is that the dependent variable is assumed to be a linear function of more than one independent variable. For example, if there are three independent variables, the model is

$$y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \varepsilon \tag{5.32}$$

where X_1, X_2, X_3 = independent variables assumed to affect the dependent variable y

ε = random error term

$\alpha, \beta_1, \beta_2, \beta_3$ = estimated coefficients

Just as in the case of simple regression, the coefficients α and β_i are estimated by finding the value of each that minimizes the sum of the squared deviations for the observed values of the dependent variable from the values of the dependent variable predicted by the regression equation. Furthermore, in order to obtain least-square estimates, multiple regression must follow each of the five assumptions required by simple regression, with two added conditions. The first condition is that none of the independent variables can be an exact linear combination of any of the other independent variables. In other words, no one variable can be an exact multiple (or linear combination) of any other independent variable. For example, X_i cannot be written as aX_2. This situation is called multicollinearity. The second condition is related to degrees of freedom. Specifically, the number of observations N must exceed the number of coefficients being estimated. In practice, the sample size should be considerably larger than the number of coefficients to be estimated in order to obtain meaningful information about the underlying relationship.

A *time series analysis of monthly data* on volumes of water sold in consecutive billing periods can be used to estimate water-use models. However, the reliability of such models will depend on (1) the ability to disaggregate sales data into classes of similar users (e.g., single-family residential, multiunit residential, small commercial, large industrial); (2) the ability to separate (or account for) the seasonal effects and weather effects in the time series data; and (3) the ability of the estimation technique to deal with nonconstant error variance and correlation of model errors through time. A theoretical time series model can be written as

$$y_t = a + \sum_{i=1}^{N} b_i S_{i,t} + \sum_{j=1}^{M} c_j W_{j,t} + \sum_{k=1}^{P} d_k X_{k,t} + \sum_{r=1}^{R} e_r C_{r,t} + \mathring{a}_t \qquad (5.33)$$

where y_t = aggregate volume of water sold to a homogeneous class of customers during a monthly or bimonthly billing period t where $t = 1, \ldots, T$

a = model intercept

S_i = set of N seasonal variables that capture the seasonal variability of water use ($i = 1, \ldots, N$)

W_j = set of M weather variables that capture the effect of actual weather conditions on water use ($j = 1, \ldots, M$)

X_k = set of P "trend forming" variables that capture changes in water use unrelated to seasonal and weather effects ($k = 1, \ldots, P$)

C_r = set of R conservation variables ($r = 1, \ldots, R$)

ε_t = error term

b_i, c_j, d_k, e_r = coefficients to be estimated

The selection and definition of variables to represent the four types of systematic forces that affect aggregate water use over time are very important. The seasonal component in water-use data can be captured in many ways. Three possible specifications used in modeling time series water-use data include (1) a seasonal index, (2) a discrete step function, and (3) a Fourier series of sine and cosine terms.

A *seasonal index* is usually expressed as the average fraction of total annual water use to be expected during a given calendar month. This fraction can be estimated using the time series data on water use. For example, the value of the index in July can be obtained by dividing water use during the month of July by total annual use for each calendar year and then calculating the average value of the index for all years in the data set. The process is repeated for each calendar month until all 12 values of the seasonal index are obtained. The seasonal index is then used as a simple variable to capture the seasonal component of water use in Eq. (5.33). A *discrete step function* can be represented by 12 indicator variables corresponding to individual calendar months (e.g., $M_1, \ldots, M_{12}$, where $M_1 = 1$, if the month in the data is January, and $M_1 = 0$ elsewhere). When bimonthly data are modeled, six indicator variables would be created, one for each bimonthly period. In order to avoid multicollinearity, only $M - 1$ indicators should be specified, where M denotes the number of monthly or bimonthly periods. A *Fourier series* of sine and cosine terms is a harmonic function that can be applied to the data to generate a smooth sinusoidal cycle of seasonal effects. In the case of monthly data, the Fourier series may include six sine and cosine harmonics that can be written as

$$\sum_{h=1}^{6} \left(a_h \sin \frac{2\pi hm}{12} + b_h \cos \frac{2\pi hm}{12} \right) \qquad (5.34)$$

where a_h and b_h are coefficients to be estimated and m is the calendar month (m = 1 for January, m = 2 for February, etc.).

The cycle corresponding to h = 1 has a 12-month period. The cycles corresponding to h = 2 are harmonics of the 6-month period. All six harmonics represent the seasonal cycle of water use, which is periodic but not directly sinusoidal. Because the lower harmonics tend to explain most of the seasonal fluctuations, in most situations, it may be possible to omit higher-frequency harmonics in Eq. (5.34), thus representing the seasonal component as

$$\sum_{i=1}^{4} b_i S_{i,t} = b_1 \text{ SIN}(1) + b_2 \text{ COS}(1) + b_3 \text{ SIN}(2) + b_4 \text{ COS}(2) \quad (5.35)$$

where b_1, b_2, b_3, b_4 = coefficients to be estimated
SIN (1) = sin ($2\pi m/12$)
COS (1) = cos ($2\pi m/12$)
SIN (2) = sin ($4\pi m/12$)
COS (2) = cos ($4\pi m/12$)

The significance of each cycle is usually tested first. The cycles with insignificant amplitudes (i.e., b_i) can then be deleted from the equation.

Air temperature and rainfall are usually used to capture the effects of weather on water use. In most cases, these two variables will be correlated with the seasonal variables. Therefore, the weather variables should be measured as deviations from their normal values for each month (or billing period). Also, lagged weather variables can be used to take into account (1) the fact that the recorded consumption in any given month represents water use which took place during the current and the previous month and (2) the short-term memory in water use (e.g., water use in month t is affected by rainfall in month $t - 1$). The following weather variables can be included in the specification of the weather effects in Eq. (5.33): (1) deviation of monthly rainfall from monthly norms, (2) deviation of monthly average of maximum daily temperatures from monthly norms, (3) deviation of the number of days with precipitation greater than 0.01 in from monthly norms, and (4) deviation of cooling-degree days from monthly norms.

These deviations can be specified both as contemporaneous and lagged measurements. The *normal* values can be calculated for the period of the time series data or for weather data extending up to 30 years back (i.e., long-term averages).

In addition to properly measuring the seasonal and weather effects, it is necessary to include the effects of variables such as the number of customers, the price of water, the cost of wastewater disposal, and other factors such as income in residential sectors and productivity in nonresidential sectors. The changes in the *number of customers* can be incorporated by expressing the dependent variable in terms of water use per customer (by dividing total volume of water by the number of customers billed). The *price of water and wastewater disposal* should be included in the model. In modeling aggregate water-use data, it is difficult to determine what measure of price should be used. The relevant measure is the marginal price faced by an individual customer. This price is the same for all customers when a uniform rate structure is used. If increasing block rates are used, then the average price determined for an average consumption level can be used, so that only actual increases in the price of water and wastewater are captured by the price variable. All nominal values of price should be converted into constant dollars using the Consumer Price Index (CPI) for all items. *Median household income* should also be included among the variables of residential models if the data are available and expressed in constant dollars. Usually, household income statistics can be obtained for each quarter from the Internal Revenue Service. Monthly values of income can be obtained by interpolating the quarterly data. Several other variables that are known to influence water use can be omitted if the changes over time are minimal (e.g., average number of persons per customer connection, average lot size). If changes in such variables are significant, then these variables should be included in the model.

Finally, the effects of the conservation programs can be accounted for in the time series model by including an indicator variable which separates the data into pre- and postprogram periods. This variable takes on the value of 0 for all months before the program, and the value of 1 for the months after program implementation. The effects of other passive and active conservation measures such as a conservation-oriented plumbing code should be included in the model. This can be accomplished by introducing a variable that measures the cumulative number of new service connections sold after the code went into effect. The *error term* can be specified as additive or multiplicative. A logarithmic transformation of the water-use variable will result in a multiplicative error term. Such a transformation will often produce a better fit of the model than untransformed water use.

The potential problems with estimating the parameters of the time series regression model are endogeneity of the price variable (i.e., the price is related to the quantity of water used), nonconstant error variance (also known as heteroskedasticity), and autocorrelation. Most regression software packages have routines that attempt to correct for the problem of autocorrelation. Nonconstant error variance and endogeneity are problems best suited for alternative regression methods such as generalized least squares.

If customer-level monthly (or billing period) data on water use for a period of 2 to 4 years can be obtained and supplemented with information on customer characteristics and external factors such as price of water and weather, then a *pooled time series cross-sectional* (TSCS) data set can be constructed and used to estimate the parameters of a multiple-regression model. The theoretical form of the model may be written as

$$y_{it} = \beta_1 + \sum_{k=2}^{K} \beta_k X_{k,it} + \varepsilon_{it} \tag{5.36}$$

where y_{it} = monthly water use of customer i in month t (instead of months, billing periods may be used)
β_1 = intercept term
β_k = regression coefficients
X_k = set of independent variables that represent all possible systematic forces which affect that use
ε_{it} = error term

The types of independent variables of residential water-use models that may be used in Eq. (5.36) are illustrated in Table 5.10. Variables that explain use in other sectors can be found in Dziegielewski et al. (2002b). Measurements of these variables (or as many variables as possible) should be obtained for *each* customer and *each* time period using such sources of information as (1) telephone or mail surveys of customers in the sample, (2) real estate and tax assessor records, (3) aerial photographs, (4) "driveby" surveys of customer premises, (5) water and wastewater prices and rate structures, and (6) meteorological stations. Many variables will have values that are constant over time or will have only one observation in the time that is available. In the latter case, their values can be assumed constant over the period for which water-use data are obtained.

Once the pooled time series cross-sectional database is complete, an appropriate form of the functional relationship between water use and its determinants must be selected. Also, in the context of the structure of systematic forces, the analyst should consider the appropriate structure of the model error. *Residential water demand models* are often estimated using one of the following functional forms:

1. Linear model

$$y_{it} = \beta_1 + \sum_{k=2}^{k} \beta_k X_{k,it} + \varepsilon_{it} \tag{5.37}$$

in which the error is additive

2. Log-linear (or double-log) model

$$\log y_{it} = \log(\beta_1) + \sum_{k=2}^{k} \beta_k (\log X_{k,it}) + \log \varepsilon_{it} \tag{5.38}$$

with multiplicative error

3. Exponential model

$$\log y_{it} = \beta_1 + \sum_{k=2}^{k} \beta_k X_{k,it} + \varepsilon_{it} \tag{5.39}$$

with multiplicative (and/or exponential) error

TABLE 5.10 Important Explanatory Variables for Residential Water-Use Models

Family characteristic variables
1. Family size (number of persons)
2. Number of children under 18
3. Household income
4. Ownership of residence
Household fixture and appliance variables
1. Number of showers
2. Number of toilets
3. Washing machine (presence of)
4. Dishwasher (presence of)
5. Garbage disposal
Frequency of appliance use variables
1. Laundry loads per week
2. Dishwasher loads per week
Outdoor feature variables
1. Lot size
2. Lawn size
3. Total irrigated area
4. Automatic sprinkling system
5. Swimming pool
Frequency of outdoor use variables
1. Lawn and landscape watering per week
2. Car washing per week
3. Hosing of concrete (blacktop) surfaces
Price variables
1. Marginal price of water
2. Rate structure
3. Wastewater charge
Weather variables
1. Monthly average of maximum daily temperatures
2. Total monthly precipitation
3. Number of days with precipitation greater than 0.01 in
4. Cooling-degree days
Other variables
1. Age of the house
2. Type of sewerage system

Good examples of a pooled time series cross-sectional multivariate regression model can be found in Dziegielewski and Opitz (1988) and Dziegielewski et al. (1993).

Most frequently, pooled time series cross-sectional analyses use the linear functional form of the model and estimate parameters using the OLS estimation technique. However, the nature of pooled time series cross-sectional data often results in the violation of one or more of the basic regression assumptions. Two common problems that have been encountered in practical research are heteroskedasticity and autocorrelation. *Heteroskedasticity* usually arises from cross-sectional components of the data. The variance of the error term ε is not constant across all observations. When this occurs, the scale of the dependent variable and the explanatory power of the OLS model tend to vary across observations. *Autocorrelation* is usually found in time series data. The disturbances ε_i are not independent of each other; they are correlated. Time series data often display a "memory" such that variation is not independent from one period to the next. For example, an earthquake or flood may affect water use in a particular community for many periods following the actual event. Note, however, that it does not always take such a large disturbance to produce autocorrelated errors. If either one of these problems exists, the coefficient estimates of the OLS model no longer have minimum variance among all linear unbiased estimators. In other words, the OLS estimators are no longer the best linear unbiased estimates.

Several diagnostic tests for heteroskedasticity and autocorrelation are commonly available in statistical software packages. Perhaps the simplest diagnostic check is to plot the residuals. If heteroskedasticity and/or autocorrelation are detected, it is advisable to specify an alternative to the OLS model. One such model, the *generalized least squares* (GLS) regression model, can be shown to provide the best linear unbiased estimators under these conditions. Instead of minimizing the sum of squared residuals as in OLS estimation, the GLS procedure produces a more efficient estimator by minimizing a *weighted* sum of the squared residuals. Observations whose residuals are expected to be large because the variances of their associated disturbances are known to be large are given a *smaller* weight. Observations whose residuals are expected to be large because other residuals are large are also given smaller weights (Kennedy, 1985). In order to produce coefficient estimates using GLS, the variance-covariance matrix of the disturbance terms must be *known* (at least to a factor of proportionality). In actual estimating situations, however, this matrix is usually not known. A procedure called *estimated generalized least squares* (EGLS) can then be employed to estimate the variance-covariance matrix of the disturbances. EGLS estimators are no longer linear or unbiased, but because they account for the effects of heteroskedastic and autocorrelated errors, they are thought to produce better coefficient estimates.

The EGLS estimation technique is used to estimate what are called *error components* models of pooled time series and cross-sectional data (see Kennedy, 1985). The general specification for the random-effects model can be written as

$$y_{it} = \beta_0 + \sum_{i=1}^{K} \beta_k X_{k,it} + \varepsilon^* \tag{5.40}$$

where

$$\varepsilon = u_i + v_t + \varepsilon_{it}$$

Notice that this model has an overall intercept and an error term that consists of three components. The u_i represents the extent to which the ith cross-sectional units intercept differs from the overall intercept. The v_t represents the extent to which the tth time period's intercept differs from the overall intercept. The u_i and v_t are each assumed to be independently and identically distributed with a mean of 0 and variance of $\sigma_u{}^2$ and $\sigma_v{}^2$, respectively. The third component, ε_{it}, represents the traditional error term that is unique to each observation. All three error components are assumed to be mutually independent. The extent to which the intercept coefficients differ across cross-sectional units and across time is assumed to be randomly distributed. Because of this, the error components model is sometimes referred to as the *random effects model*. A good example of an error components model can be found in Chesnutt and McSpadden (1990).

5.8 ANALYSIS OF WATER SAVINGS

A precise measurement of water savings that can be attributed to various demand management programs is difficult because the observed water use often shows great variability among different users and it also significantly varies over time for the same user. For example, the amounts of water used inside and outside a residential home can vary substantially from month to month and from household to household. This variability is caused by many factors, including conservation practices. Because of the great variability in water use, the observed changes in water use over time, or differences in use between individual customers or groups of customers, may be caused by influences unrelated to the customer participation in a conservation program. Therefore, the most important consideration in measuring water conservation savings is the design of a measurement procedure that is capable of correctly measuring not only the changes in water use but also separating these changes into those caused by the program and those caused by changes in weather, prices, economic factors, and other confounding factors. The precision of the measurements of water savings depends on whether the study design was capable of isolating and controlling for (1) the characteristics of the conservation program that could significantly influence the results of the estimation of water savings, (2) the characteristics of the customer groups targeted by the program that could also influence the results, and (3) the characteristics that are external to both the conservation program design and the targeted customer groups. Research in evaluation designs identified a number of factors referred to as *outside effects* or *externalities* (Dziegielewski et al., 1993).

Once a conservation practice is adopted, the baseline demand that represents water use without the practice cannot be directly measured, and the unaltered demand has to be reconstructed somehow. In practice, all study designs employ comparisons of water-use behavior (and other customer characteristics in some cases) over time and/or between groups of customers. Possible types of comparisons are illustrated by Fig. 5.3. Implementation of a conservation program divides the time continuum into two periods, namely, pretreatment conditions and posttreatment conditions. It also divides the water users into two groups—the control group of nonparticipants and the treatment group of program participants. The conditions of a valid study design are achieved by a careful selection of a sample of water users and the use of proper methods of data analysis.

There are two basic approaches for estimating water conservation savings: statistical techniques and leveraged approaches. These two approaches will be discussed.

5.8.1 Statistical Estimation of Savings

Statistical comparison methods produce estimates of conservation savings by comparing water use between a participant group and a control group (or changes in water use before and after the program). The *comparison-of-means method* is derived from the statistical theory of randomized controlled experiments which utilizes a treatment and control design. Conservation savings are estimated as the difference in the mean level of water use between the treatment group and the control group, i.e.,

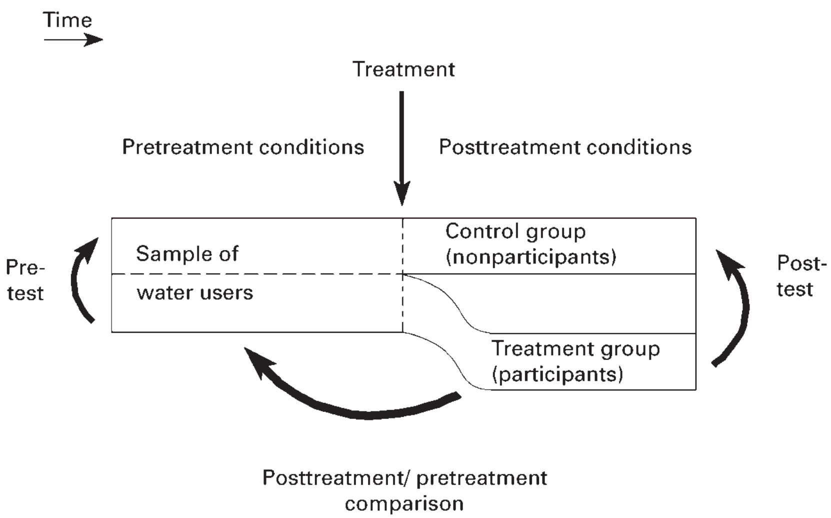


FIGURE 5.3 Evaluation designs for measuring water conservation savings.

$$d = \bar{q}_t - \bar{q}_c \tag{5.41}$$

where

$$\bar{q}_t = \frac{1}{n_1} \sum_{t=1}^{n_1} q_t \tag{5.42}$$

and

$$\bar{q}_c = \frac{1}{n_2} \sum_{c=1}^{n_2} q_c \tag{5.43}$$

where d = conservation effect (difference between means)
$\bar{q}_t$ = mean water use in treatment sample
$\bar{q}_c$ = mean water use in control sample
q_t = water use of customer t in treatment sample
q_c = water use of customer c in control sample
n_1 = number of customers in treatment sample
n_2 = number of customers in control sample

The data can be used to test whether the observed difference d can be attributed to chance, or whether it is indicative that the two samples come from populations of unequal means. Given that the parent population distribution of the differences in means is unknown or not normal, and that the population standard deviation of water use in each group (σ_t and σ_c) are unknown but assumed equal, the sampling distribution of the differences in mean water use should follow a t distribution for large samples. The *central limit theorem* (and the concept of repeated sampling) implies that for large sample sizes, the sampling distribution of the difference in means between groups will approach a normal distribution. This allows statistical inference about population parameters when the population distribution is unknown or not normal. However, in order for the sampling distribution of the difference in two means to be approximated by the normal distribution, the population standard deviations (σ_t and σ_c) must be known. If σ_t and σ_c are not known but are assumed equal, one may use the sample estimates of σ_t and α_c (s_t and s_c) and the t distribution for statistical inference. For large sample sizes, the t distribution will approximate the normal distribution.

A calculated t statistic can be used to test the hypothesis that the differences between mean water use of the treatment and control samples is 0. In order to calculate a t statistic, the *standard error of the difference between the two means* must be determined using the following formula:

$$S_d = \sqrt{\frac{(n_t - 1)\, s_t^2 + (n_c - 1)\, s_c^2}{n_t + n_c - 2} \; \frac{n_c + n_t}{n_c n_t}} \tag{5.44}$$

where S_d = estimated standard error of conservation effect d
s_t = standard deviation of water use in treatment sample
s_c = standard deviation of water use in control sample
n_t = number of customers in treatment sample
n_c = number of customers in control sample

The *t statistic* is then calculated as

$$t = \frac{\bar{q}_t - \bar{q}_c}{S_d} = \frac{d}{S_d} \tag{5.45}$$

Using the properties of the t distribution, one can test the null hypothesis stating that the true population mean of the treatment group is equal to the population mean of the control group, or the alternative hypothesis stating that the population means of the treatment and control groups are different. The value of the t statistic calculated by Eq. (5.45) uses the sample estimates to infer whether the difference d is large enough to reject the null hypothesis that the true difference in means is equal to 0. To test the null hypothesis, the resultant value of t must be compared to statistical tables of the t distribution. These tables may be found in any standard statistics textbook. Depending on the type of assertion to be made about the difference between means, one may use either a one-tail or two-tail (also called one-sided or two-sided) test of significance.

The two-tail test should be used to test the hypothesis that the difference between means, d, is 0, against the alternative hypothesis that the difference in means is positive or negative. The one-sided test of significance should be used to test the null hypothesis of no difference in means against the alternative hypothesis that, after the treatment, the mean water use in the treatment group is *lower* than the mean water use in the control group.

In order to produce reliable estimates, the comparison-of-means method must satisfy two requirements: (1) the two random variables representing water use in the treatment and control groups must be drawn from the same population distribution, and (2) the distribution is normal. The first assumption is often violated. In the classic experimental design on which the comparison-of-means method is based, the experimenter has careful control over all factors that might affect the variable under consideration. Therefore, by carefully designing the samples to be used in the experiment, any difference between the treatment and control groups can be attributed solely to the "treatment." However, the water planner is not likely to have complete control over the confounding factors that affect household water use. Although statistical theory suggests that randomly assigning households into treatment and control groups will result in groups that are less likely to be systematically different in terms of water use, this does not ensure that the two groups are identical with respect to household income, average number of persons per household, yard size, and many other factors. Therefore, unless a great deal of matching or sampling work is done, water use cannot be considered a random variable. It is related, if not caused, by the uncontrolled-for factors that differ between the treatment and control groups. In other words, one runs the risk of incorrectly attributing observed changes in water use to the treatment (e.g., a retrofit), when in fact they are caused by the different average values of external factors in each group. With respect to the second assumption, it must be stressed that empirical distributions of water use have most often been found not to follow a normal distribution. Typically, distributions of water use show a long right-hand tail, and thus do not conform to the symmetric bell-shaped appearance associated with the normal distribution. The violation of this assumption is not a fatal flaw, however, if a normalizing transformation of the data is used. For example, taking the log of water use should at least pull in the right-hand tail and minimize the leverage of "contaminated" or outlying data points or they can be screened out by rejecting values greater than $x + 3s$.

When adhering to a strict experimental design, the comparison-of-means method is more likely to produce meaningful and reliable results in situations where (1) the expected conservation effect is large when compared to mean water use, (2) the variance in water use is small compared to the conservation effect, (3) the mean and variance in water use are very similar (in terms of size) for both groups prior to treatment, and (4) the sample sizes in the treatment and control groups are large. The comparison-of-means method can produce reliable and informative results if used in conjunction with experimental designs and large sample sizes.

Multivariate regression models represent the most sophisticated method of comparing water-use data over time or between groups of customers while controlling for the effects of a large number of external factors. One can choose from a variety of regression methods depending on the types of available data and the acceptable level of estimation complexity. The statistical models described in Section 5.7.2 can be used for this purpose.

5.8.2 Time Series Analysis of Conservation Effects

A time series of the volumes of water sold in consecutive billing periods can be used to measure conservation effects of full-scale programs while controlling for external influences other than the conservation program. The reliability of the estimates will depend on (1) the ability to disaggregate sales data into classes of similar users (e.g., single-family residential, small commercial), (2) the ability to separate (or account for) the seasonal effects and weather effects in the time series data, and (3) the ability of the estimation technique to deal with nonconstant error variance and correlation of model errors through time. The effects of a conservation program under investigation can be measured by including an indicator variable which separates the time series data into pre- and postprogram periods.

However, it is also important to capture the effects of other passive and active conservation measures which are adopted by water customers independently of the program under evaluation.

If customer-level monthly (or billing period) data on water use for a period of 2 to 4 years can be supplemented with information on customer characteristics, and such external factors as price of water and weather conditions, then a pooled time series cross-sectional data set can be constructed and used to estimate the parameters of a multiple-regression model. The important explanatory variables will depend on the customer class. In the residential sectors they usually include information on family characteristics, household features and appliances, frequency of appliance use, outdoor features, and frequency of outdoor uses. This measurement technique can produce very accurate estimates of actual water savings for some programs, especially those targeting the residential sector.

The major drawback of the multivariate regression models is that they are relatively expensive and time-consuming. They require large amounts of data and a lot of expertise from the analyst. They also need large sample sizes and are less appropriate for nonresidential water users.

5.8.3 End-Use Accounting System

The accumulating experience in the evaluation of conservation programs indicates that it is very difficult to obtain measurements of water savings with a high level of precision using a single best method. The most precise estimates can be achieved by taking advantage of the strong features of the statistical methods with engineering methods. The known strengths, weaknesses, and biases of each approach can be used to narrow down the confidence bands surrounding the actual water savings.

Engineering (or mechanical) estimates are obtained using laboratory measurements or published data on water savings per device or conservation practice. These data can be combined with assumptions regarding the magnitude of factors expected to impact on the results of the conservation program in order to generate estimates of program savings. However, the resultant estimates can be very sensitive to the underlying assumptions and relationships. For example, the savings resulting from the installation of an ultralow-flush toilet replacing a standard toilet will depend on assumptions regarding flushing volumes and frequency of flushing. The resultant savings can range from 19.5 gal per person per day (3.9 gal $\times$ 5 flushes per person per day) to 7.6 gal per person per day (1.9 gal $\times$ 4 flushes per person per day). The high estimate is almost 3 times greater than the low estimate. The validity of the assumptions used in the above example can easily come under attack, since they rely on subjective conclusions and a great deal on the professional judgment of the engineer or analyst.

Despite their obvious shortcomings, engineering estimates may be considered appropriate for providing preliminary estimates of potential savings when field measurements are not available. They can also be used to verify statistical estimates by setting limits on a possible range of savings. However, they become most useful in leveraged techniques where they can be used to augment and strengthen statistical models.

The most promising method of leveraging information involves the use of information from one approach within the procedures of another. For example, engineering estimates or special metering measurements can be used as independent variables in statistical models. Statistical models of urban water demands do not accommodate the needs of planning and evaluation of demand management programs because of an inability to disaggregate water demands down to the end-use level. Because many demand-side programs target specific end uses, the absence of end-use water demand models severely impairs the development of effective demand management policies. Without adequate end-use models, the effects of various demand management programs cannot be measured precisely. In order to enhance the ability of water planners to formulate, implement, and evaluate various demand management alternatives, it is necessary to disaggregate the usually observed sectoral demands during a defined season of use into the applicable end uses. Only such a high level of disaggregation will permit water planners to make all necessary determinations in estimating water savings of various programs.

The first step is to disaggregate the observed water demands into their specific components or end uses. Figure 5.4 illustrates how water demand of a homogeneous sector of water users can be disaggregated into its seasonal and end-use components. A rational representation of each end use is made using a *structural end-use equation* of the following form:

$$q = [(M_1S_1 + M_2S_2 + M_3S_3)U + KF]A \tag{5.46}$$

where q = average quantity of water in a given end use
M_1, M_2, M_3 = efficiency classes of end use (design parameters)
S_1, S_2, S_3 = fractions of end uses within efficiency class
U = usage rate (or intensity of use)
K = average flow rate of leaks
F = fraction of end uses with leaks (incidence of leaks)
A = presence of end use in a given sector of users

A graphical representation of this structural end-use relationship is given in Fig. 5.5. An application of Eq. (5.46) to the toilet end use in the residential sector would require the knowledge of all end-use parameters. An example of the uses of this equation for analyzing the toilet end use is presented in Table 5.11. Other end uses and effects of improvements in their efficiency can be estimated using similar parameters and data.

The structural end-use relationship [Eq. (5.46)] is dictated by the need to distinguish between changes in water demand caused by active and passive demand

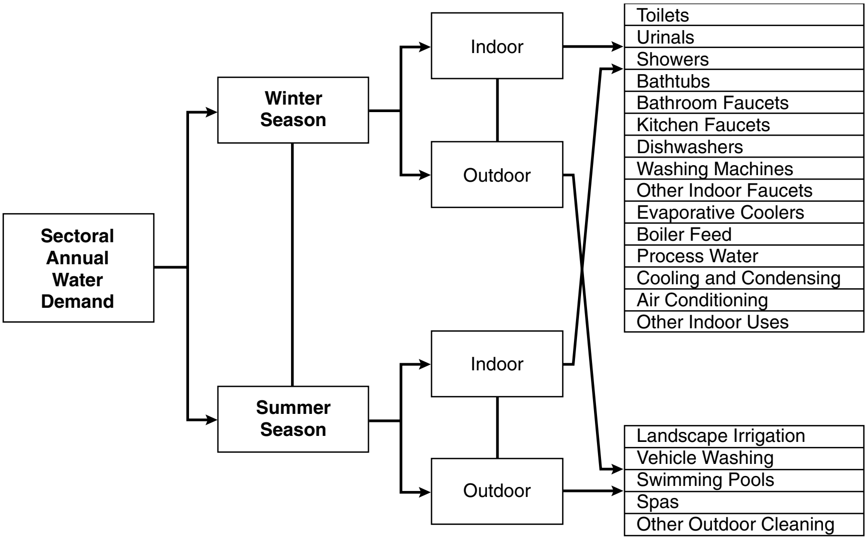


FIGURE 5.4 Disaggregation of annual water demands.

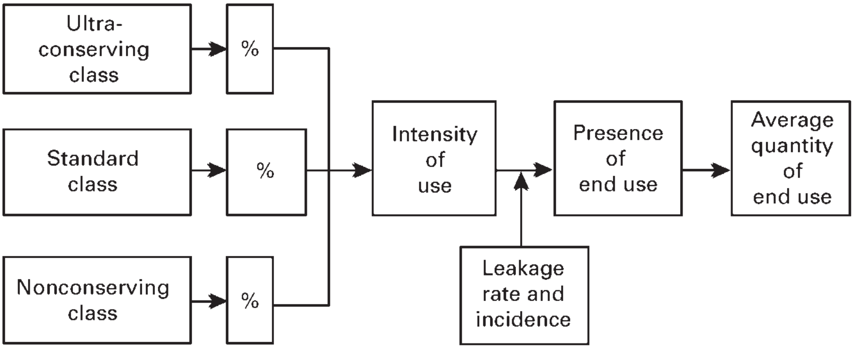


FIGURE 5.5 Structural end-use relationships.

management programs from the changes caused by other factors. The structure of end uses which exists in the service area at any point in time will not remain constant over the planning horizon. It will change in response to changes in the determinants of water use such as income, household size, and housing density. The effects of various interventions of demand management programs must be counted relative to the baseline forecasts of water demands, which capture the effects of the relevant external factors. The parameters of each end use are affected by the external factors. For example, changes in price will cause a decrease in the incidence of leaks in the short run and will affect the distribution of end uses among the classes

TABLE 5.11 Example of Estimating Water Savings with End Use Accounting System (Toilet End Use)

End-use parameter	Before change	After change	Net effect (savings)
Inefficient class rate	5.50	5.50	
Inefficient class fraction	0.35	0.25	−0.10
Standard class rate	3.50	3.50	
Standard class fraction	0.55	0.50	−0.05
Efficient class rate	1.60	1.60	
Efficient class fraction	0.10	0.25	+0.15
Intensity of use, fpd	14.00	14.00	
Presence of end use	1.00	1.00	
Leakage rate, gpd	20.00	20.00	
Incidence of leaks	0.15	0.15	
Average quantity, gpd	59.10	52.35	−6.75

of efficiency in the long run. The other two parameters in the end-use equation (i.e., intensity and presence) also will be affected by changes in price.

For example, changes in price will cause a decrease in the incidence of leaks in the short run and will affect the distribution of end uses among the classes of efficiency in the long run. The other two parameters of the end-use equation (i.e., intensity and presence) also will be affected by changes in price. The ideal forecasting model would be capable of predicting the parameters of the structural end-use equations as a function of such influencing variables as price, income, household size, housing density, and weather. For example, in the case of lawn irrigation end use, the parameters of Eq. (5.46) should be modeled as functions of the explanatory variables.

The structure of the end-use equation allows the planner to estimate the net effects of long-term conservation programs by tracking the values of end-use parameters over time. It also accommodates the handling of such issues as interaction and overlapping of multiple programs, customer-initiated conservation effects, and the relationships between long-term and short-term (e.g., drought emergency) programs. Each of these effects creates problems in measuring the effectiveness of efficiency improvement programs.

5.9 SUMMARY

Urban water supply systems are designed to deliver water "on demand" as the system operators have no direct control over the quantity of water taken out by the customers. Accordingly, water demands are usually taken as given quantities that have to be matched with supplies. However, statistical studies of historical rates of water use have shown that water use fluctuates in response to various influencing factors. As a result, the per capita use rates in urban areas can fluctuate from year to year in response to changing weather conditions and other conditions. The usage rate also exhibits strong long-term trends caused by changes in the mix of housing types, commercial activities, new growth, prices of water, and other factors. Because these influencing forces are likely to operate in the future, future demand cannot be assumed to be a simple product of the projected population and the historical rate of per capita water use. The adoption of water conservation measures, price of water, housing density, urban growth policies, and types of landscaping are important determinants of water use and can be viewed as instruments for managing water demands.

Analytical techniques described in this chapter can help water planners to isolate and quantify the effects of many explanatory factors, and the impacts on water demand of anticipated changes in these factors. An even greater challenge will be the development of models that are able to estimate the long-term effects of water-demand management programs and policies with a high degree of confidence.

5.10 REFERENCES

American Water Works Association. 1986. *1984 Water Utility Operating Data.* Denver, Colo.

Chesnutt, T. W., and C. N. McSpadden. 1990. *A Model-Based Evaluation of the Westchester Water Conservation Program.* Metropolitan Water District of Southern California, Los Angeles.

Cochran, W. G. 1963. *Sampling Techniques.* John Wiley, New York.

DeOreo, W. B., J. P. Heaney, and P. Mayer. 1996. Flow Trace Analysis to Assess Water Use. *Journal AWWA,* 88(1):79–90.

DeOreo, W. B., P. W. Mayer, E. M. Opitz, B. Dziegielewski, J. C. Kiefer, W. Y. Davis, and J. O. Nelson. 1998. *Residential End Uses of Water: Final Report.* AWWA Research Foundation, Denver, Colo.

Dziegielewski, Benedykt, and Eva Opitz. 1988. *Phoenix Emergency Retrofit Program: Impacts on Water Use and Consumer Behavior.* Phoenix Water and Wastewater Department, Phoenix.

Dziegielewski, Benedykt, Eva Opitz, and Dan Rodrigo. 1990. *Seasonal Components of Urban Water Use in Southern California.* Metropolitan Water District of Southern California, Los Angeles.

Dziegielewski, Benedykt, Eva M. Opitz, Jack C. Kiefer, and Duane D. Baumann. 1993. *Evaluation of Urban Water Conservation Programs: A Procedures Manual.* American Water Works Association, Denver, Colo.

Dziegielewski, B., S. Sharma, T. Bik, X. Yang, and H. Margono. 2002a. *Analysis of Water Use Trends in the United States: 1950–1995. Project Completion Report.* Southern Illinois University, Carbondale, Ill. (forthcoming).

Dziegielewski, B., S. Sharma, T. Bik, X. Yang, H. Margono, and R. Sa. 2002b. *Predictive Models of Water Use: Analytical Bibliography.* Southern Illinois University, Carbondale, Ill. (forthcoming).

Fowler, Floyd J. Jr. 1988. *Survey Research Methods.* Sage Publications, Newbury Park, Calif.

Kennedy, P. 1985. *A Guide to Econometrics.* MIT Press, Cambridge, Mass.

Kish, Leslie. 1965. *Survey Sampling.* John Wiley, New York.

Schlenger, Donald. 1991. Current Technologies in Automatic Meter Reading. *Waterworld News* 7(3):14–17.

Solley, Wayne B., Robert R. Pierce, and Howard A. Perlman. 1998. *Estimated Use of Water in the United States in 1995.* U.S. Geological Survey Circular 1200, Washington, D.C.

USEPA. 1997. *Community Water System Survey,* Vol. II.

USEPA. 2000. *Factoids: Drinking Water and Ground Water Statistics for 2000.* EPA 816-k01-004. *www.epa.gov/safewater/data/00factoids.pdf.*

CHAPTER 6
WATER PRICING AND DROUGHT MANAGEMENT

Messele Z. Ejeta
California Department of Water Resources
Sacramento, California

Larry W. Mays
Department of Civil and Environmental Engineering
Arizona State University
Tempe, Arizona

6.1 INTRODUCTION

Droughts continue to rate as one of the most severe weather-induced problems around the world. Global attention to natural hazard reduction includes drought as one of the major hazards. Changnon (1993) gave seven lessons or truths that have emanated out of studying the major droughts from 1932 to 1992 in the United States. These lessons are summarized below:

1. A major drought is a pervasive condition affecting most portions of the physical environment as well as the socioeconomic structure.
2. Droughts are a major but unpredictable part of the climate of all parts of the United States. Moreover, they occur infrequently, and this results in a decay in the attention to drought preparedness and mitigation.
3. Responses and adjustments to drought problems can be sorted into two classes: (*a*) short-term fixes and (*b*) long-term improvements.
4. Although many long-term adjustments have been made as a result of the major droughts of the last 60 years, many factors make today's society generally more vulnerable to drought than ever before.

5. Agriculture, in general, cannot escape from experiencing major drought losses in the future, even with healthier crop strains and increased irrigation.
6. Opportunities for improvement in water management exist and could make the nation's water resources more impervious to drought. However, many water-related problems are localized and at the substate scale and often do not get needed attention.
7. Drought is ubiquitous: everything and everybody is affected, and yet no one (everyone) is in charge.

The shortage of water supply during drought periods is such a significant factor for the general welfare that its effect cannot be easily undermined. Domestic water supply shortages during these periods in particular have been crucial in some cases, and as a result, various measures targeting different means of reducing water demand during such periods were initiated by different water supply agents. These measures, which may be considered semiempirical to empirical, included water metering, leak detection and repair, rate structures, regulations on use, educational programs, drought contingency planning, water recycling and reuse, and pressure reduction. Such efforts are collectively termed water conservation, although there has not been a uniform definition among authors.

On the other hand, different researchers and scientists have tried to develop more scientific methods for water conservation during drought periods. These methods have been aimed at water conservation through price increases of the water supply to the customers. The results elucidated the fact that water is more of a commodity than it is a public resource. However, the several models developed so far which relate reduction in demand for water due to the increase in its price, through price elasticity, used different variables that range from the income of the customers to hydrologic conditions. The relations developed used regression analysis, and as a result the differences and the variations of the variables considered are significant enough that the estimated demand is subject to uncertainty. Thus the demand may be better expressed by an estimated value and a probability distribution.

The basics of the *price-elasticity approach* presumes that the demand can be adjusted to the available supply. This may happen on an average basis; however, the demand has a random distribution about the available supply. By similar reasoning, the available supply corresponding to a given return period of weather conditions may have a random distribution about the expected value. All the aforementioned uncertainties call for risk evaluation to determine the probability that the demand exceeds the available supply, for water supply project planning. Conversely, the price of water supply for a given tolerable risk level can be determined.

This chapter first discusses various efforts reported in the literature for water conservation and then culminates with the new idea of a *risk-based approach*. The different water-conservation practices are briefly discussed, giving coverage of the price-elasticity formulation. The basic reasons that make it necessary for a risk-based approach are described. Some risk-level indices, which have been used for the evaluation and prediction of a drought period, are given and their limitations are explained. A method for evaluation of the damage associated with certain levels of drought severity is developed. This new approach relates the risk, the price, and the return period. It is found through this relationship that risk is sensitive to the return period and to the price changes.

6.2 BACKGROUND OF WATER CONSERVATION

6.2.1 Drought Management Options

Experiences from past droughts have shown that the action of water managers can greatly influence the magnitude of the monetary and nonmonetary losses from drought. There have been a variety of *drought management options* that have been undertaken in response to anticipated shortages of water, which can be categorized as (1) demand reduction measures, (2) improvements in efficiency in water supply and distribution system, and (3) emergency water supplies (Dziegielewski, 1986). A topology of drought management options is given in Table 6.1.

Not only is water conservation necessary during drought periods, but the economic merits are also important to consider. In the United States, federal mandates urge that opportunities for water conservation be included as a part of the economic evaluation of proposed water supply projects (Griffin and Stoll, 1983). Water conservation during drought periods, however, requires important attention because our demand of water may exceed the available resources in the demand environment. Conservation may be achieved through different activities. According to the U.S. Water Resources Council (1979a), these activities include, but are not limited to,

1. Reducing the level and/or altering the time pattern of demand by metering, leak detection and repair, rate structure changes, regulations on use (e.g., plumbing codes), education programs, drought contingency planning
2. Modifying management of existing water development and supplies by recycling, reuse, and pressure reduction
3. Increasing upstream watershed management and conjunctive use of ground and surface water (Griffin and Stoll, 1983)

The effort to conserve water started out with metering rather than providing a flat rate. Both domestic and sprinkling demands reduced significantly as a result of the introduction of water meters (Hanke, 1970). Grunewald et al. (1976) stated: “Traditionally, water utility managers have adjusted water quantity [rather] than prices as changes in demand occurred.”

TABLE 6.1 A Topology of Drought Management Options

I. Demand Reduction Measures
  1. Public education campaign coupled with appeals for voluntary conservation
  2. Free distribution and/or installation of particular water-saving devices:
     - 2.1 Low-flow showerheads
     - 2.2 Shower flow restrictors
     - 2.3 Toilet dams
     - 2.4 Displacement devices
     - 2.5 Pressure-reducing valves
  3. Restrictions on nonessential uses:
     - 3.1 Filling of swimming pools
     - 3.2 Car washing
     - 3.3 Lawn sprinkling
     - 3.4 Pavement hosing
     - 3.5 Water-cooled air conditioning without recirculation
     - 3.6 Street flushing
     - 3.7 Public fountains
     - 3.8 Park irrigation
     - 3.9 Irrigation of golf courses
  4. Prohibition of selected commercial and institutional uses:
     - 4.1 Car washes
     - 4.2 School showers
  5. Drought emergency pricing:
     - 5.1 Drought surcharge on total water bills
     - 5.2 Summer use charge
     - 5.3 Excess use charge
     - 5.4 Drought rate (special design)
  6. Rationing programs:
     - 6.1 Per capita allocation of residential use
     - 6.2 Per household allocation of residential use
     - 6.3 Prior use allocation of residential use
     - 6.4 Percent reduction of commercial and institutional use
     - 6.5 Percent reduction of industrial use
     - 6.6 Complete closedown of industries and commercial establishments with heavy uses of water

II. System improvements
  1. Raw water sources
  2. Water treatment plant
  3. Distribution system:
     - 3.1 Reduction of system pressure to minimum possible levels
     - 3.2 Implementation of a leak detection and repair program
     - 3.3 Discontinuing hydrant and main flushing

TABLE 6.1 (*Continued*)

III. Emergency water supplies
- 1. Interdistrict transfers:
  - 1.1 Emergency interconnections
  - 1.2 Importation of water by trucks
  - 1.3 Importation of water by railroad cars
- 2. Cross-purpose diversions:
  - 2.1 Reduction of reservoir releases for hydropower production
  - 2.2 Reduction of reservoir releases for flood control
  - 2.3 Diversion of water from recreation water bodies
  - 2.4 Relaxation of minimum streamflow requirements
- 3. Auxiliary emergency sources:
  - 3.1 Utilization of untapped creeks, ponds, and quarries
  - 3.2 Utilization of dead reservoir storage
  - 3.3 Construction of a temporary pipeline to an abundant source of water (major river)
  - 3.4 Reactivation of abandoned wells
  - 3.5 Drilling of new wells
  - 3.6 Cloud seeding

Source: Dziegielewski, 1986.

In general, some of the major measures followed for water-conservation efforts with references are as follows:

1. Use restrictions (no car washing, or hosing down of sidewalks, alternate-day lawn and garden watering and the like) (Moncur, 1989).

2. Increasing rate structures, also called inverted-block rates, inclining-block rates, increasing blocks, inverted pyramid rates (Jordan, 1994).

3. A lump-sum charge and a commodity charge per unit volume imposed in addition to the normal rates (Carver and Boland, 1980).

4. Allowing the market process to operate, that is, adopting marginal cost pricing, even for normal periods, rather than averaging price (Moncur, 1989).

5. Attempting to decrease the amount used by industries by trying to utilize existing technology to design and install production processes using less water per unit of output (Grebenstein and Field, 1979).

6. Reducing withdrawals for production processes by recycling (Grebenstein and Field, 1979).

7. Passing water-conservation acts, requiring builders to install ultralow-flow fixtures in all new projects (Jordan, 1994).

8. Forcing the public to be bound to treated wastewater for new recreational use. Officials in Phoenix, for instance, considered not allowing new recreational lakes unless treated wastewater was used (Maddock and Hines, 1995).

Increasing the price of domestic water supply has been a focus of several studies. These studies were conducted to analyze the effect of urban water pricing and how it contributes to water conservation during a drought period (Agthe and Billings, 1980; Moncur, 1989). However variations have been observed in the approaches followed. According to Jordan (1994), water pricing is an effective way of conserving water, compared to the other measures mentioned above. An increase in the price of water contributes to water conservation because of the fact that customers have limited money. For every percent increase in the price, there is some decrease in the demand, which is explained through price elasticity. A significant number of studies have been undertaken in different regions to determine price elasticity associated with pricing. Section 6.2.2 explains price elasticity.

6.2.2 Price Elasticity of Water Demand

The elasticity of demand is the responsiveness of consumers' purchases to varying price. The most frequently used elasticity concept is *price elasticity,* which is defined as the percentage change in quantity taken if price is changed 1 percent. Young (1996) states that "the price elasticity of demand for water measures the willingness of consumers to give up water use in the face of rising prices, or conversely, the tendency to use more as price falls." Two different ways have been followed to formulate the price elasticity of demand for water: one based upon average price and the other based upon marginal price. Agthe and Billings (1980) state that the elasticity determined based upon average price overestimates the result. Therefore they recommend (as do several others) that the marginal price be used.

Howe and Linaweaver (1967) defined the *price elasticity of water* as

$$\eta_P = \frac{\Delta d}{\bar{d}} \div \frac{\Delta p}{\bar{p}} \tag{6.1}$$

where η_P = price elasticity
$\bar{d}$ = the average quantity of water demanded
$\bar{P}$ = average price
Δd = change in demand
ΔP = change in price

For a continuous demand function, the following more general formula is applicable.

$$\eta_P = \frac{dd}{d} \div \frac{dP}{P} \tag{6.2}$$

Table 6.2 is a summary of some of the values of price elasticity of water demand reported in the literature.

The use of the price elasticity of water has been applied to some cities with some important achievements having been obtained. The following schematic may depict the general trend of this principle, as derived from the conclusion reached by Jordan (1994):

$$\uparrow \text{(Price)} \rightarrow \downarrow \text{(water demand)} \ \& \uparrow \text{(revenue)} \tag{6.3}$$

An increase by less than 40 percent of the price resulted in a 10 percent decrease in the demand in Honolulu—the announced goal of the restrictions imposed in the drought episodes of 1976 to 1978 and in 1984 (Moncur, 1987). This was achieved using a price elasticity of only −0.265. In Tucson, Arizona, an inverted rate structure was claimed to have been credited with reducing public demand from about 200 gal per capita per day (gcd) to 140–160 gcd (Maddock and Hines, 1995).

The way in which water utilities are structured is probably the most important factor, which complicates the study of price elasticity. For instance, some customers who own homes or who pay water bills, more or less, react to the price change, whereas those who rent apartments or who do not pay for water bills are almost indifferent to it. Furthermore, water necessity for residential, commercial, and industrial purposes are not equally important. Because of this reason, different researchers had to study demand elasticity by categorizing water distribution systems for industrial, commercial, and residential uses. The demand patterns under these categories are not uniform. One of the most comprehensive studies on price elasticity of water demand done by Schneider and Whitlatch (1991) for six user categories (residential, commercial, industrial, government, school, and total metered) showed different results for these categories. Residential water use is further complicated by different factors: many residents who rent housing do not pay for water and as such are indifferent to demand regulations; the patterns for indoor and outdoor water demand differ quite significantly and hence necessitate different approaches of demand analysis. The climatic conditions of a given area and the time of the year are also worth mentioning. These are probably the reasons why apparently different elasticity values are reported for the eastern and the western United States and for winter and summer uses.

From the studies enumerated so far, a general conclusion is reached: demand is elastic to price increase. Almost all research has reinforced this hypothesis. However, differences exist between the elasticity values calculated for different geographic locations. For instance, Howe (1982) obtained values of −0.57 and −0.43 for the eastern and the western United States, respectively. On the other hand, no clear consistency exists in the way that elasticity is calculated: some use average price, some use marginal price, and still some include the intramarginal rate structure. Although some of the studies targeted alleviating water shortage problems during drought periods, they did not approach the problem from the perspective of risk analysis.

TABLE 6.2 Summary of Some of the Price Elasticity Values

No.	Researchers	Research area	Year	Estimated price elasticity	Estimated income elasticity	Remarks
1	Howe & Linaweaver	Eastern U.S.	1967	−0.860		
2	Howe & Linaweaver	Western U.S.	1967	−0.52		
3	Wong	Chicago	1972	−0.02	0.20	
4	Wong	Chicago suburb	1972	−0.28	0.26	
5	Young	Tucson	1973	−0.60 to −0.65		Exponential and linear models used
6	Gibbs	Metropolitan Miami	1978	−0.51	0.51	Elasticity measured with the mean marginal price
7	Gibbs	Metropolitan Miami	1978	−0.62	0.82	Elasticity measured with the average price
8	Agthe & Billings	Tucson	1980	−0.27 to −0.71		Long-run model
9	Agthe & Billings	Tucson	1980	−0.18 to −0.36		Short-run model
10	Howe		1982	−0.06		
11	Howe	Eastern U.S.	1982	−0.57		
12	Howe	Western U.S.	1982	−0.43		
13	Hanke & de Maré	Malmö, Sweden	1982	−0.15		
14	Jones & Morris	Metropolitan Denver	1984	−0.14 to −0.44	0.40 to 0.55	Linear and log-log models used
15	Moncur	Honolulu	1989	−0.27		Short-run model
16	Moncur	Honolulu	1989	−0.35		Long-run model
17	Jordan	Spalding County, Georgia	1994	−0.33		A price elasticity of −0.07 was also reported for no rate structure, but increased price level

6.2.3 Demand Models

It is important to have demand related to the drought severity. Several studies have expressed demand as a function of different variables. Mays and Tung (1992) gave a general form of demand models as

$$d = f(x_1, x_2, \ldots, x_k) + \varepsilon \tag{6.4}$$

where f is the function of explanatory variables $x_1, x_2, \ldots, x_k$ and ε is a random error (random variable) describing the joint effect on q of all the factors not explicitly considered by the explanatory variables.

Several explicit linear, semilogarithmic, and logarithmic models have been developed through different researches. Billings and Agthe (1980), for example, gave the following water demand function for Tucson, Arizona (notations modified to fit the notations adopted for this study).

$$\ln d = -7.36 - 0.267 \ln P + 1.61 \ln I - 0.123 \ln \mathrm{DIF} + 0.0897 \ln W \tag{6.5}$$

where d = monthly water consumption of average household, 100 ft^3
P = marginal price facing average household, cents per 100 ft^3
DIF = difference between actual water and sewer use bill minus what would have been paid if all water was sold at marginal rate, \$
I = personal income per household, \$/month
W = evapotranspiration minus rainfall, in

Equation (6.5) implicitly relates demand to the hydrologic index, W. The positive coefficient of W shows that demand increases exponentially with W, which indirectly indicates increases of demand with the dryness of weather conditions. The general trend of the average demand with the return period, therefore, may be shown as given by the demand curve in Fig. 6.1. Demand increases with the return period of the drought severity because the more severe the drought, the more the customers are prompted to use more water. Different demand curves are illustrated in Fig. 6.2 for different price levels. As shown in this figure, the higher the price, the lower the demand for a given hydrologic conditions.

Equation (6.5) may be rearranged as

$$d = 0.00006362P^{-0.267}I^{1.61}(\mathrm{DIF})^{-0.123}W^{0.0897} \tag{6.6}$$

or in more general terms,

$$d = a'P^{b'}I^{c'}(\mathrm{DIF})^{d'}W^{e'} \tag{6.7}$$

where a', b', c', d', and e' are constants. The price elasticity of demand for Eq. (6.6) is -0.267. Therefore, changing the price while keeping the other variables constant results in different average demand values, $\bar{d}_{P_i}$. Again, varying W while keep-

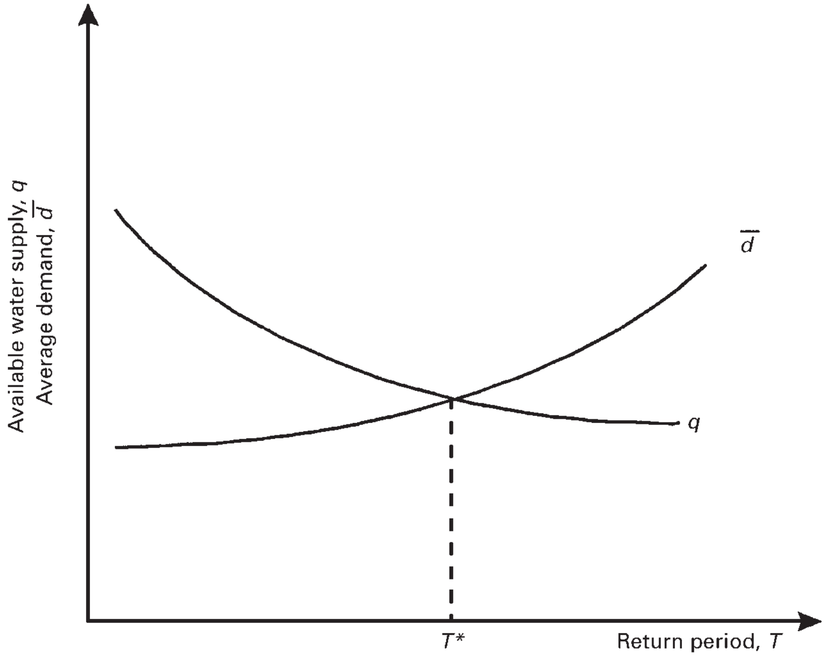


FIGURE 6.1 Water supply availability and average demand as related to the return period, *T*.

ing the other variables constant gives a general relation of the average demand associated with the return period *T*.

As given in Eq. (6.7), it can be seen that the demand d is related to the hydrologic index *W*, which is also related to the return period. The available supply (flow) q is also related to the return period (Hudson and Hazen, 1964). Thus the general relationships between demand and return period and supply and return period which are shown in Fig. 6.1 are based on these trends.

6.2.4 The Need for a Risk-Based Approach

A few past studies analyzed risk only by defining it as the monetary (financial) loss. They did not consider the risk as the probability of the supply not meeting the demand. To make these two connotations of risks distinctive, the terms *financial risk* and *probabilistic risk* are introduced and used differently. Many of the previous studies on risk did not explicitly define financial and probabilistic risks.

Financial risk can be simply stated as the monetary loss associated with a certain damage. *Probabilistic risk,* which is explicitly used in this paper, may be formulated as the probability $p(\ )$ that the demand d_{P_i} at price P_i for level i exceeds the available supply q_T, expressed as

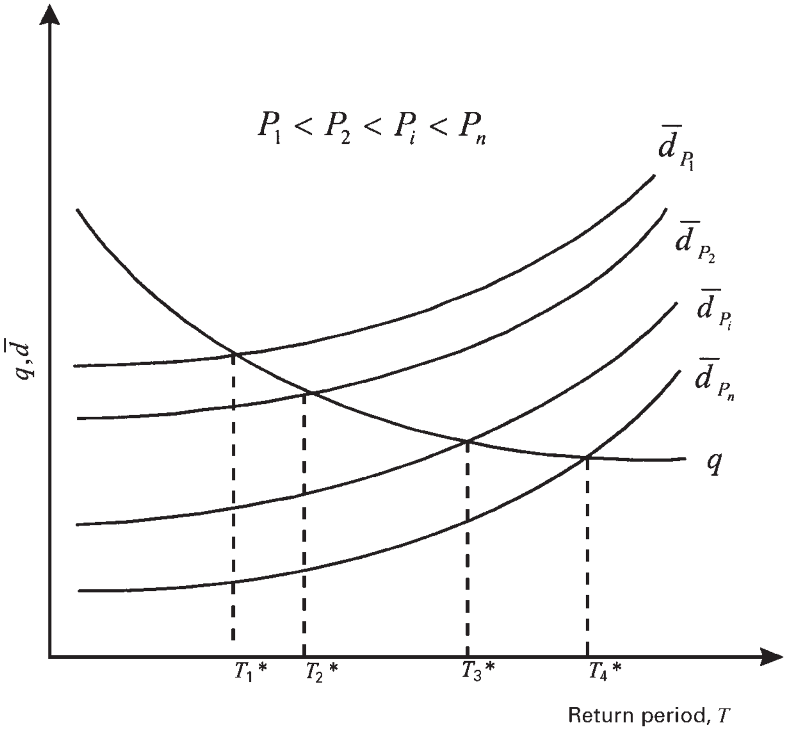


FIGURE 6.2 Water supply and average demands for different price values as related to T.

$$\text{Risk} = p(d_{P_i} > q_T) \tag{6.8}$$

Municipal water supply shortage problems have been manifesting themselves in different regions at different times for a long time. A study by Dixon et al. (1996) for California showed that projections of future water supply and demand (including environmental uses) indicate that the gap between supply and demand will widen to 4.1 million acre·ft in average water years and 7.4 million acre·ft in drought-water years by 2020. In 1977 in Fairfax County, Virginia, the drought was so severe that drastic measures such as the closing of schools and businesses were actively being considered (Sheer, 1980).

Two major groups of actions are undertaken by water agencies in order to avert some serious consequences of impending water shortages caused by droughts: (1) measures that reduce demands and (2) measures that enhance existing supplies. Developing practical methods for determining the necessary prices and devising structures of water rates that would achieve the desired reductions in water use are the most critical needs for establishing effective drought pricing policies (Dziegielewski et al., 1991).

There are different uncertainties involved with either one of these measures. In trying to reduce demands by increasing the price, uncertainty is involved in that the demand volume may not be equal to the limited available supply. This is simply because the demand depends on so many factors that cannot be totally controlled, irrespective of the price increase. On the other hand, enhancing the existing supply may cost more than the risk of not undertaking this task at all. By way of risk analysis, it is possible to optimize between the economic loss and the cost of enhancing the existing supply—such as emergency supply construction. Many scientists in different professions agree that the level of risk as a decision support system is a good indicator for sound decisions. Decisions in which the effects are portrayed relatively in the long run may finally result in adverse effects. Such effects are incurred at the expense of nonconservative risk-level designs.

Suter (1993) gave the following reasons for using the risk assessment approach for decision making:

1. Cost of estimating all environmental effects of human activities is impossibly high.
2. Regulatory decisions must be made on the basis of incomplete scientific information.

He concluded that a risk-based approach balances the degree of risk to be permitted against the cost of risk reduction and against competing risks. Lansey et al. (1989) also suggested that reliability analysis (a complement of risk analysis) be viewed as an alternative to making a decision without an analytical structure.

It has not been a common practice by responsible bodies to systematically incorporate in the decision process the risk of water supply shortages during sustained drought periods. Bruins (1993) indicated that "…governments often respond to drought through crisis management rather than pre-planned programs (i.e., risk management)." Wilhite (1993) also criticized that until recently, nations had devoted little effort toward drought planning, preferring instead the crisis management approach.

A consensus among water managers and researchers regarding water supply during drought is that the key to adequate management in urban areas lies in predrought preparation, especially as it relates to conservation and planning for future water needs (Dziegielewski et al., 1991). All the above accounts prompt us to focus on the necessity of risk-based design, especially when we deal with phenomena such as drought that are very difficult to predict accurately with regard to timing and magnitude.

6.2.5 Drought Severity as Risk Indices

Every natural phenomenon with associated detrimental effects to human beings and their environment need our keen attention to how and when it occurs.

Unfortunately, the degree of severity of some such phenomena including drought is difficult to determine, as accurately as desirable, before they occur. A study by the National Research Council (1986) indicated that there is not a firm rationale or explanation of the drought mechanism. It added that though empirical relations have been documented so far, why and when these relations trigger the occurrence of significant drought is not understood.

In the absence of such rationale, it is worth studying the degree (level) of risk, such as in the case of droughts, based on the available indices. The level of risk is apparently reflected by the severity of the drought. Severe drought implies a relative shortage of the required water supply, which in turn can be expressed by a certain level of the risk that the demand is not met. Thus calibrating drought severity may be used to indirectly determine the risk level.

No single definite method has been in use as a drought severity indicator. Nonetheless there are some methods being used in different fields. According to Wilhite (1993), the simplest drought index in widespread use is the percent of normal precipitation. This, indeed, is a good approach to infer the status of the available supply. However, it does not render an obvious forecast to enable a risk management body to be prepared for a forthcoming drought period.

Sheer (1980) tried to calculate the risk that the reservoir of a water supply system becomes empty by blending together the severest hydrologic and hydraulic conditions of different time periods. Specifically, he considered a condition in which the demands were the highest, the reservoir storage the lowest, and the date when these conditions occurred the beginning of one of the worst drought years. This simulation resulted in 4 out of 26 years in which the reservoir was empty, and it is concluded that the risk is $^4/_{26}$.

Although the approach is reasonable enough to indicate what would have happened had the conditions been met, the authors hardly believes that it fully reflects the realistic situation. One simple reason is that if the actual conditions were as bad as the ones selected, the demand could be higher and might result in more years of an empty reservoir, since the demands under such conditions would be much higher over the considered time span. Another reason may be that the risk in that study is not fully analogous with the usual convention. This is to say that the risk is based on the demand exceeding the supply, which is reached long before the reservoir becomes empty.

As the best alternative, risk analysis may be viewed in relation to the uncertainty associated with the different variables. Tung (1996) points out that the most complete and ideal description of uncertainty is the probability density function of the quantity subject to uncertainty. It is, therefore, very feasible to consider the probability density function of the demands about a fixed available supply during drought and thus derive the risk as the cumulative probability function of the supply being exceeded.

To be able to calculate the risk, the level of the drought severity must be determined (forecast). There are several drought severity indices, which have been used so far. Some of them are used to assess the severity of a drought event that has already happened, while a few others are used for forecasting. The Palmer Drought Severity Index (PDSI) and the Sheer Steila Drought Index (DI) are examples of the former category, while the Surface Water Supply Index (SWSI) and the Southern Oscillation Index (SOI) are examples of indices used for drought forecasting.

Palmer (1965) expressed the severity of a drought event by developing the following equation (Steila, 1972; Puckett, 1981):

$$\mathrm{PDSI}_i = 0.897\ \mathrm{PDSI}_{i-1} + \tfrac{1}{3} Z_i \tag{6.9}$$

where Z is an adjustment to soil moisture for carryover from one month to the next, expressed as

$$Z_i = k_j[PPT_i - (\alpha_j PE_i + \beta_j G_i + \gamma_j R_i - \delta_j L_i)] \tag{6.10}$$

where PPT_i = precipitation
PE_i = potential evapotranspiration
G_i = soil moisture recharge
R_i = surface runoff (excess precipitation)
L_i = soil moisture loss for month i.

Subscript j represents one of the calendar months and i is a particular month in a series of months. The coefficients α_j, β_j, γ_j, and δ_j are the ratios for long-term averages of actual to potential magnitudes for E, G, R, and L based on a standard 30-year climatic period.

The SWSI gives a forecast of a drought event. It is a weighted index that generally expresses the potential availability of the forthcoming season's water supply (U.S. Soil Conservation Service, 1988). It is formulated as a rescaled, weighted equation of nonexcedence probabilities of four hydrologic components: snowpack, precipitation, streamflow, and reservoir storage (Garen, 1993).

$$\frac{\mathrm{SWSI} = \alpha p_{\mathrm{snow}} + \beta p_{\mathrm{prec}} + \gamma p_{\mathrm{strm}} + \omega p_{\mathrm{resv}} - 50}{12} \tag{6.11}$$

where α, β, γ, and ω are weights for each hydrologic component and add up to unity; p_i is the probability of nonexcedence (in percent) for component i; and the subscripts snow, prec, strm, and resv stand for the snowpack, precipitation, streamflow, and reservoir storage hydrologic components, respectively. This index has a numerical value for a given basin, which varies between -4.17 to $+4.17$. The following are the ranges for the index for practical purposes: $+2$ and above, -2 to $+2$, -3 to -2, -4 to -3 and -4 and below. These ranges are associated with the qualitative expressions of abundant water supply, near normal, moderate drought, severe drought, and extreme drought conditions, respectively.

The SWSI has been in use to forecast monthly surface water supply forecasts of different basins (see, for example, *U.S. Soil Conservation Service Monthly Report,* 1988). In fact, it gives a forecast of both wet and dry (drought) months. On the other hand, Wilhite (1993) reports that several scientists agree that it has been possible to forecast drought for up to 6 months in Australia by using the SOI, which is based on forecast meteorological conditions.

6.3 RISK-PRICE RELATIONSHIP

Risk can be defined as the probability that the loading exceeds the resistance (Chow et al., 1988; Mays and Tung, 1992). Analogously, the risk in water distribution systems is defined as the probability that the demand exceeds the available supply where the demand is considered as the loading and the supply as the resistance. For future planning purposes, it is not certain when a drought event of a certain severity level will occur.

In planning for urban water supply projects, therefore, it is important to determine the probability distribution parameters of the demand and the supply. Both demand and supply are related to hydrologic indices. Also, operation and management of an existing water distribution system can be handled better through a risk analysis approach when the forthcoming period's (say month's) conditions of weather or water supply availability can be predicted ahead of time.

One of the common ways to represent uncertain events such as demand and supply under drought conditions is using an appropriate probability distribution of these variables. On the other hand, both variables are related to the return period T of the drought. The available supply data of many years can be arranged in descending order of magnitude for drought indication. These arranged flow data can be plotted versus the return period T, which is a measure of hydrologic conditions, as shown in Fig. 6.1. Two basic ways can be considered for selecting the representative flow data in relation to the return period. The first one is selecting one extreme value for each unit of time, e.g., the lowest monthly flows in a period of years. The second is selecting the lowest monthly flows in a period of years (Hudson and Hazen, 1964). Both of these procedures give a general relationship between available supply (flow) and its corresponding return period, as given in Fig. 6.1.

6.3.1 Developing Risk-Price Relationships

Demand depends on many uncertain factors and consequently is uncertain for a given return period drought event. The uncertainty can be represented through a probability distribution function as illustrated in Fig. 6.3, which indicates the risk at two different return periods.

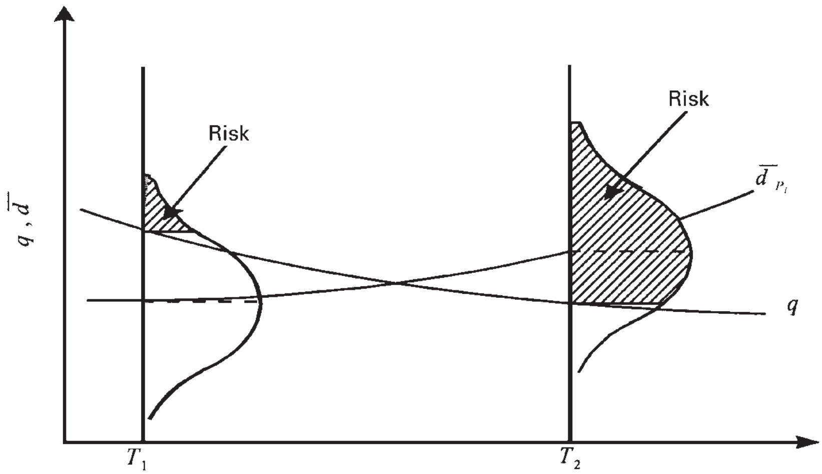


FIGURE 6.3 Probability distribution of demand at different return periods.

For decision purposes, the design may be fixed at the condition where the demand equals the available supply for a given price level. Beyond this point the demand exceeds the supply and there will be some associated risk. As shown in Fig. 6.2, the intersection points and the region beyond represent different values of water price and the associated risk. The illustrations in Figs. 6.1 and 6.2 show that for return periods larger than the critical return period T^* at the intersection point of supply and demand, the demand at the given price is greater than the supply. As the price decreases, the shortage volume increases thereby increasing the risk. Thus a graph of risk versus price may be plotted as shown in Fig. 6.4.

The regions beyond each of the intersection points in Fig. 6.2 have some corresponding risk levels, that is, the probability $p(\ )$ that the available supply falls below the demand corresponding to the price adopted at level i, $i = 1, 2, 3, \ldots, n$. Such a relation can help water supply planners to determine a municipal water supply price based on a predetermined tolerable risk or to assess the risk associated with a certain price level. Although not yet demonstrated by data analysis, the price-risk relationship indicates that price is infinite at no risk and risk is close to 1.00 at zero price (Fig. 6.4).

Risk Evaluation Procedures. The general procedures for evaluating the risk of a system's loading exceeding a system's capacity are considered under two different scenarios. For water supply systems, the demand may be considered as the loading and the supply as the capacity. The two scenarios are (1) when the loading is uncertain and the capacity is certain, and (2) when both the loading and the capacity are uncertain. Risk evaluation in the first case involves consideration of the probability density function of one variable (the demand) which is computa-

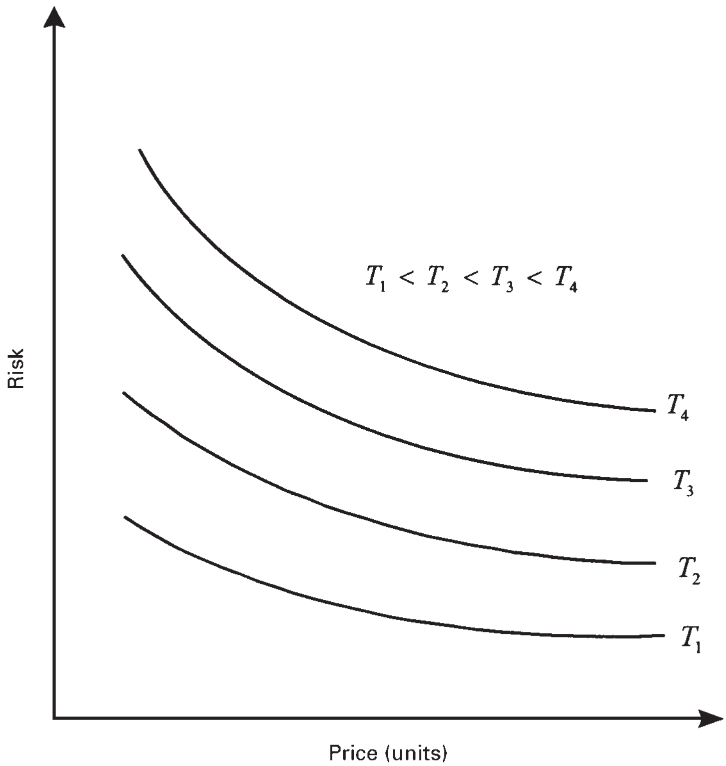


FIGURE 6.4 Risk-price relationships for different return periods.

tionally simpler. Risk evaluation in the second case involves composite risk evaluation.

Suppose that the probability density function of loading L is $f(L)$. The probability p that the loading will exceed a fixed and known capacity C^* is given as (Chow et al., 1988; Mays and Tung, 1992)

$$p(L > C^*) = \int_{C^*}^{\infty} f(L)\, dL \tag{6.12}$$

This relationship holds true when the capacity C is a deterministic quantity, which corresponds to the first scenario. Analogously, if the probability density function of demand d_{P_i} at price level P_i is $f(d_{P_i})$, the risk of demand d_{P_i} at price level P_i exceeding the supply q_T for return period T is expressed as

$$\text{Risk}_{(P_i,T)} = \int_{q_T}^{\infty} f(d_{P_i})\, dd \tag{6.13}$$

Using this definition for risk, the risk-price relationship may be developed for each *T*. The higher the price the lower the demand is, and consequently the lower the risk.

When the capacity is also uncertain but may be represented by a probability density function $g(C)$, i.e., the second scenario, the composite risk is used. The general formula for risk in this case is (Fig. 6.5).

$$\text{Risk} = \int_{-\infty}^{\infty} \left[\int_{C^*}^{\infty} f(L)\, dL \right] g(C)\, dC \tag{6.14}$$

Again in a similar analogy, the corresponding composite risk where both demand and supply are considered to be uncertain (for a given price and return period) is expressed as (Fig. 6.6)

$$\begin{aligned} \text{Risk}_{(P_i, T)} &= {}_p(\bar{d}_{P_i} > \bar{q}_T) \\ &= \int_{-\infty}^{\infty} \left[\int_{\bar{q}_T}^{\infty} f(q_T)\, dq \right] f(d_{P_i})\, dd \end{aligned} \tag{6.15}$$

A similar relationship as the one shown in Fig. 6.4 can also be developed for the composite risk from these relationships. In both cases [Eqs. (6.13) and (6.15)], the risk at a given price and return period is computed as the probability of the demand exceeding the supply. The difference is in the certainty of the supply in the former equation and its uncertainty in the latter one.

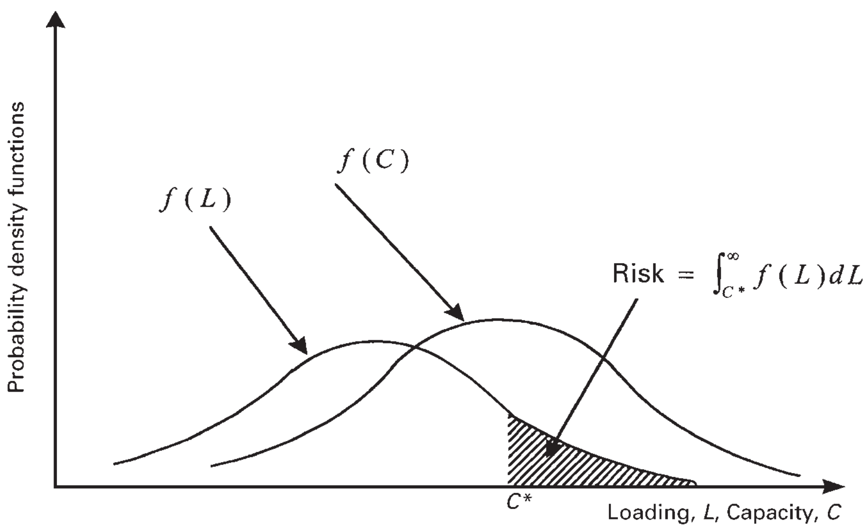


FIGURE 6.5 Probability distribution functions of loading and capacity.

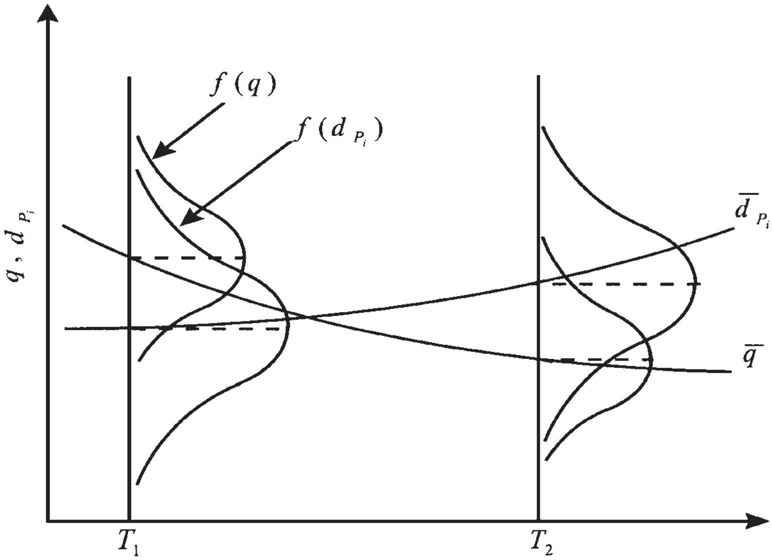


FIGURE 6.6 Probability distribution of both demand and supply at different return periods.

Methodology of Risk Evaluation. Numerical evaluations of risk using the above equations call for the approach to determine the quantitative values of different statistical parameters of the loading and/or the capacity. The risk equations consist of complex probability distribution functions, which become difficult to integrate. Because of this, alternative ways of evaluating the value of risk are often utilized. The safety margin and safety factor approaches (see Chow et al., 1988; Mays and Tung, 1992) are generally used for the computation of the risk from the probability distributions of the loading and/or the capacity. The safety margin approach is illustrated below with numerical data and the safety factor approach will be introduced in Sec. 6.4.5.

The safety margin SM is generally given as the difference between the loading and the capacity or SM $= C - L$. Thus the risk in terms of the safety margin is given as

$$\text{Risk} = p(C - L < 0) = p(\text{SM} < 0) \tag{6.16}$$

If C and L are independent random variables, the mean value and the standard deviation of SM are given, respectively, as

$$\mu_{\text{SM}} = \mu_C - \mu_L \tag{6.17}$$

$$\sigma_{\text{SM}}{}^2 = \sigma_C{}^2 + \sigma_L{}^2 \tag{6.18}$$

By taking water demand and the available supply as the loading and the capacity, respectively, the risks at different price levels for different return periods can be easily computed. Using the safety margin approach,

$$\text{Risk} = p[(d_{P_i} - q_T) < 0] = p(\text{SM} < 0) \tag{6.19}$$

$$\mu_{SM} = \mu_q - \mu_{d_P} \tag{6.20}$$

$$\sigma_{SM}{}^2 = \sigma_q{}^2 + \sigma_{d_P}{}^2 \tag{6.21}$$

where μ_q is the mean supply and μ_{dP} is the mean demand at price level P, respectively. Assuming that the safety margin is normally distributed, the risk is expressed as

$$\text{Risk} = p\left(\frac{\text{SM} - \mu_{\text{SM}}}{\sigma_{\text{SM}}} < \frac{0 - \mu_{\text{SM}}}{\sigma_{\text{SM}}}\right) = p\left(z < \frac{-\mu_{\text{SM}}}{\sigma_{\text{SM}}}\right) = p\left(z < \frac{-(\mu_q - \mu_d)}{(\sigma_q{}^2 + \sigma_d{}^2)^{1/2}}\right)$$

$$= \Phi_z\left(-\frac{\mu_{\text{SM}}}{\sigma_{\text{SM}}}\right) = \Phi_z\left(\frac{-\mu_q - \mu_d}{\sigma_q{}^2 + \sigma_d{}^2}\right)^{1/2} \tag{6.22}$$

where z is the standard normal variable with mean 0 and standard deviation 1.

However, before using these equations it is further required that the mean and the standard deviation estimates of demand and/or supply must be estimated. The expected value of demand at different price levels can be estimated using the price elasticity formula. Its standard deviation, on the other hand, can be estimated from the first-order analysis of uncertainty of the demand model (equation). If a dependent variable Y is a function of independent variables $\mathbf{X}$ ($\mathbf{X} = X_1, X_2, \ldots, X_k$) such that $Y = g(\mathbf{X})$, the first-order approximation of Y is given as

$$Y \approx g\,(\bar{x}) + \sum_{i=1}^{k} \left[\frac{\delta g}{\delta X_i}\right]_{\bar{x}} (X_i - \bar{x}_i) \tag{6.23}$$

in which $\bar{\mathbf{x}} = (\bar{\mathbf{x}}_1, \bar{\mathbf{x}}_2, \ldots, \bar{x}_k)$, a vector containing the means of k random variables (Mays and Tung, 1992). The variance of Y, Var[Y] or $\sigma_Y{}^2$, is estimated by the following equation, which can be derived from the first-order analysis of uncertainty of Eq. (6.23).

$$\sigma_Y{}^2 = \text{Var}\,[Y] \approx \sum_{i=1}^{k} \alpha_i^2 \sigma_i^2 + 2 \sum_{1<}^{k} \sum_{j}^{k} a_i a_j \,\text{Cov}\,[X_i, X_j] \tag{6.24}$$

where $a_i = [\delta g/\delta X_i]_{\bar{x}}$ and σ_i^2 is the variance corresponding to random variable X_i. When the X_i's are independent random variables, Cov $[X_i, X_j] = 0$.

The foregoing discussion in general indicated that for a given return period for design, water supply planners can decide the price of the water supply for an affordable risk level or can determine the risk at a given affordable water price. The flowchart given in Fig. 6.7 summarizes the basic steps used to develop the risk-price-return period relationships.

Risk Evaluation Example. Based on the safety margin analysis given by Eqs. (6.19) to (6.22) and the price elasticity of demand definition [Eq. (6.1)], it is pos-

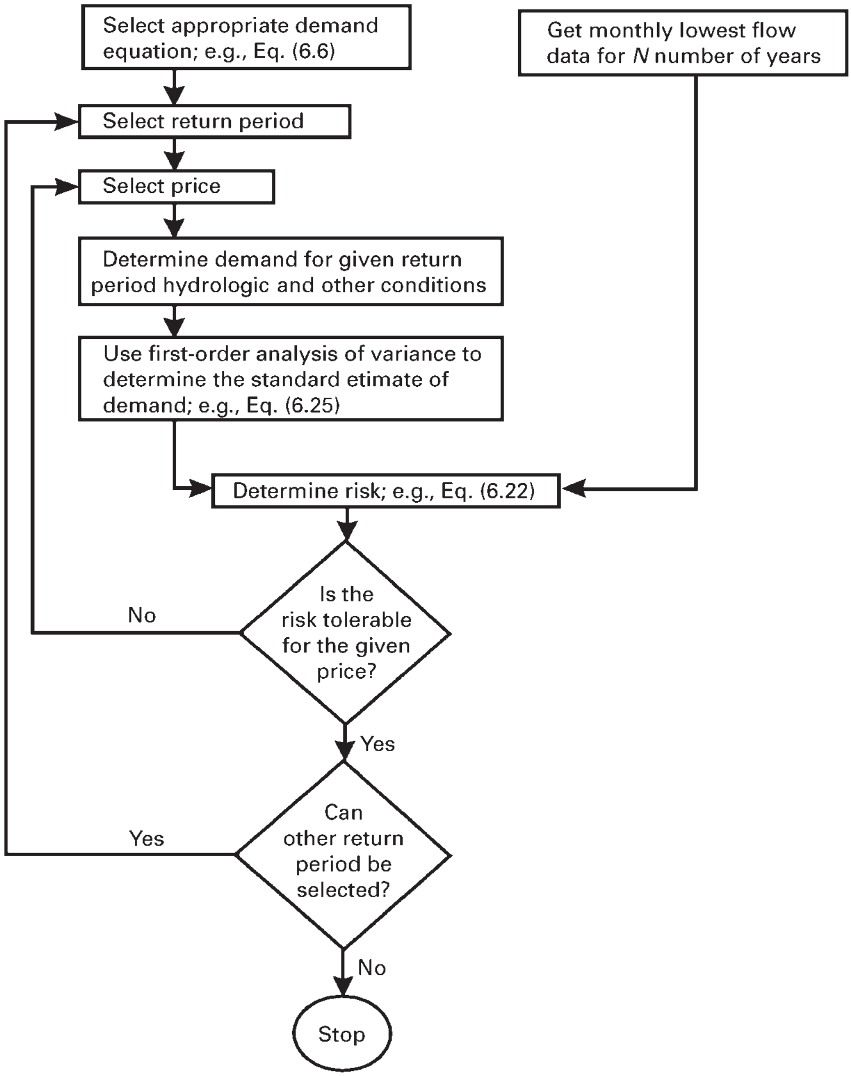


FIGURE 6.7 Flowchart for the proposed planning procedure.

sible to determine the risk values for a given return period and different price levels. For a given return period of drought, the expected demand when the price is increased by a certain amount can be determined by Eq. (6.25). Table 6.3 lists demand for an initial price level P_i and also the available supply for different return periods. Equation (6.1) for the price elasticity of demand is rearranged to solve for d_i as

$$d_i = \frac{d_{i-1}\,[1 + \eta_P(P_i - P_{i-1}) / (P_i + P_{i-1})]}{1 - \eta_P(P_i - P_{i-1}) / (P_i + P_{i-1})} \tag{6.25}$$

Equation (6.25) is used to determine the demand at a given price level and a given return period. Price increases of up to 200 percent and a price elasticity of −0.5 are used to compute the demand reduction due to the increases in the price for each of the return periods. The risks associated with different price levels and different return periods are determined based on approximate estimates of the standard error for supply as 2.0 units and for demand as 4.0 units, for which σ_{SM} equals 4.47. The results thus obtained are given in Table 6.4 and also plotted as shown in Fig. 6.8. The plots show that the risk is not significantly sensitive to the price change for small price increases. The plot in Fig. 6.9 shows how sensitive risk is to the return period *T*. It is inferred from these two plots that planning and/or overcoming the shortage of water supply during drought periods requires a strong commitment to increase the price sufficiently.

6.4 OPERATION AND MANAGEMENT PLANNING UNDER SUSTAINED DROUGHT CONDITIONS

The price-elasticity formulation indicates that when the available supply is less than the demand, the latter can be adjusted to the former by increasing the price. In other words, for a drought event of severity index greater than the one at which the demand equals the available supply (Fig. 6.10), it is possible to force the demand curve down to the supply curve by increasing the price. However, the fact

TABLE 6.3 (Hypothetical) Data of Demand and Supply for Different Return Periods

Return period *T*, years	Demand at price level P_1, units	Available flow *q*, units
1	8.0	12.0
5	8.5	11.0
10	9.5	9.5
25	11.0	8.0
50	13.0	7.0

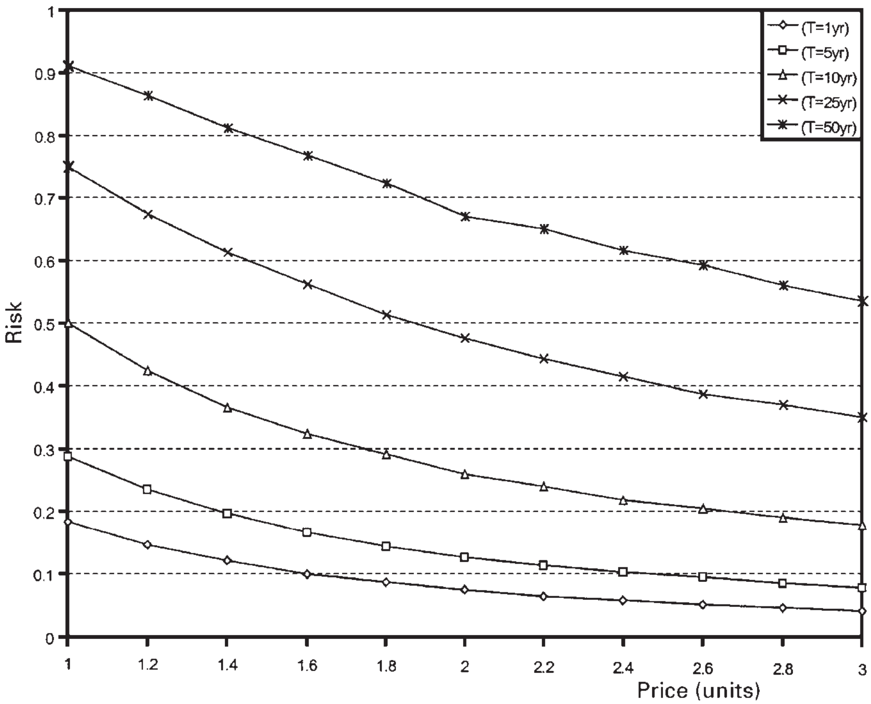


FIGURE 6.8 Risk-price-return period relationships.

that it has not been easy to forecast drought conditions well ahead of time and the uncertainty in its magnitude and length requires operation and management of water supply systems that will attempt to smooth out the effect of the drought. Such operation and management efforts will be based on data of a short time interval. The efforts in effect are a supplement to the planning procedure already mentioned above. The planning basically turns out to be a one-time decision, while operation and management especially under sustained drought conditions involve routine decisions. The severity of the drought could be so high that emergency water supply construction projects may be considered. An estimate of the damage that would result from a sustained drought period is required and, most of all, may be used to determine if an emergency water supply should be implemented.

6.4.1 Economic Aspects of Water Shortage

The shortage of water supply during drought periods results in different types of losses in the economy including, but not limited to, agricultural, commercial, and industrial. In agriculture, lack of water supply results in crop failures; in commerce, it may result in a recession of the business; and in industry, it may result in

TABLE 6.4 Risk Values for Different Return Periods and Price Increases of up to 200%

	Return period, *T*, years				
Price (unit)	1	5	10	25	50
1.0	0.183	0.288	0.500	0.749	0.910
1.2	0.147	0.236	0.425	0.674	0.864
1.4	0.123	0.198	0.367	0.614	0.813
1.6	0.102	0.169	0.326	0.564	0.770
1.8	0.090	0.147	0.295	0.516	0.726
2.0	0.079	0.131	0.264	0.480	0.674
2.2	0.069	0.119	0.245	0.448	0.655
2.4	0.064	0.109	0.224	0.421	0.622
2.6	0.058	0.102	0.212	0.394	0.599
2.8	0.054	0.093	0.198	0.378	0.568
3.0	0.050	0.087	0.187	0.359	0.544

underproduction of commodities. The loss in each production or service sector depends on the purpose of the sector. For instance, the economic impact of drought on agriculture depends on the crop type (Easterling, 1993). There is no single common way of assessing the economic impact of drought on any one of the sectors. Evaluating and comparing what actually happens during a drought period with what would have happened had there been no drought may be one way of assessing the effects of drought (Dixon et al., 1996). Dixon et al. (1996) adopted the concept of willingness-to-pay to value changes in well-being. They define *willingness-to-pay* as the maximum individuals would have been willing to pay to avoid the drought management strategies imposed by water agencies.

On the other hand, since water is supplied during a drought period at a greater price, it can be viewed as a revenue generator. Therefore, when the demand exceeds the available supply, the revenue collected by the water supply agency will be less than what could have been collected had there been more supply than that actually available. In other words, if the demand exceeds the supply, the problem is not only limited to lack of water but there will also be economic loss since the customers would pay for more supply if there were enough. Depending on the risk level, it is possible to reach a decision of whether supply augmentation is necessary or the pressure for more demand could be tolerated with the available supply.

Some water shortage relief efforts can be undertaken so that emergency water supplies may be made available to the users. This can be implemented by well drilling, trucking in potable supplies, or transporting water through small-diame-

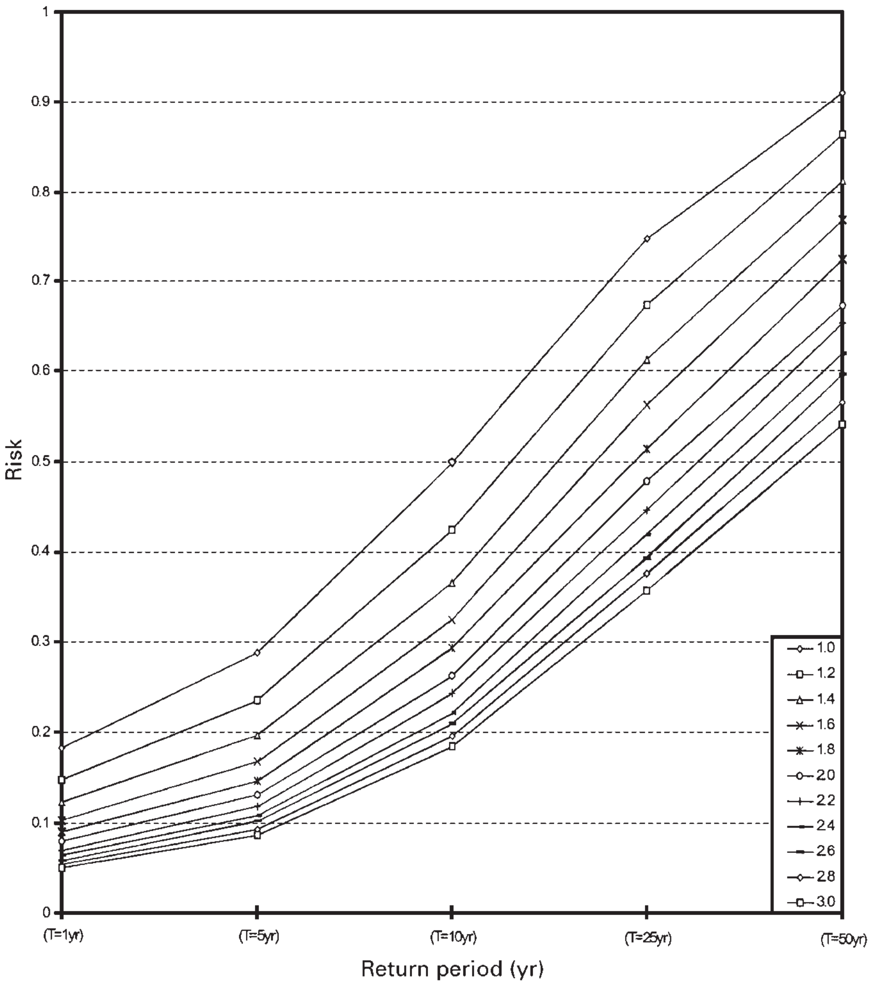


FIGURE 6.9 Risk-price-return period relationships. The legend indicates the price in units.

ter emergency water lines. In such cases, it may be required that the emergency supply construction costs be paid by the users (Dziegielewski et al., 1991). The estimation of the expected financial loss can be used to determine and inform the users of its extent and advise them of the necessity, if any, of paying for the emergency supply construction costs.

If the option for emergency supply construction is justified, then the design needs to take into consideration the possibilities of optimization. The construction can be designed such that the financial risk and the cost of construction are at optimum. Figure 6.11 illustrates this optimization process.

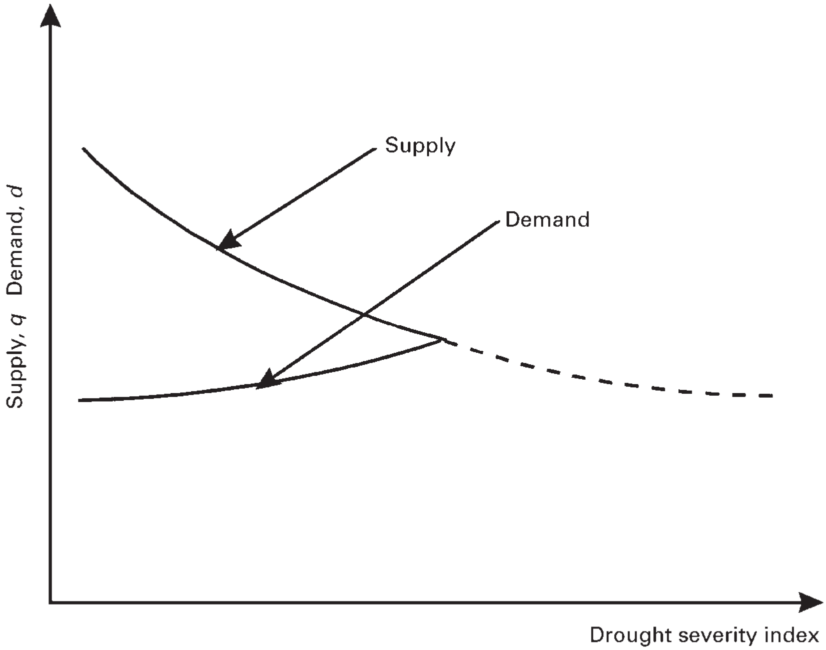


FIGURE 6.10 Water demand during a sustained drought period as adjusted to the available supply (the broken line shows the adjustment).

The economic loss (damage) can be calculated with the help of Eqs. (6.29) and (6.30) (given in Sec. 6.4.2), and the cost of emergency construction must be determined from the physical conditions at the disposal of the water supply agency.

6.4.2 Damage Assessment

The damage that would result if a certain drought event occurred is used as one of the main decision factors. Since the time of occurrence of the drought event that causes the damage is difficult to determine, only the expected value is assessed by associating its magnitude with its probability of occurrence.

Chow et al. (1988) define the expected annual damage cost D_T for the event $x > x_T$ as

$$D_T = \int_{x_T}^{\infty} D(x)\, f(x)\, dx \tag{6.26}$$

where $f(x)\, dx$ is the probability that an event of magnitude x will occur in any given year and $D(x)$ is the damage cost that would result from that event. The event x in this case can be assumed as the demand and x_T can be the available supply during a drought event of return period T.

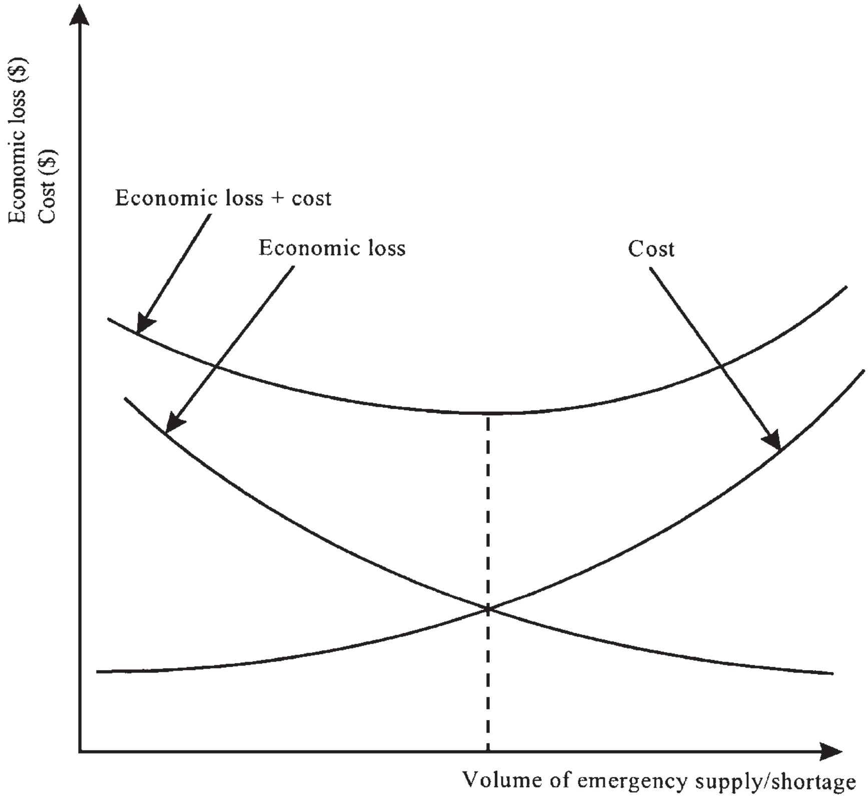


FIGURE 6.11 Optimization for emergency water supply construction.

Breaking down the expected damage cost into intervals,

$$\Delta D_i = \int_{x_{i-1}}^{x_i} D(x)\, f(x)\, dx \tag{6.27}$$

from which the finite difference approximation is obtained as

$$\begin{aligned} \Delta D_i &= \frac{D(x_{i-1}) + D(x_i)}{2} \int_{x_{i-1}}^{x_i} f(x)\, dx \\ &= \frac{D(x_{i-1}) + D(x_i)}{2} [p(x \geq x_{i-1}) - p(x \geq x_i)] \end{aligned} \tag{6.28}$$

Thus the annual damage cost for a structure designed for a return period T is given as

$$D_T = \sum_{i=1}^{\infty} \frac{D(x_{i-1}) + D(x_i)}{2} \left[p(x \geq x_{i-1}) - p(x \geq x_i)\right] \tag{6.29}$$

To determine the annual expected damage in Eq. (6.29), the damage that results from drought events of different severity levels must be quantified.

The magnitude of the drought (in monetary units) may be obtained by estimating the volume of water shortage that would result from that drought. In other words, not having the water results in some financial loss to the water supply customer. The resulting financial loss to the customer from a certain drought event is thus considered as the damage from that drought event.

As was shown in Fig. 6.1, after the critical return period T^*, the divergence between the demand and the supply increases with the return period. Expressing the demand and the supply as a function of return period T of drought events enables one to estimate the annual expected water supply shortage volume as

$$S_V = \int_{T^*}^{T} [d(T) - q(T)]\, dt \tag{6.30}$$

The shortage volume S_V is illustrated by the shaded area in Fig. 6.12. The shortage volume for a drought event of a higher return period than the critical one results in a higher shortage volume and consequently a higher associated damage. The relationship between the shortage volume and the associated damage generally depends on several factors including the water-use category—residential, industrial, commercial, agricultural, and so on. To use the procedure presented herein for assessing the damage that results from a certain water shortage volume, the damage given by Eq. (6.29) must be developed for a specific user category.

6.4.3 Operation and Management

For operation and management of an existing municipal water supply system during a sustained drought period, administrative decisions may be based on short-time forecasting of the hydrologic conditions. A forecast of, say 1 month ahead of the available supply, helps the supply managers to preadjust the expected demand to the forecasted available supply by increasing the price. In other words, the expected demand can be, in principle, suppressed to the forecasted available supply by increasing the price. Howe (1993) points out that since price presumably affects the quantities users demand, price can be used to adjust demand to the available supply. The basic factor in the decision will be the damage that would occur if the adjustment were not undertaken. This is the reason why we need to focus on the assessment of such damages.

It may be easily conceived from the above reasoning that it is possible to express price increase as some function of damage. If dP is an elementary increase

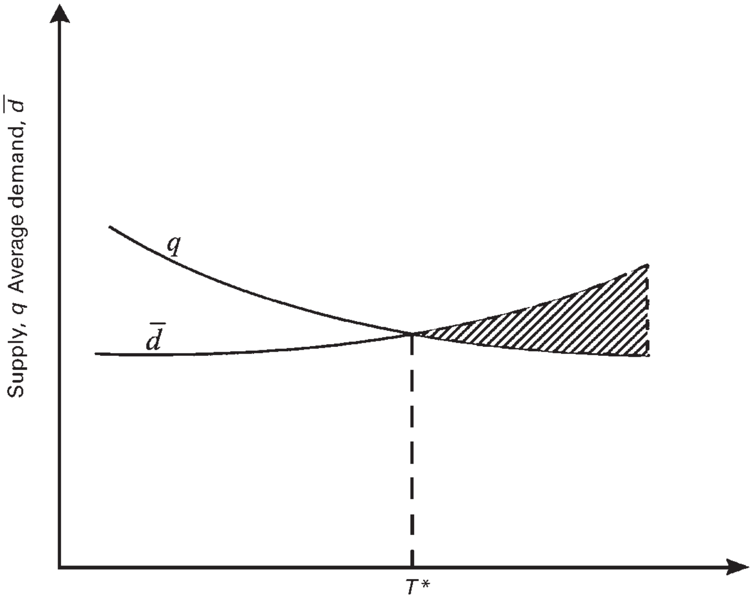


FIGURE 6.12 Demand and supply showing water shortage volume when demand exceeds supply.

in price due to a certain level of drought, the following general relationship may be formulated:

$$dP = \phi\,(\xi) \tag{6.31}$$

where ξ is an implicit variable for the drought severity level.

The amount of decrease in the demand attained as a result of the increase in the price may be determined from the concept of price elasticity of demand for water, which is rewritten in finite difference form as

$$\Delta d = \frac{\eta_P(d_{x_{i-1}} + d_{x_i})/2}{(P_{x_{i-1}} + P_{x_i})/2}\,\Delta P \tag{6.32}$$

The equation for the increased price P_{x_i} is obtained from Eq. (6.32) as

$$P_{x_i} = \frac{P_{x_{i-1}} + P_{x_{i-1}}(d_{x_i} - d_{x_{i-1}})/\eta_P(d_{x_{i-1}} + d_{x_i})}{1 - (d_{x_i} - d_{x_{i-1}})/\eta_P(d_{x_{i-1}} + d_{x_i})} \tag{6.33}$$

Also Eq. (6.31) can be written in finite difference form as

$$\Delta P = P_{x_i} - P_{x_{i-1}} = \phi(\xi) \tag{6.34}$$

$$P_{x_i} = P_{x_{i-1}} + \phi(\xi) \tag{6.35}$$

A close-up look at Eq. (6.33) indicates that the price P_{x_i} at drought event level x_i is greater than the price $P_{x_{i-1}}$ at drought event level x_{i-1}, as expected. To achieve this, the price must increase from $P_{x_{i-1}}$ to P_{x_i} by the amount $\phi(\xi)$, as shown by Eq. (6.35). Thus by increasing the price, the supply deficiency of water during sustained drought periods can be overcome or minimized. In fact, the price can be forced to rise to the level that limits the demand of water to that amount which is available. Doing so will theoretically enable us to adjust the portion of the demand curve beyond the critical drought severity index (Fig. 6.11) down to the supply curve. However, this may not be readily accepted by the customers and thus arises the uncertainty. In essence, there will result a positively skewed distribution tendency of the customers for the demand, and hence the analysis of the associated uncertainty comes into picture.

6.4.4 The $\phi(\xi)$ Function

To fully make use of Eq. (6.33) or (6.35) for water demand abatement through price increase, an explicit form of the drought function, $\phi(\xi)$, in which ξ is the drought severity index must be determined. Different approaches have been followed to develop indices for a drought event. Presuming that the SWSI is one of the alternatives available to forecast a drought severity level, ϕ(SWSI) will be used herein. The subscripts of d and P may be substituted by the numerical values of SWSI. For instance, if a drought month of SWSI $= -2.00$ is forecast to follow a normal month of SWSI $= 0.00$, the price $P_{-2.00}$ can be determined based on Eq. (6.33), with the price during the normal month $P_{0.00}$ known. Since the SWSI depends on p_{snow}, p_{prec}, p_{strm}, p_{resv}, the following general relation between SWSI and the variables may be conceived.

$$\text{SWSI} = \psi(p_{\text{snow}}, p_{\text{prec}}, p_{\text{strm}}, p_{\text{resv}}) \tag{6.36}$$

Apparently, then,

$$\phi(\xi) \equiv \phi(\text{SWSI}) = \phi'(p_{\text{snow}}, p_{\text{prec}}, p_{\text{strm}}, p_{\text{resv}}) \tag{6.37}$$

Once a fully explicit model is developed for Eq. (6.37), it becomes possible to recompute the expected demand using Eq. (6.25). As an alternative for this, the following equation may also be derived from Eq. (6.32):

$$P_{x_i} = P_{x_{i-1}} + \frac{(d_{x_i} - d_{x_{i-1}})\,(P_{x_i} + P_{x_{i-1}})/2}{\eta_P(d_{x_i} + d_{x_{i-1}})/2} \tag{6.38}$$

The second term on the right-hand side in Eq. (6.38) is equivalent to the $\phi(\xi)$ function mentioned earlier. It is to be noted that Eq. (6.33) gives an explicit equation to determine the price at drought event level x_i, while Eq. (6.38) gives a term equivalent to the $\phi(\xi)$. The SWSI can be used to indicate if a drought may occur and to determine its severity level if it occurs. It is to be recalled that, as indicated by Eq. (6.33), the demand at drought event level x_i can be adjusted to the estimated available supply q_{x_i} by increasing the price from its value at drought event level x_{i-1} to a new value at drought event level x_i.

6.4.5 Uncertainty and Risk in Demand

Although it is presumed that demands can be adjusted to the available supply, there is uncertainty. Demand is a variable and may not meet the available supply irrespective of the increase in the price. Hobbs (1989) points out that future demands are random because they depend upon weather, consumer tastes and preferences, household income, water rates, and level of economic development. These reasons naturally cause the demand to have some positively skewed probabilistic distribution.

Some organizations and researchers have used different probability distributions for demand. Charles Howard and Associates (in 1984) and Norrie (in 1983) used a gamma distribution for demand for Seattle (Hobbs, 1989). Also, it may be possible that the statistics of the distribution of the demand about the available supply is not uniform at different drought severity levels.

A general trend of the supply with the drought severity index and the distribution of the demand about the supply may be represented as shown in Fig. 6.13. A general gamma probability density function given by the following equation (Montgomery and Runger, 1994) is assumed.

$$f_x(x; \lambda, r) = \frac{\lambda^r x^{r-1} e^{-\lambda x}}{\Gamma(r)} \qquad x > 0 \qquad \lambda > 0, \qquad \text{and} \qquad r > 0 \tag{6.39}$$

where

$$\Gamma(r) = \int_0^\infty x^{r-1} e^{-x}\, dx \qquad r > 0 \tag{6.40}$$

Taking the SWSI as the drought severity level indicator and the demand as the variable x in the gamma function given in Eq. (6.40), a general relationship between the supply q, the SWSI, and the density function of the demand $f(d)$ can be given as illustrated in Fig. 6.13. Negatives of the SWSI values normally adopted are used in this figure to indicate the increase of severity with the index in absolute terms.

Let d_{x_i} be the random demand at drought severity level x_i, and by implication let q_{x_i} be the corresponding known available supply at drought severity level x_i. Tung (1996) defines *reliability* as the probability that the resistance is greater than the loading. In a similar analogy, the reliability of a water supply system may be defined as the probability that the available supply is greater than the expected

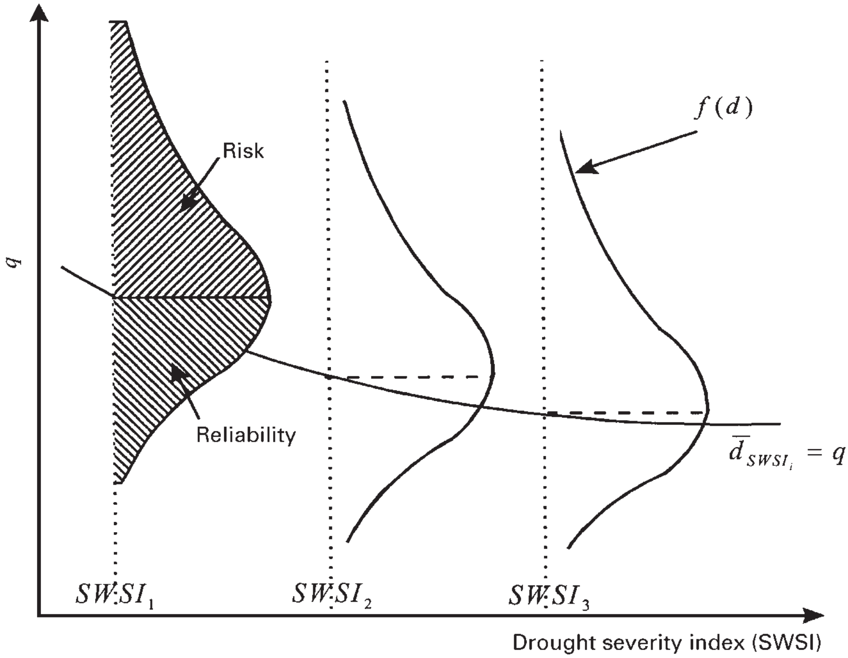


FIGURE 6.13 Expected water demand as adjusted to the available supply under sustained drought conditions and its probability distribution.

demand. Thus, the reliability of supply is the probability that q_{x_i} is greater than d_{x_i}, expressed as

$$R = p(q_{x_i} > d_{x_i}) \tag{6.41}$$

and the risk is

$$\text{Risk} = 1 - p(q_{x_i} > d_{x_i}) = p(d_{x_i} > q_{x_i}) \tag{6.42}$$

where R is the reliability that the available supply is greater than the estimated demand and p is the probability. As illustrated in Fig. 6.13, a higher demand above the available supply implies a higher risk and a lower reliability. The risk defined by Eq. (6.42) can also be expressed in terms of the *safety factor,* SF, which may be defined as

$$\text{SF} = \frac{q_{x_i}}{d_{x_i}} \tag{6.43}$$

where the corresponding risk formula is

$$\text{Risk} = p(\text{SF} < 1) \tag{6.44}$$

For different drought severity indices, different risk–safety factor relationships can be developed. This is illustrated in Figs. 6.14 and 6.15. Once such relationships are developed, it is easier for water supply managers to decide the tolerable risk for a given drought severity index. It is to be noted here that the reliability analysis is just complementary to the risk analysis, whereas the safety factor approach is simply an alternative to the safety margin approach discussed in Methodology of Risk Evaluation in Sec. 6.3.1.

6.4.6 Operation and Management Strategy

It is indicated in the foregoing sections that urban water supply operation and management during sustained drought periods requires preparation at least by the water supply agents. Sound preparation procedures entail good strategy to be used. Most of all, collection of enough data affecting the water supply during a forthcoming period of time enables the supply agents to be prepared better for smoothing out the effect from a forecast drought event. Such efforts must be undertaken continuously during a sustained drought period. The flowchart shown in Fig. 6.16 will help water supply operation and management during drought periods. As indicated in the flowchart, the important data to forecast are the forthcoming period's (say month's) weather conditions or available flow and the expected demand. A comparison between the expected demand (after some necessary price increase) and the available flow clearly indicates if there will be a

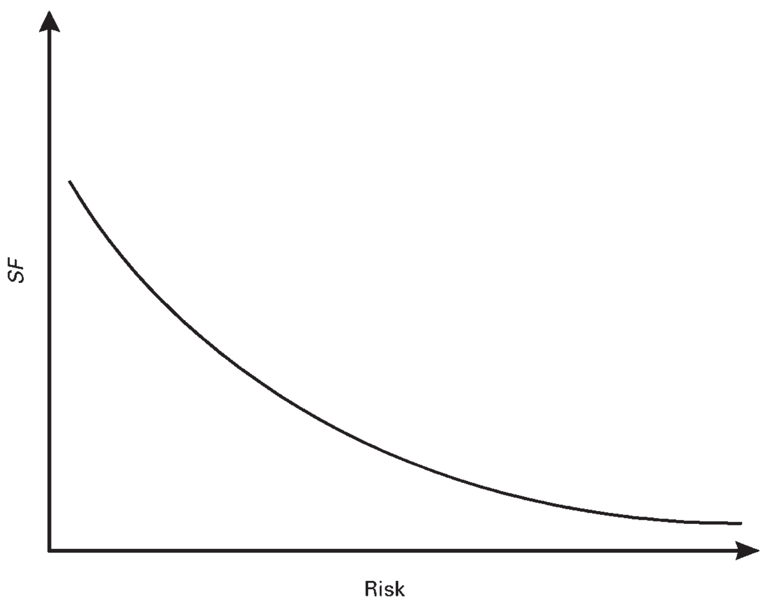


FIGURE 6.14 General illustration of risk–safety factor relationship.

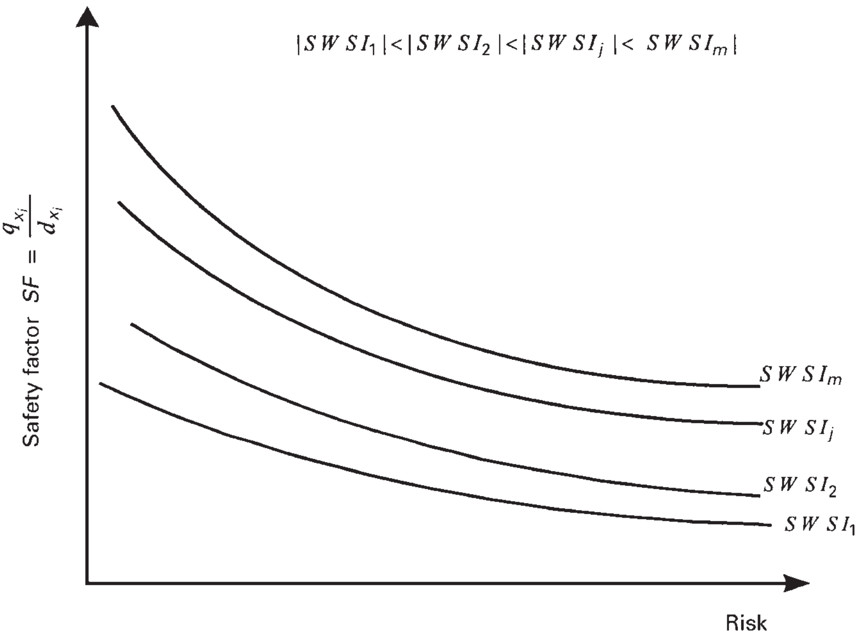


FIGURE 6.15 Risk–safety factor relationship for different SWSIs.

drought event. If the expected demand is found to be greater than the available flow, it implies that a drought event exists and hence necessary measure(s) should be sought.

6.4.7 Risk Evaluation Procedure under Sustained Drought Conditions

Developing the Procedure. The flowchart given in Fig. 6.16 for evaluating the risk under sustained drought conditions entails an explicit method of risk assessment. A certain gamma distribution must be determined among a possibly infinite number of such distributions. Specifically, the parameters of the distribution, λ and r, must be selected. The chi-square distribution, perhaps the most widely used gamma distribution, is selected for this example. The value of λ for chi-square distribution equals 1/2, whereas an r value of 1 is selected from possible values of ½, 1, 1½, …

Figure 6.17 shows the relative values of the available supply q_{x_i} and the expected demand d_{x_i} which are plotted on a chi-square distribution with $\lambda = 1/2$ and $r = 1$. The shaded area shows the probability that the demand exceeds the supply, or equivalently, the probability that the safety factor SF is less than 1. The area under the chi-square distribution that lies to the right of the d_{x_i} limit (or the $\chi_{p/2,\, n-1}{}^2$

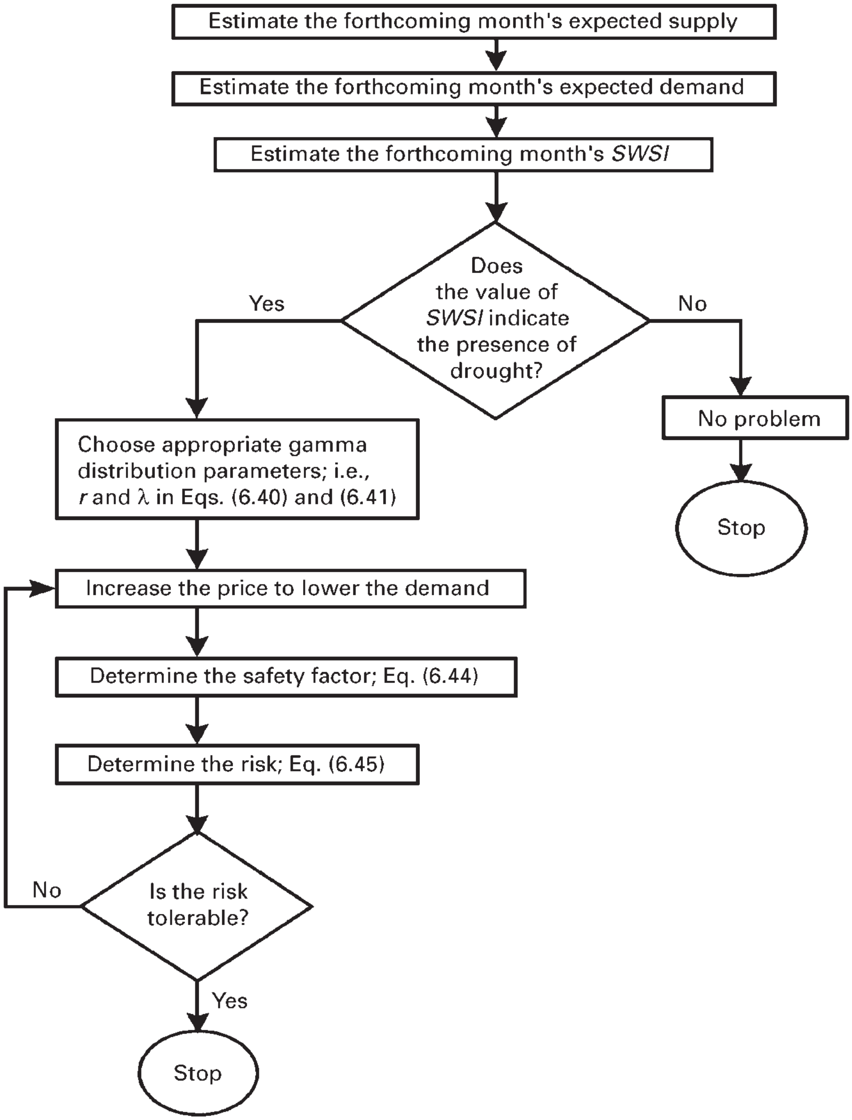


FIGURE 6.16 Flowchart for adjusting demand to available supply under sustained drought conditions.

limit) is simply $p/2$, which in essence represents a probability value. Also the area that lies to the left of the $\chi^2_{1-p/2,\,n-1}$ limit is $p/2$. The letter n represents the sample size used in determining the expected average demand, and the value $(n - 1)$ represents the degrees of freedom for the error in estimating the expected demand. Therefore, the risk which is represented in the figure by the shared area is one-half of $1 - p$, i.e., Risk $= \frac{1}{2}(1 - p)$. Equivalently, the risk can be derived from Eq. (6.43) as

$$\text{Risk} = p\left(\frac{q_{x_i}}{d_{x_i}} < 1\right) = p\left(\frac{d_{x_i}}{q_{x_i}} > 1\right)$$
$$= p(d_{x_i} - q_{x_i} > 0) \tag{6.45}$$

Since it is assumed that the expected demand is distributed as chi-square and the available supply is considered deterministic, the quantity $(d_{x_i} - q_{x_i})$ is also distributed as chi-square. As shown in Fig. 6.17, the $\chi_{p/2,\,n-1}{}^2$ value where $d_{x_i} = q_{x_i}$ is $\chi_{0.5,\,n-1}{}^2$. Therefore, the risk can be computed as

$$\text{Risk} = p(d_{x_i} - q_{x_i} > \chi_{0.5,\,n-1}{}^2) \tag{6.46}$$

Figure 6.18 gives a simplification procedure that will help determine the risk value for a given demand level. Figure 6.18*a* represents the actual demand level on the chi-square distribution. However, to use some of the readily available probability values of the chi-square distribution at varying levels of the given variable, the distribution must start from the origin (Fig. 6.18*c*). Subtracting q_x from d_x displaces the actual demand distribution to the left by q_x, as shown in Fig. 6.18*b*. Adding $\chi_{0.5,\,n-1}{}^2$ to $d_x - q_x$ results in the desired final distribution. Referring to this figure and simplifying Eq. (6.46) results in

$$\text{Risk} = p(d_{x_i} - q_{x_i} + \chi_{0.5,\,n-1}{}^2) - p(\chi_{0.5,\,n-1}{}^2) \tag{6.47}$$

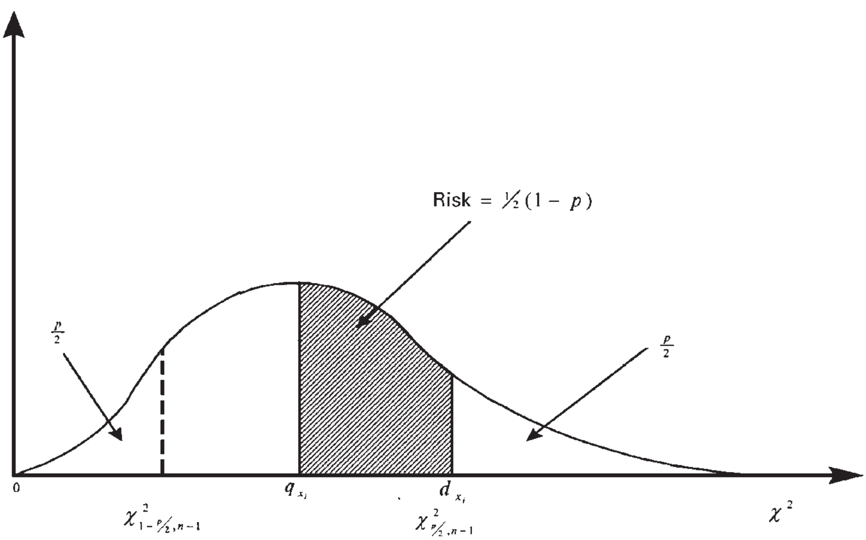


FIGURE 6.17 Available supply and expected demand as represented on a chi-square distribution.

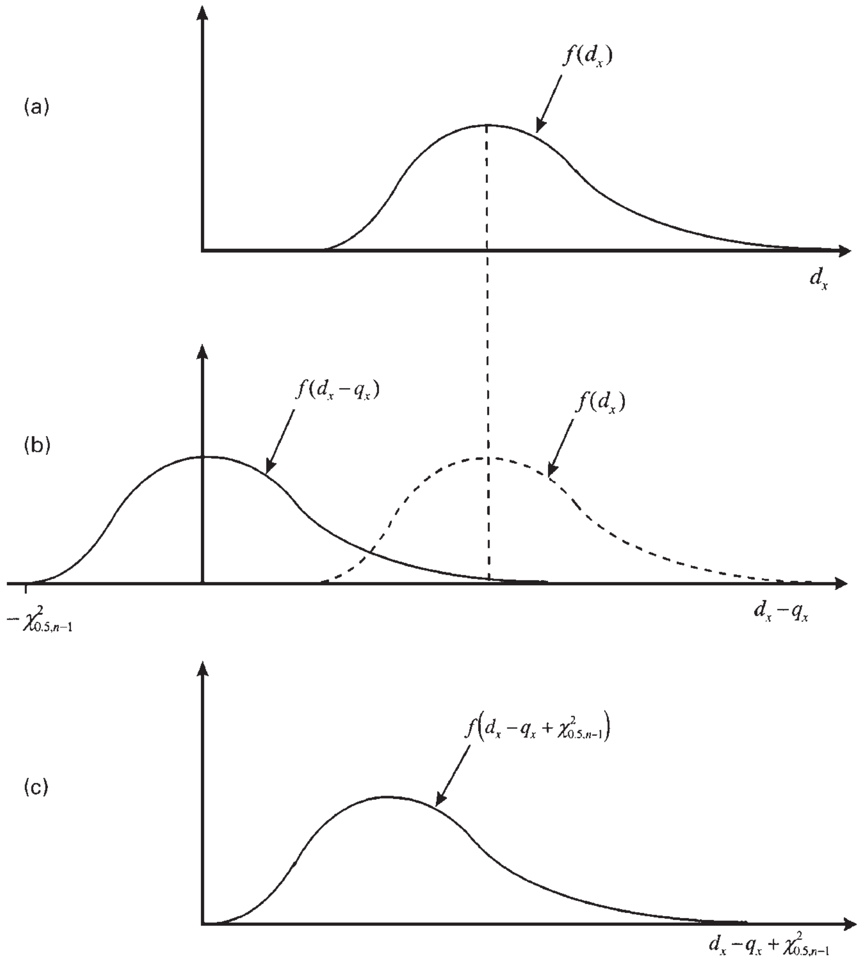


FIGURE 6.18 Representation of demand by the chi-square distribution: (*a*) general representation, (*b*) displaced distribution so that the mean demand is put at the origin, and (*c*) demand distribution in (*b*) displaced to the right by $\chi_{0.5,n-1}{}^2$.

where $p(\chi_{0.5,\,n-1}{}^2) = 0.5$. Therefore,

$$\text{Risk} = p(d_{x_i} - q_{x_i} + \chi_{0.5,\,n-1}{}^2) - 0.5 \tag{6.48}$$

It may be noted here that the maximum value of risk given by Eq. (6.48) is 0.5. This is in agreement with the definition of the safety factor and the procedure developed in this work. Referring back to Fig. 6.10 shows that under sustained drought conditions, the primary objective is to lower the demand down to the available supply, lowering it to a level below the available supply not being the primary focus. However, should the latter be considered, a little modification can be applied to Eq. (6.48) simply by dropping the constant 0.5 for $\chi_{0.5,\,n-1}{}^2$.

Risk Evaluation Example Using Data from the City of Phoenix. To show numerical evaluation of risk, data was obtained from the city of Phoenix (Table 6.5). As indicated in the table, there is no uniform water price for the city. The price arrangement is divided into two categories: inside city and outside city uses. The costs for both categories are further divided into three seasons: low months (December, January, February, March); medium months (April, May, October, November); and high months (June, July, August, September) (Kiefer, 1994). In view of the procedure developed in this work that is sought for drought conditions, the high months' current price of \$1.53/100 ft^3 and an applicable environmental charge of \$0.08/100 ft^3 are used. In addition, the service charge[1] of \$5.16 for inside city use for a meter size of 5/8-in is selected. It is to be noted that if a customer having a 5/8-in meter size doesn't use more than the 600-ft^3 limit during the months of October through May or the 1000-ft^3 limit during the months of June through September, she or he pays only the monthly service charge. In general, such a customer pays \$5.16 plus \$1.61 for every 100 ft^3 above the specified limit during a given month. Table 6.5 shows the general data from which the above information was extracted.

Based on the data for water price given in Table 6.5, the data given in Table 6.6, an assumed price elasticity of -0.267 [a value assessed for Tucson (Agthe and Billings, 1980)] and an assumed certain dry month (say SWSI $= 2$), the risk-price relationship is determined as given in Table 6.7. Although the data in Table 6.6 show that, in June 1997, the demand was below the available supply, a demand as high as 736.9 gallons per day (gpd) was recorded in July 1989 (Kiefer, 1994). To show the application of the foregoing procedure, an expected demand of 700 gpd is assumed as the typical demand during a given month of the selected drought period. The price elasticity concept [Eq. (6.25)] is used to lower the demand first, and then the risk at each price level is computed. With the assumption of a sample size of 20, the risk value is determined for each safety factor computed.

The risk–safety factor result obtained (Table 6.7) is drawn as shown in Fig. 6.19. It may be noted that to bring the high demand down to the available supply, it requires several price increase steps. Thus only part of the risk–safety factor data are plotted.

6.5 SUMMARY AND CONCLUSIONS

The damages caused due to lack of urban water supply during drought conditions may result in adverse and undesirable effects to the general welfare. The efforts

[1] The monthly service charge is for 600 ft^3 for the months of October through May inclusive and 1000 ft^3 for the months of June throught September inclusive.

TABLE 6.5 City of Phoenix Water Rates (Effective March 1, 1997)

Part I. Monthly service charge		
Meter size, in	Inside city, $	Outside city, $
5/8	5.16	7.74
1	5.61	8.42
$1^1/2$	8.88	13.32
2	9.78	14.67
3	39.06	58.59
4	47.24	70.86
6	51.33	77.00
Part II. Volume charges (above the limits)		
Months	Inside city, $	Outside city, $
Low months	1.01	1.52
Medium months	1.19	1.79
High months	1.53	2.30
Environmental charge	0.08	0.12

TABLE 6.6 Inside City Data for the City of Phoenix for Low-Density Residential Water Use for the Month of June 1997

Code	Conservation, million gal (A)	% of total* (B)	No. of accts (C)	Avg. conservation, gpd (ft^3/d) (A/30)C (D)	Total production, million gal (E)	Supply share [B(E)/100] (F)	Avg. supply available, gpd (ft^3/d) (F/30C) (G)
0†	2435	24	132,584	612 (82)	11,247	2672	672 (90)

*Total water use during the same month equals 10,251,693,804 gal.
†This code is for single-family residential water use.
Source: City of Phoenix Water-Wastewater and Water Services Departments.

which may be taken basically include water conservation practices and/or new water development projects. In the case when the availability of this resource is limited due to the drought, water conservation is a viable target. Marginal pricing of water demand is found to be one of the best alternatives in urban water conservation efforts. This is achieved due to the fact that urban water demand is elastic to price changes. In other words, an increase in urban water price decreases its demand.

TABLE 6.7 Price-Risk Relations for a Typical Drought Month

Available monthly supply, ft^3	Monthly demand, ft^3	Demand subject to surcharge, ft^3	Basic service charge, $	Price per additional 100 ft^3, $	Total price, $	Difference $d_x - q_x$	Safety factor (SF)	Risk
2695	2800.00	1800.00	5.16	1.61	34.14	105.00	0.963	0.500
2695	2791.13	1791.00	5.16	1.65	34.78	96.13	0.966	0.500
2695	2784.28	1784.28	5.16	1.68	35.28	89.28	0.968	0.500
2695	2777.00	1777.00	5.16	1.72	35.82	82.00	0.970	0.500
2695	2769.83	1769.83	5.16	1.76	36.37	74.83	0.973	0.500
2695	2762.65	1762.65	5.16	1.80	36.93	67.65	0.976	0.500
2695	2755.49	1755.49	5.16	1.83	37.50	60.49	0.978	0.500
2695	2748.34	1748.34	5.16	1.87	38.07	53.34	0.981	0.500
2695	2741.21	1741.21	5.16	1.92	38.66	46.21	0.983	0.500
2695	2734.10	1734.10	5.16	1.96	39.26	39.10	0.986	0.500
2695	2727.00	1727.00	5.16	2.00	39.87	32.00	0.988	
2695	2719.92	1719.92	5.16	2.05	40.48	24.92	0.991	
2695	2712.85	1712.85	5.16	2.09	41.11	17.85	0.993	
2695	2705.80	1705.80	5.16	2.14	41.75	10.80	0.996	
2695	2698.77	1698.77	5.16	2.18	42.40	3.77	0.999	
2695	2691.75	1691.75	5.16	2.23	43.07	−3.25	1.001	
2695	2684.75	1684.75	5.16	2.28	43.74	−10.25	1.004	

For urban water planning purposes during a drought condition, the inclusion of a third dimension termed as the risk, the probability that the demand exceeds the supply, plays an important role in deciding, for a given drought condition, the tolerable risk for a given affordable price or vice versa. Thus the risk-price-return period relationship developed will help as one of the decision support systems for urban water planning under drought conditions.

The use of the risk–safety factor approach under sustained drought conditions may be also useful. It not only has the advantage of lowering the demand during such conditions, but it is also a more effective way of controlling water use. This is because the water-use control framework is enforced at the water meters of the individual customers.

Although water demand is believed to be elastic to price, it is apparently not significant for small price increases. The fact that the risk is highly sensitive to the return period and less sensitive to the price demands that there be a strong commitment to conserve water under adverse drought conditions. The price-elasticity study has been geographically limited, mainly to the United States. More research

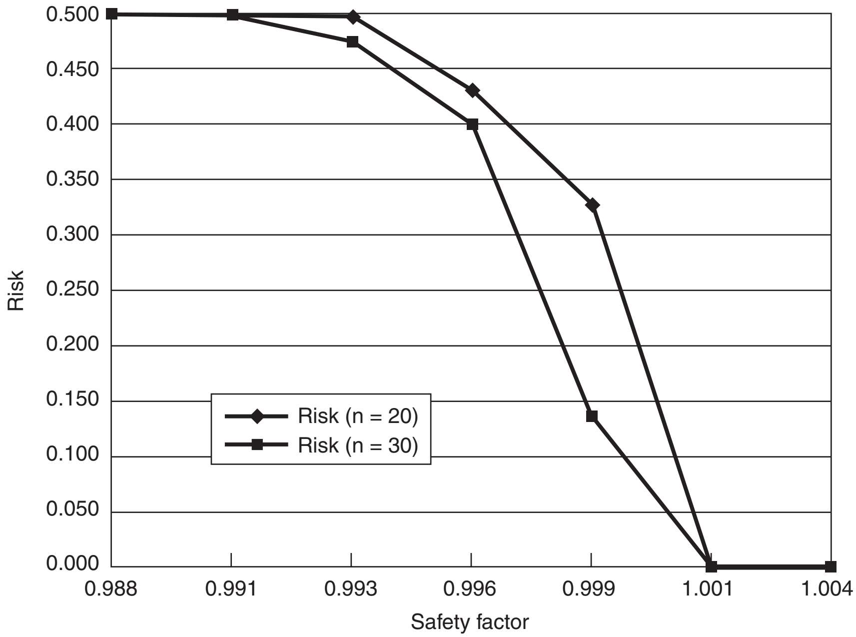


FIGURE 6.19 Risk–safety factor relationship for sample sizes of 20 and 30.

is needed in regions with a different socioeconomic status and hydrometeorological setup, that is, in regions where the incomes of the customers are relatively low and the weather conditions are more uncertain and/or the drought occurrences are more frequent. It is likely that the use of the concept of price elasticity will be more effective in such areas.

6.6 REFERENCES

Agthe, D. E., and R. B. Billings (1980). Dynamic Models of Residential Water Demand, *Water Resources Research,* 16(3):476–480.

Billings, R. B., and D. E. Agthe (1980). Price Elasticities for Water: A Case of Increasing Block Rates, *Land Economics,* 56(1):73–84.

Bruins, H. (1993). Drought Risk and Water Management in Israel: Planning for the Future, in Donald A. Wilhite (ed.), *Drought Assessment, Management and Planning: Theory and Case Studies,* Kluwer Academic Publishers, Boston.

Carver, P. H., and J. J. Boland (1980). Short- and Long-Run Effects of Price on Municipal Water Use, *Water Resources Research,* 16(4):609–616.

Changnon, S. (1993). Are We Doomed to Fail in Drought Management? in Martin Ruess (ed.), *Water Resources Administration in the United States: Policy, Practice, and Emerging Issues,* Michigan State University Press, East Lansing, Mich., 194–202.

Chow, V. T., D. R. Maidment, and L. W. Mays (1988). *Applied Hydrology,* McGraw-Hill, New York.

Dixon, L. S., et al. (1996). *Drought Management Policies and Economic Effects on Urban Areas of California, 1987–1992,* Rand Corporation, Santa Monica, Calif.

Dziegielewski, B., et al. (1986). Drought Management Options, in *Drought Management and Its Impact on Public Water Systems,* report on a colloquium sponsored by the Water Science and Technology Board, Sept. 5, 1985, National Academy Press, Washington, D.C.

Dziegielewski, B., et al. (1991). *National Study of Water Management During Drought: A Research Assessment for the US Army Corps of Engineers Water Resources Support Center,* Institute for Water Resources, IWR Report 91-NDS-3.

Easterling, W. E. (1993). Assessing the Regional Consequences of Drought: Putting the MINK Methodology to Work on Today's Problems, in Donald A. Wilhite (ed.), *Drought Assessment, Management and Planning: Theory and Case Studies,* Kluwer Academic Publishers, Boston.

Garen, D. C. (1993). Revised Surface-Water Supply Index for Western United States, *Journal of Water Resources Planning and Management,* ASCE, 119(4):437–454.

Gibbs, K. (1978). Price Variables in Residential Water Demand Models, *Water Resources Research,* 14(1):15–18.

Grebenstein, C. R., and B. C. Field (1979). Substituting for Water Inputs in U.S. Manufacturing, *Water Resources Research,* 15(2):228–232.

Griffin, R. C., and J. Stoll (1983). The Enhanced Role of Water Conservation in the Cost-Benefit Analysis of Water Projects, *Water Resources Bulletin,* 19(3):447–457.

Grunewald, O. C., et al. (1976). Rural Residential Water Demand: An Econometric and Simulation Analysis, *Water Resources Bulletin,* 12(5):951–961.

Hanke, S. H. (1970). Demand for Water under Dynamic Conditions, *Water Resources Research,* 6(5):1253–1261.

Hanke, S. H., and L. de Maré (1982). Residential Water Demand: A Pooled, Time Series, Cross-Section Study of Malmö, Sweden, *Water Resources Bulletin,* 18(4):621–625.

Hobbs, B. F. (1989). An Overview of Integrated Water Supply System Availability, in Larry W. Mays (ed.), *Reliability Analysis of Water Distribution Systems,* ASCE, New York.

Howe, C. W. (1982). The Impact of Price on Residential Water Demand: Some New Insights, *Water Resources Research,* 18(4):713–716.

Howe, C. W. (1993). Water Pricing: An Overview, *Water Resources Update,* 92:3–6.

Howe, C. W., and F. P. Linaweaver, Jr. (1967). The Impact of Price on Residential Water Demand and Its Relation to System Design and Price Structure, *Water Resources Research,* 3(1):13–32.

Hudson, Jr., H. E., and R. Hazen (1964). Drought and Low Stream Flows, in Ven Te Chow (ed.), *Handbook of Applied Hydrology,* McGraw-Hill, New York.

Jones, C. V., and J. R. Morris (1984). Instrumental Price Estimates and Residential Water Demand, *Water Resources Research,* 20(2):197–202.

Jordan, J. L. (1994). The Effectiveness of Pricing as a Stand-Alone Water Conservation Program, *Journal of the American Water Resources Association,* 30(5):871–877.

Kiefer, J. C. (1994). City of Phoenix Water Use Monitoring Program: A Multi-Objective Study of Single-Family Household Water Use (Unpublished Technical Report), Planning and Management Consultants, Ltd., Phoenix.

Lansey, K. E., et al. (1989). Methods to Analyze Replacement-Rehabilitation of Water Distribution System Components, in Larry W. Mays (ed.), *Reliability Analysis of Water Distribution Systems,* ASCE, New York.

Maddock, T. S., and W. G. Hines (1995). Meeting Future Public Water Supply Needs: A Southwest Perspective, *Journal of the American Water Resources Association,* 31(2):317–329.

Mays, L. W., and Y. K. Tung (1992). *Hydrosystems Engineering and Management,* McGraw-Hill, New York.

Moncur, J. E. T. (1987). Urban Water Pricing and Drought Management, *Water Resources Research,* 23(3):393–398.

Moncur, J. E. T. (1989). Drought Episodes Management: The Role of Price, *Water Resources Bulletin,* 25(3):499–505.

Montgomery, D. C., and G. C. Runger (1994). *Applied Statistics and Probability for Engineers,* John Wiley, New York.

National Research Council (1986). *Drought Management and Its Impact on Public Water Systems, Colloquium 1,* National Academy Press, Washington, D.C.

Palmer, W. C. (1965). Meteorological Drought, U.S. Weather Bureau, Research Paper No. 45.

Puckett, L. J. (1981). *Dendroclimatic Estimates of a Drought Index for Northern Virginia: Geological Survey Water Supply Paper 2080,* U.S. Government Printing Office, Washington, D.C.

Schneider, M. L., and E. E. Whitlatch (1991). User-Specific Water Demand Elasticities, *Water Resources Planning and Management,* 117(1):52–73.

Sheer, D. P. (1980). Analyzing the Risk of Drought: The Occoquan Experience, *Journal of American Water Works Association,* 72(12):246–253.

Steila, D. (1972). *Drought in Arizona: A Drought Identification Methodology and Analysis,* University of Arizona, Tucson.

Suter II, G. W. (1993). *Ecological Risk Assessment,* Lewis Publishers, Chelsea, Mich.

Tung, Y. K. (1996). Uncertainty and Reliability Analysis, in Larry W. Mays (ed.), *Water Resources Handbook,* McGraw-Hill, New York.

U.S. Soil Conservation Service Monthly Report (Jan. 1988). Colorado Water Supply Outlook, United States Department of Agriculture, U.S. Government Documents.

U.S. Water Resources Council (1979a). Procedures for Evaluation of National Economic Development (NED) Benefits and Costs in Water Resources Planning (Level C), Final Rule, Federal Register, 44(242):72891–72976.

Wilhite, D. A. (1993). Planning for Drought: A Methodology, in Donald A. Wilhite (ed.), *Drought Assessment, Management and Planning: Theory and Case Studies,* Kluwer Academic Publishers, Norwell, Mass., pp. 87–108.

Wong, S. T. (1972). A Model on Municipal Water Demand: A Case Study of Northern Illinois, *Land Economics,* 48(1):34–44.

Young, R. A. (1973). Price Elasticity of Demand for Municipal Water: A Case Study of Tucson, Arizona, *Water Resources Research,* 9(4):1068–1072.

Young, R. A. (1996). Water Economics, in Larry W. Mays (ed.), *Water Resources Handbook,* McGraw-Hill, New York.

CHAPTER 7
COMPUTER PROGRAMS FOR INTEGRATED MANAGEMENT

Messele Z. Ejeta
California Department of Water Resources
Sacramento, California

Larry W. Mays
Department of Civil and Environmental Engineering
Arizona State University
Tempe, Arizona

7.1 INTRODUCTION

The facts that every living being depends on water to live and that water's availability is limited in terms of both quantity and quality, makes water a resource over which there is growing competition. The various competitors for water have made it often challenging in space and time to fully satisfy their needs. The viable solution under such conditions is balancing out. This may be achieved through integrated hydrosystems management.

The concept of integrated hydrosystems management and computer programs used for integrated hydrosystems management are discussed in this chapter. The term *integrated hydrosystems management* is used for various types of water systems. A review of computer programs for hydrosystems management developed prior to the 1960s up to the present day shows enormous evolution and revolution of computing applications to hydrosystems. These efforts, which started in the early days of computer programming for the simplification of calculations of analytical functions, have now reached the age of what is being referred to in computing technology as *artificial intelligence.* In essence, it has become possible to write computer programs that not only evaluate a hydrosystems problem, but also draw preliminary conclusions based on the results and recommend appropriate actions based on the conclusions.

Various definitions have been given in the past to integrated resource management in general and water resources management in particular by different authors and institutions involved in the study of water resources. In addition, various terms such as hydrosystems management, integrated water management, integrated regional water management, water resources management, river basin management, watershed management, and total water management have been used to refer to the management of water resources on a large scale. Herein, the term integrated hydrosystems management is consistently used unless otherwise specified.

The definitions of integrated hydrosystems management as used by various institutions and individuals are cited and an attempt is made to give a definition that considers its wide range of aspects. The evolution of simulation computer programs and the structure of optimization techniques for hydrosystems problems are revisited. Examples of a relatively new set of computer programs, generally termed *decision support systems* (DSSs), are reviewed. Some of the examples of DSSs given for integrated hydrosystems management manifest the possibility of incorporating or at least monitoring water policy issues in the process of allocating water to all the competing users.

Computer programs and computing techniques useful for hydrosystems management have been categorized into simulation programs, optimization techniques, and DSSs. The discussion of different optimization techniques ranging from mathematical programming to heuristic search techniques including genetic algorithms and simulated annealing shows the potential resources available for computer programming for integrated hydrosystems management. Incorporating established water resources operation policies in simulation and optimization computer programs may help develop DSSs that can be used for integrated hydrosystems management. The study of a few such computer programs manifests the relative importance of these computer programs for integrated hydrosystems management. Only a limited number of DSSs for this purpose have been developed and used in the past. However, the availability of technical resources including database management systems, simulation computer programs, optimization techniques, and advanced computing technology provide the opportunity for more exploration to develop DSS for integrated hydrosystems management.

7.2 INTEGRATED HYDROSYSTEMS MANAGEMENT

7.2.1 Definition

Mitchell (1990) noted that integrated water management may be contemplated in at least three ways. These include (1) the systematic consideration of the various dimensions of water: surface and groundwater, quality and quantity; (2) the implication that while water is a system it is also a component which interacts with other systems; and (3) the interrelationships between water and social and economic development. In the first thought, the concern is the acceptance that water comprises an ecological system, which is formed by a number of interdependent components. In the second one, the interactions between water, land, and the environment, which involve both terrestrial and aquatic issues, are addressed. Finally, the concern is with the relationships between water and social and economic development, since availability or lack of water may be viewed as an opportunity for or a barrier against economic development. Each aspect of integrated hydrosystems management depends on and is affected by other aspects. Loucks (1996) points out that "integrated water resources systems planning and management focuses not only on the performance of individual components, but also on the performance of the entire system of components."

Water policy issues, of which limited effort was made in the past to incorporate them into hydrosystems computer programs, are some of the major factors that affect integrated hydrosystems management. Grigg (1998) describes water policy as dealing with finding satisfactory ways to allocate resources to balance between diverse and competing objectives of society and the environment. He referred to *integrated water management* as blending together actions and objectives favored by different players to achieve the best total result. Mitchell (1990) states that integration in water management deals with "…problems that cut across elements of the hydrological cycle, that transcend the boundaries among water, land and environment, and that interrelate water with broader policy questions associated with regional economic development and environmental management." The policies that are needed for integrated water resources management require coordination and collaboration among governments and agencies engaged in water management (Viessman, 1998). Grigg (1998) notes that improving coordination is the most promising route to the conceptual and perhaps utopian vision of integrated water management.

AACM, a consulting company in Australia, and the Center for Water Policy Research, Australia, in 1995 defined integrated resource management, of which water resources is a part, as the coordinated management of land and water resources within the region, with the objectives of controlling and/or conserving the water resource, ensuring biodiversity, minimizing land degradation, and achieving specified and agreed-upon land and water management and social objectives (Hooper, 1995). The American Water Works Association Research Foundation (AWWARF) (1996) defined the concept of total water management, which comprehends wide aspects of integrated hydrosystems management, through the following statements.

> Total Water Management is the exercise of stewardship of water resources for the greatest good of society and the environment. A basic principle of Total Water Management is that the supply is renewable, but limited, and should be managed on a sustainable use basis. Taking into consideration local and regional variations, Total Water Management:

- Encourages planning and management on a natural water systems basis through a dynamic process that adapts to changing conditions;
- Balances competing uses of water through efficient allocation that addresses social values, cost effectiveness, and environmental benefits and costs;
- Requires the participation of all units of government and stakeholders in decision-making through a process of coordination and conflict resolution;
- Promotes water conservation, reuse, source protection, and supply development to enhance water quality and quantity; and
- Fosters public health, safety, and community good will.

Table 7.1 shows an elaboration by Grigg (1998) of the definition of total water management as related to the concept of coordination. He emphasized what is implied by each of the important phrases used in the definition. These phrases which are apparently the central aspects of integrated hydrosystems management include society and environment, stakeholder, watershed and natural water systems, means of water management, time-wise, intergovernmental, water quality and quantity, local and regional concerns, and competing uses. Integrated hydrosystems management is as much challenging as compromising between these different aspects in making decisions.

The foregoing definitions and discussions indicate that integrated hydrosystems management is multiobjective. It is necessary both for economic efficiency (which is measured in monetary units) and for environmental quality (which may be measured in terms of pollutant concentration). Shortly, it balances between societal welfare and ecosystem sustainability. To summarize, integrated hydrosystems management in a watershed involves a multidisciplinary approach of developing and using water resources by making possible balances between all the competing water uses and through coordination between all parties without causing detrimental consequences to the ecosystem and/or future requirements.

7.2.2 History

Jamieson and Fedra (1996) report that practitioners have recognized the concept of integrated hydrosystems management since the early 1970s. The United Nations in the Dublin Statement endorsed this perception in 1992. Thus, integrated hydrosystems management on a regional scale is a newly emerging approach. The lack of a clear definition of a water resources region has, perhaps, contributed to the absence in the past of a regional approach to hydrosystems management.

River basin boundaries often differ from political boundaries. Groundwater flow has obviously never been dictated by political boundaries, and neither has the movement of atmospheric water. Furthermore, the question of the size of a region has been a challenge and will probably remain so in the near future. Viessman (1998) states that it is not clear that integrated regional water plans can be fit within the geographic limits of large river basins or watersheds. Vlachos (1998)

TABLE 7.1 Types of Coordination from Total Water Management Definition

Type of coordination	Phrase from total water management definition	Discussion	Effectiveness ranking
Society and environment	The exercise of stewardship of water resources for the greatest good of society and the environment	This statement provides a general organizing framework for balancing. It is adequately understood, but needs more explanation.	1
Stakeholder	Requires the participation of all…stakeholders in decision making through a process of coordination and conflict resolution	Process is known as stakeholder and public involvement. Good and improving. A central issue of democratic government.	2
Watershed and natural water systems	Encourages planning and management on a natural water systems basis	It is recognized and currently popular that water management on a basin or watershed basis is desirable. Further progress will require more effort.	3
Means of water management	Promotes water conservation, reuse, source protection, and supply development	This means to coordinate different ways to meet needs and sustain the environment. A central planning and management issue.	4
Time-wise	Through a dynamic process that adapts to changing conditions	This requires valid planning methods to preserve institutional memory and keep processes on track and requires much improvement.	5
Intergovernmental	Requires the participation of all units of government in decision making through a process of coordination and conflict resolution	Intergovernmental coordination is given as separate from stakeholders because of the different kinds of authorities that government has.	6
Water quality and quantity	To enhance water quality and quantity	This is handled through water quality law and regulation. Many problems still require solution.	7
Local and regional concerns	Taking into consideration local and regional variation	This is a difficult issue requiring intergovernmental cooperation in arenas which lack adequate incentives and often cannot be mandated. It is not working too well.	8

TABLE 7.1 (*Continued*)

Type of coordination	Phrase from total water management definition	Discussion	Effective-ness ranking
Competing uses	Balances competing uses of water through efficient allocation that addresses social values, cost-effectiveness, and environmental benefits and costs	This is handled through state and federal water law regulations, court decisions, and other institutions. A very difficult arena.	9

Source: Grigg (1998).

poses a very important question: Can integrated planning and management work in the vast expanses of the Nile, the Amazon, the Paraná/La Plata, or should it be restricted to more regional, specific sociopolitical conflicts of rather well-defined geographic, cultural, environmental, physiographic, and economic boundaries?

Defining a water resources region now appears to be driven more by the watershed approach than the other factors mentioned above. A national forum convened in January 1994 by the Conservation Fund and the National Geographic Society clearly recognized the critical need for the watershed approach for integrated hydrosystems management rather than political jurisdiction or boundaries. Similarly, the Environmental Advisory Board (EAB) of the U.S. Army Corps of Engineers (USACE) recommended in 1994 that the watershed or ecosystem approach be used as the holistic, integrated concept on which to base (water resources) planning (Bulkley, 1995). Furthermore, the U.S. General Accounting Office (1994) listed the importance of the watershed approach for integrated management. Accordingly, watershed boundaries

1. Are relatively well defined
2. Can have major ecological importance
3. Are systematically related to one another hierarchically and thus include smaller ecosystems
4. Are already used in some water management efforts
5. Are easily understood by the public

Many water resources projects in the past lacked the integrated planning aspect. Hall (1998) states that throughout history, water management systems have been developed in a linear fashion; i.e., they had a piecemeal development in which the components of integrated water management were put into place as the need for each component arose. Similarly to the management aspect, the modeling aspect of water resources has also followed the same piecemeal fashion. Jamieson and Fedra (1996) state that "although the principle of integrated river basin management models has been aspired to in many countries, more often than not the problems have been considered in a piecemeal fashion, with experts from different disciplines using separate computer programs (water resources, surface-water pollution control, groundwater contamination, etc.), to tackle parts of the overall problem in a reactive way." Uncoordinated hydrosystems modeling efforts often result in incompatibilities. As a result, these systems have not been sufficient and effective enough, thus leading to the emergence of the regional management and modeling approach.

7.2.3 Importance

We are becoming increasingly aware, with time, of the fact that our water supplies are limited both in quantity and quality. Because water has multiple and often competing uses, hydrosystems are interrelated with other physical and socioeconomic systems. In some locations, when water supplies become extremely limited, their further use is based on the determination of which user has the oldest "right" to them, or on a judgment about which uses have the highest priority (Hall, 1998). Hall (1998) also warns that unless dealt with appropriately, the forces of population growth; urbanization; and increased water demands for home, industry, and agriculture, coupled with an increasingly global economy and culture, will in the future produce spreading, perilous degradation of water quality everywhere, and a continuously widening gap between water needs and the availability of useful water in all too many locations. As a solution to this problem, he suggested a different approach, which includes: (1) management across political boundaries; (2) the collective management of atmospheric water, surface waters, and groundwater; (3) the combined management of water quality and water quantity.

Schultz (1998) gives the criteria for water resources management projects at present and those criteria emerging as new ones in the future. Accordingly, the factors that have to be satisfied include: (1) economic benefits; (2) technical efficiency; and (3) performance reliability. The approach, which seems to become more and more dominant, includes

1. Principle of sustainable development
2. Ecological quality
3. Consideration of macroscale systems and effects
4. Planning in view of changes in natural and socioeconomic systems

It is evident from these comparisons that hydrosystems projects are geared toward integrated management. These new planning approaches for integrated hydrosystems management necessitate new ways of modeling. Schultz (1998) adds that new planning tools are required to plan and design water resources systems on the basis of the new criteria. He concludes "since no planning tools following the four new criteria are available, we are faced with a vacuum."

In a different argument, an integrated hydrosystems project needs to be evaluated on the following important factors: technical, economic, financial, environmental, and sociopolitical. Technically, it must be feasible to build; economically, it must be reasonably affordable; financially, it must have a source; environmentally, its effect must be mitigated with ease; and sociopolitically, it must be acceptable to the public. The project can be successful if effective coordination prevails between the parties involved and if such parties are mandated to monitor clearly defined scope and regional coverage. In this argument also, integrated hydrosystems management is perceived to be a viable approach in planning efficient and successful water resources projects. In England and Wales, for example, regional water authorities whose boundaries were defined by the watersheds of the country enabled the replacement of 1600 separate water service entities with 10 regional watersheds (Bulkley, 1995).

7.3 COMPUTER PROGRAMS FOR INTEGRATED MANAGEMENT

The implementation of the ideals of integrated hydrosystems management necessitates analytical tools to simplify or assist the balancing-out process. Water policies need to be transformed into forms that can be understood and interpreted using analytical tools such as computer programs. Consequently, robust computer programs that not only solve the problems with analytical structure or mathematical formula but also are capable of reducing and incorporating water policies into analytical structure are required. Furthermore, they may be required to interpret the result of the computations, give conclusions based on the result, and make appropriate recommendations based on the conclusions reached.

As we apply computing technology to water resources modeling, the lack of conventional terms is apparent. In general, the terms *computer programs* and *models* have been used interchangeably. Models are also used to refer to the applications developed for specific computer programs. Herein, we will refer to computer programs as the tools developed to solve any problem in a specific general area of problems. In contrast, we will refer to a model as a specific representation of a physical or real system that is prepared to be solved by a given computer program. Thus, a computer program is generic whereas a model is specific. For instance, we refer to the U.S. Environmental Protection Agency's EPANET as a computer program and a representation of a particular city's water distribution system problem that can be solved by EPANET as a model.

Although tremendous work has been done in the past to develop computer programs for integrated hydrosystems management, only a few exist that address the overall framework of problems associated with integrated hydrosystems management. A few of the reasons may be attributable, among others, to

1. the lack of clear definition and better understanding of integrated hydrosystems management
2. the variation of water needs with space and time
3. the evolution (revolution) of computer programming

Computer modeling approaches that at least partly tried to address some of the concepts of integrated hydrosystems management are highly based on interfacing simple computer programs written and used for the analysis of specific hydrosystems problems. At the core of some advanced computer programs used for integrated hydrosystems management lie simple simulation modules, rule-based simulation modules (also known as expert systems), and optimization modules of hydrosystems problems. While many simulation and optimization modules have been developed and interfaced over the years by different institutions and agencies, the incorporation of rule-based simulation modules in computer programs for integrated hydrosystems management appears to have emerged recently as a sound approach. By incorporating rule-based simulation modules, it has become easier to manage decisions that involve several factors and water policies.

Section 7.3.1 discusses the development of simulation computer programs that emerged in the United States over the past few decades for the simulation of various types of hydrosystems problems. Section 7.3.2 discusses the basic mathematical structure of optimization computer programs, which may be viewed as generic tools that can be customized to specific hydrosystems problems.

7.3.1 Development of Hydrosystems Simulation Computer Programs

Over the past few decades, water resources professionals have witnessed the development of quite a number of hydrosystems simulation computer programs. Wurbs (1995) points out that a tremendous amount of work has been accomplished during the past 3 decades in developing computer programs for use in water resources planning and management. The majority of these programs, perhaps most of the earliest computer programs to be developed for water resources problems, may be viewed as simulation computer programs.

As information technology advances, hydrosystems simulation computer programs have generally gone through an evolutionary process. Figure 7.1 depicts the evolution of hydrosystems computer programs as classified into five generations (derived from the explanation given by Jamieson and Fedra, 1996). The first-generation codes, which tremendously simplified calculation of analytical functions through generic computer codes, are but mediocre by today's standards. One may draw an analogy between the coming into being of these codes and the transition

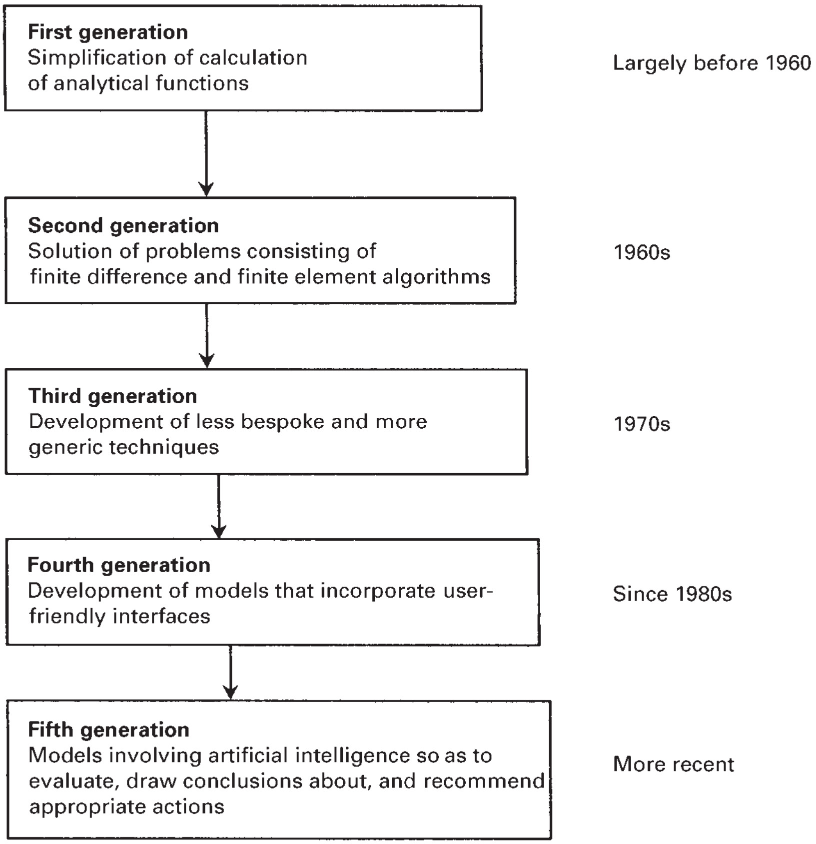


FIGURE 7.1 Schematic diagram showing the evolution of hydroinformatics. (After Jamieson and Fedra, 1996)

of computation methods from using the slide rule to scientific calculators. In both cases, similar jobs are done, but the new method highly reduced the time required for numerical computations. The succeeding generations of programs enhanced the robustness of the programs and/or the ease with which the programs can be used. The fifth generation of programs is embodied with artificial intelligence that not only performs analytical computations but also draws some preliminary conclusions and recommends appropriate actions.

7.3.2 Optimization Formulations

Background and General Formulation. Various optimization techniques, in general, and their application to various hydrosystems problems, in particular, have shown remarkable progress over the past 3 decades. The progress of the application of these techniques has occurred alongside the revolution of computer programs, and as such, similar explanations can be given to the development of simulation computer programs and optimization techniques over the past 3 or more decades. Figure 7.2 gives the development of the application of optimization techniques to hydrosystems problems, which is analogous to that given for simulation computer programs in Fig. 7.1.

The general formulation for optimization problems in water resources, which are generally nonlinear programming (NLP) problems, can be expressed in terms of state (or dependent) variables (**x**) and control (or independent) variables (**u**) as (Mays, 1997; Mays and Tung, 1992)

$$\text{Optimize } f(\mathbf{x}, \mathbf{u}) \tag{7.1}$$

subject to process simulation equations

$$G(\mathbf{x}, \mathbf{u}) = 0 \tag{7.2}$$

and additional constraints for operation on the dependent (**u**) and independent (**x**) variables

$$\underline{\mathbf{w}} \leq w(\mathbf{x}, \mathbf{u}) \leq \overline{\mathbf{w}} \tag{7.3}$$

where $\underline{\mathbf{w}}$ and $\overline{\mathbf{w}}$ represent lower and upper bounds, respectively. The term *optimize* in Eq. (7.1) refers to either maximization or minimization, whereas the constraint equation [(Eq. (7.3)] dictates the feasibility of the objective with respect to each and every constraint. In other words, the solution to the simulation equation [Eq. (7.2)] must satisfy the constraints defined by Eq. (7.3). The process simulation equations basically consist of the governing physical equations of mass, energy, and momentum.

Many hydrosystems problems can be formulated as discrete-time optimal control problems. The basic mathematical definition of a discrete-time optimal control problem is stated as

$$\underset{\mathbf{u}}{\text{Min}}\; Z = \sum_{t=1}^{T} f_t(\mathbf{x}_t, \mathbf{u}_t, t) \tag{7.4}$$

subject to

$$\mathbf{x}_{t+1} = g_t(\mathbf{x}_t, \mathbf{u}_t, t) \qquad t = 1, 2, \ldots, T \tag{7.5}$$

$$\mathbf{x}_t \geq 0 \qquad \mathbf{u}_t \geq 0 \tag{7.6}$$

where $\mathbf{x}_t$ = vector of state variables at time t
$\mathbf{u}_t$ = vector of control variables at time t
T = number of decision times

A few possible optimization formulations for different hydrosystems problems are given below.

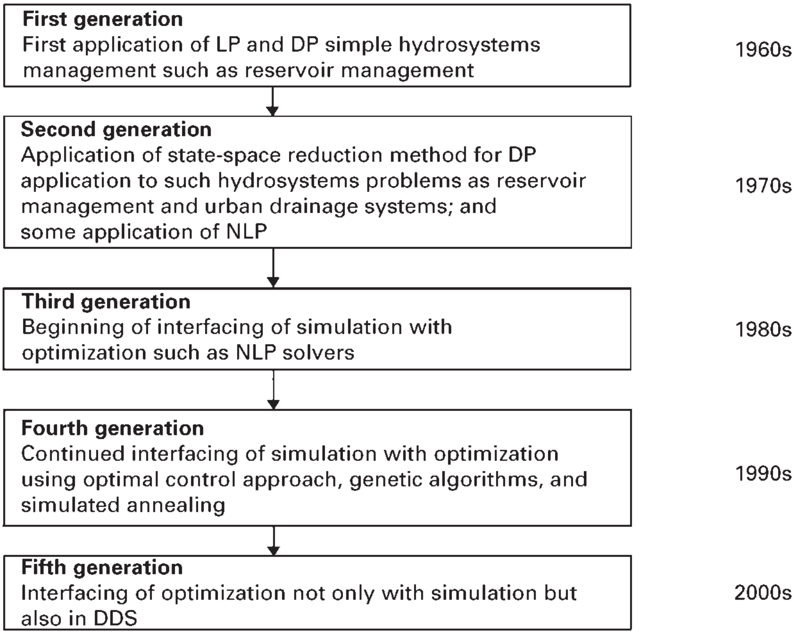


FIGURE 7.2 Schematic diagram showing the developments in the application of optimization techniques to hydrosystems problems.

Water Distribution System Operation. Mays (1997) defines the optimization problem for water distribution system operation in terms of the nodal pressure heads, **H**; pipe flows, **Q**; tank water surface elevations, **E**; pump operating times, **D**; and water quality parameter, **C**, as follows.

$$\text{Minimize energy costs } = f(\mathbf{H}, \mathbf{Q}, \mathbf{D}) \tag{7.7}$$

subject to

$$\mathbf{G}(\mathbf{H}, \mathbf{Q}, \mathbf{D}, \mathbf{E}, \mathbf{C}) = \mathbf{0} \tag{7.8}$$

$$\mathbf{w}(\mathbf{E}) = \mathbf{0} \tag{7.9}$$

$$\underline{\mathbf{H}} \leq \mathbf{H} \leq \overline{\mathbf{H}} \tag{7.10}$$

$$\underline{\mathbf{D}} \leq \mathbf{D} \leq \overline{\mathbf{D}} \tag{7.11}$$

$$\underline{\mathbf{E}} \leq \mathbf{E} \leq \overline{\mathbf{E}} \tag{7.12}$$

$$\underline{\mathbf{C}} \leq \mathbf{C} \leq \overline{\mathbf{C}} \tag{7.13}$$

where Eqs. (7.8) and (7.9) express the energy and flow constraints and the pump operation constraints. Equations (7.10) to (7.13) express the bound constraints on the nodal pressure head, pump operating times, tank water surface elevations, and water quality, respectively.

Reservoir System Operation for Water Supply. The optimization for this kind of hydrosystems problem can be expressed as (Mays, 1997)

$$\text{Maximize benefits} = \sum_{0}^{T} f(\mathbf{S}_t, \mathbf{U}_t, t) \tag{7.14}$$

subject to

$$\mathbf{G}(\mathbf{S}_{t+1}, \mathbf{S}_t, \mathbf{U}_t, \mathbf{I}_t, \mathbf{L}_t) = \mathbf{0} \qquad t = 0, \ldots, T - 1 \tag{7.15}$$

$$U_t \leq \mathbf{U}_t \leq \overline{\mathbf{U}}_t \qquad t = 1, \ldots, T \tag{7.16}$$

$$S_t \leq \mathbf{S}_t \leq \overline{\mathbf{S}}_t \qquad t = 1, \ldots, T \tag{7.17}$$

$$P[\mathbf{S}_t \geq S_t] \leq \alpha_t^{\min} \qquad t = 1, \ldots, T \tag{7.18}$$

$$P[\mathbf{S}_t \leq S_t] \leq \alpha_t^{\max} \qquad t = 1, \ldots, T \tag{7.19}$$

$$\mathbf{w}(\mathbf{S}_t, \mathbf{U}_t) = \mathbf{0} \tag{7.20}$$

where $\mathbf{S}_t$ and $\mathbf{U}_t$ are the vectors of reservoir storage and releases and t represents discrete time period. Equation (7.15) defines the system of equations of conservation of mass for the reservoirs and river reaches. $\mathbf{S}_{t+1}$ and $\mathbf{S}_t$ are, respectively, the vectors of reservoir storage at the beginning of time period $t + 1$ and t, $\mathbf{I}_t$ is the vector of hydrologic inputs, and $\mathbf{L}_t$ is the vector of reservoir losses. Equations (7.16) and (7.17) define the bound constraints on reservoir releases and storages, respectively. Equations (7.18) and (7.19) define the bound constraints on reservoir storage in probabilistic form where $P[\ ]$ denotes the probability and $\alpha_t^{\min}$ and $\alpha_t^{\max}$ represent the minimum and maximum reliability or tolerance levels, respectively. Equation (7.20) expresses the other constraints on reservoir operation.

Groundwater Management Subsystems. The general groundwater management problem can be expressed mathematically as (Mays, 1997)

$$\text{Optimize } Z = f(\mathbf{h}, \mathbf{q}) \tag{7.21}$$

subject to

$$\mathbf{G}(\mathbf{h}, \mathbf{q}, \mathbf{c}) = \mathbf{0} \tag{7.22}$$

$$\underline{\mathbf{q}} \leq \mathbf{q} \leq \overline{\mathbf{q}} \tag{7.23}$$

$$\underline{\mathbf{h}} \le \mathbf{h} \le \overline{\mathbf{h}} \tag{7.24}$$

$$\underline{\mathbf{C}} \le \mathbf{C} \le \overline{\mathbf{C}} \tag{7.25}$$

$$\underline{\mathbf{w}}(\mathbf{h}, \mathbf{u}) \le \mathbf{0} \tag{7.26}$$

where **h** and **q** in the objective function are vectors of heads and pumpages (or recharges), respectively. **C** is a parameter that measures quality such as chlorine content and so on. Equation (7.22) gives the general groundwater flow constraints, which represent a system of equations governing groundwater flow and transport. Equations (7.23) and (7.24) represent, respectively, the upper and the lower bounds on the pumpages (recharges) and on the heads. Equation (7.25) gives the groundwater quality constraints, whereas Eq. (7.26) may be taken as additional constraints which can be included to impose restrictions such as water demands, operating rules, and budgetary restrictions. It may be noted here that the lower and upper bounds on pumpages may or may not exist, whereas those on the head can be the bottom elevation of the aquifer and the groundwater surface elevations for the unconfined cells, respectively.

7.3.3 Interfacing Optimization and Simulation Computer Programs

The general form of the objective functions and the constraints in hydrosystems problems including the foregoing examples can be linear, nonlinear, or differential equations. The solution of each equation needs a different approach. Several computer codes have been written for each of these types of formulations.

For those hydrosystems optimization problems, which involve solving general governing differential equations of mass, energy, and momentum, the approach used can be solving the optimization problem directly by embedding finite difference or finite element equations of the governing process equations (Mays, 1997). This approach is relatively tedious to apply to real-world problems. Alternatively, an appropriate process simulator can be used to solve the constraint process simulation equations when they need to be evaluated for the optimizer. Consequently, the following general and simpler optimization problem can be used.

$$\text{Minimize } F(\mathbf{u}) = f(\mathbf{x}(\mathbf{u}), \mathbf{u}) \tag{7.27}$$

subject to

$$\underline{\mathbf{w}} \le \mathbf{w}(\mathbf{x}(\mathbf{u})) \le \overline{\mathbf{w}} \tag{7.28}$$

Different techniques have been successfully applied to solve optimization problems that are formulated in the above form. The most common techniques are given below.

Mathematical Programming. *Mathematical programming* includes linear programming and nonlinear programming problems (Jeter, 1986). Herein we will refer to the mathematical programming approach as interfacing simulation computer programs with nonlinear programming codes such as GRG2. This programming technique has been found useful in several hydrosystems problems such as groundwater management systems (Wanakule et al., 1986), and water distribution systems operation (Brion and Mays, 1989; Sakarya and Mays, 1999).

Various computer programs are available that solve either linear programming problems, nonlinear programming problems, or both. Table 7.2 gives a summary of some of the more popular optimization computer programs in the United States.

Differential Dynamic Programming. *Differential dynamic programming* (DDP) is a stagewise, nonlinear programming procedure that has been successfully applied to hydrosystems problems that are based on discrete-time optimal control, such as multireservoir operation and groundwater hydraulics (Mays, 1997).

A modified form of DDP, known as *successive approximation linear quadratic regulator* (SALQR), has been used for optimization problems in which nonlinear simulation equations are made linear in the optimization step (Culver and Shoemaker, 1992). For example, Carriaga and Mays (1995a,b) applied DDP to reservoir release optimization to control sedimentation. SALQR was applied to the operation of multiple reservoir systems to control sedimentation in alluvial river networks by Nicklow and Mays (2000, 2001); to the operation of soil aquifer treatment systems by Tang and Mays (1999); and to the determination of optimal freshwater inflows to bays and estuaries by Li and Mays (1995).

Genetic Algorithms and Simulated Annealing. *Genetic algorithms* (GA) are nonconventional search techniques patterned after the biological processes of natural selection and evolution (Tang and Mays, 1999). Genetic algorithms can be useful for the selection of parameters to optimize the performance of a system and for testing and fitting quantitative models (Chambers, 1995). Every solution of the optimization problem is represented in the form of a string of bits (integers or characters) that consist of the same number of elements, say n. Each candidate solution represented as a string is known as an *organism* or a *chromosome*. The variable in a position on the chromosome and its value in the chromosome are called the *gene* and the *allele*, respectively. For example, if $n = 3$, a general chromosome is $x = (x_1, x_2, x_3)$ where x_1, x_2, and x_3 are the genes on this chromosome in the three positions (Murty, 1995).

Genetic algorithms for optimization problems are developed by first transforming the problem into an unconstrained optimization problem so that every string of length n can be looked upon as a solution vector for the problem (Murty, 1995). Five tasks are required in the performance of a GA to solve the optimization

TABLE 7.2 Summary of Some of the Most Popular Optimization Computer Programs in the United States

Model name	Developed by	Model purpose	Remarks
LINDO	Lindo Systems, Inc.	Solves linear, quadratic, and integer programming problems	A user-friendly linear interactive and discrete optimizer (hence, the name LINDO)
LINGO	Lingo Allegro USA, Inc.	Solves linear and nonlinear programming problems	A sophisticated matrix generator; helps the user create large constraints objective function terms by writing one-line code
GRG2	Univ. of Texas	Solves nonlinear programming problems	Uses the generalized reduced gradient algorithm to find the optimal solution
GINO		Solves nonlinear programming problems	This model is a microcomputer version of GRG2
GAMS	GAMS Development Corporation	Solves linear programming problems	
MINOS	Saunders and Murthagh	Solves linear and nonlinear programming problems	Uses different algorithms when the problem has linear objective function and constraints, nonlinear objective function and linear constraints, and nonlinear objective function and constraints
GAMS/ ZOOM		Solves mixed integer programming problems	Adapted ZOOM (Zero/One Optimization Method).
GAMS/ MINOS		Solves linear and nonlinear programming problems	Adapted MINOS (Modular In-Core Nonlinear Optimization System)

problem: encoding, initialization of the population, fitness evaluation, evolution performance, and working parameters (Adeli and Hung, 1995).

The decision variable vector is encoded as a chromosome using mostly a binary number coding method. Therefore if there are m decision variables and if each decision variable is encoded as an n-digit binary number, then a chromosome is a string of $n \times m$ binary digits as shown in Fig. 7.3.

A population of chromosomes is initialized which requires randomly generating the initial population in such a way that all values for each bit have equal probability of being selected. The fitness measure at every feasible solution is equal to

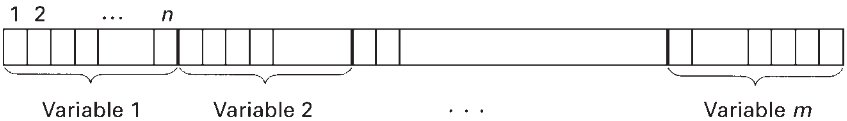


FIGURE 7.3 Typical representation of a chromosome of m decision variables of n bits each, used for encoding.

the objective function value at that point. Thus, a fitness evaluation is used to determine the probability that a chromosome will be selected as a parent chromosome to generate new chromosomes. Evolution performance involves selection, crossover, and mutation. Selection chooses the chromosome to survive for a new generation. Crossover is used to recombine two chromosomes (parent strings) and generate two new chromosomes (offspring strings) based on a predefined crossover criterion. Mutation serves as an operator to reintroduce "lost alleles" into the population based on a predefined mutation criterion. Working parameters guide the genetic algorithm and include chromosome length, population size, crossover rate, mutation rate, and stopping criterion.

Simulated annealing (SA) stems from an algorithm that is used for the application of statistical thermodynamics concepts to combinatorial optimization problems. A solution to a combinatorial optimization problem is based on statistical mechanics in which the best solution is obtained from a large set of feasible solutions. In essence, it is a type of local search (descent method) heuristic that starts with an initial solution and has a mechanism for generating a neighbor of the current solution. For minimization problems, if the generated neighbor has a smaller objective value, it becomes the new current solution; otherwise the current solution is retained. The process is repeated until a solution is reached with no possibility of improvement in the neighborhood (Murty, 1995).

This algorithm has the disadvantage that the local search stops at a local minimum (see Fig. 7.4). This can be avoided by running the local search several times starting randomly from different initial solutions. By doing so, the global minimum can be taken as the best of the local minima found.

A better approach to find the global minimum was introduced in 1953 by Metropolis et al. (Murty, 1995). In this attempt, annealing was applied to the search of the minimum energy configuration of a system after it is melted. At each iteration, the system is given a small displacement and the change in the energy of the system, δ, is calculated. If $\delta < 0$, the change in the system is accepted; otherwise, the change is accepted with probability $\exp(-\delta/T)$ where T is a constant times the temperature. This optimization technique has been applied to different problems in engineering, such as groundwater restoration (Skaggs and Mays, 1999), operation of water distribution systems (Sakarya et al., 1998; Goldman 1998; Goldman and Mays, 1999; Sakarya and Mays, 2000), for water quality purposes.

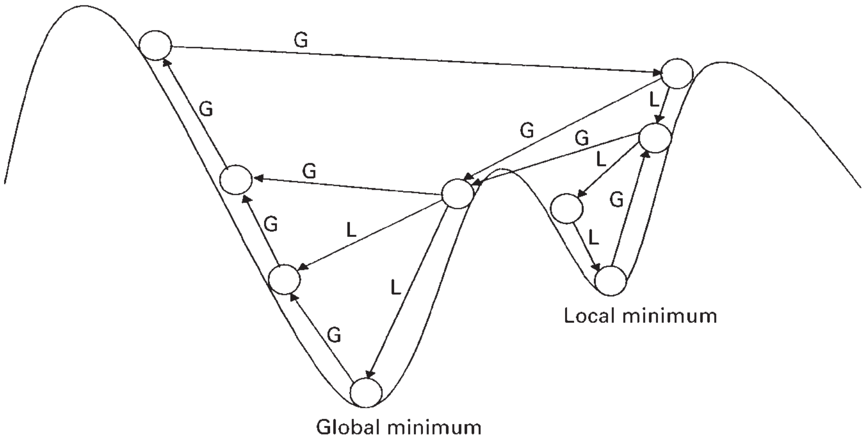


FIGURE 7.4 The objective function landscape. G = global, L = local. (After Topping et al., 1993)

Comparison of Heuristic Search Methods to Other Optimization Techniques. Whereas the heuristic search methods involve trial solutions, mathematical programming and DDP/SALQR follow some given procedures. On the other hand, mathematical programming and DDP/SALQR require derivative information. The optimal solution found by the mathematical programming approach may result in a very short operating time during one time interval that cannot be followed for practical purposes. In the simulated annealing approach, this problem can be minimized by setting the minimum period of operation (Sakarya et al., 1998).

The mathematical programming approaches find the optimum solution in much shorter operating times than the heuristic search approaches. Tang and Mays (1998) have developed a new methodology for the operation of soil aquifer treatment systems, formulated as a discrete-time optimal control problem. This new methodology is based upon solving the operations problem using a genetic algorithm interfaced with the one-dimensional unsaturated flow model HYDRUS (Kool and van Genuchten, 1991). The same problem has been solved by Li et al. (2000) using SALQR interfaced with the HYDRUS model. The computer time for a 10-cycle operation with the SALQR algorithm was reported as 654 CPU seconds, while with the genetic algorithm, it needed about 46,600 CPU seconds [about 13 hours (h)] on the same computer to obtain the optimal solution for a three-cycle operation.

Sakarya et al. (1998) have compared two newly developed methodologies, a mathematical programming approach and a simulated annealing approach, for determining the optimal operation of a water distribution system considering both quantity and quality aspects. Both methodologies formulate the problem as a discrete-time optimal control problem. The mathematical programming approach interfaces the GRG2 model (Lasdon and Warren, 1986), a generalized reduced gradient procedure, with the U.S. Environmental Protection Agency EPANET model (Rossman, 1994) for water distribution system analysis. The simulated annealing approach is also interfaced with the EPANET model. The study showed that while different results were obtained for total pump operation hours, the total 24-h energy costs were comparable.

7.3.4 Computer-Based Information Systems: Supervisory Control Automated Data Acquisition (SCADA)

SCADA is a computer-based system that can control and monitor several hydrosystems operations such as pumping, storage, distribution, and wastewater treatment. Several such systems have been developed in the past for different water supply agencies. For instance, the Metropolitan Sewer District of Cincinnati planned to integrate a SCADA system in the 1980s to monitor its wastewater treatment plants and pump stations. This system was planned for an area which consisted of 7 major treatment plants, 30 package wastewater plants serving individual subdivisions, and about 130 pump stations (Clement, 1996). A SCADA system developed in 1986 for the city of Honolulu had the capability of controlling and monitoring 57 source pumping stations, 126 storage reservoirs, and 73 booster pumping stations (Wada et al., 1986). In general, SCADA systems are designed to perform the following functions:

- Acquire data from remote pump stations and reservoirs and send supervisory controls
- Allow operators to monitor and control water systems from computer-controlled consoles at one central location
- Provide various types of displays of water system data using symbolic, bar graph, and trend formats
- Collect and tabulate data and generate reports
- Run water control software to reduce electrical power costs

Remote terminal units (RTUs) are used to process data from remote sensors at pump stations and reservoirs. The processed data are transmitted to the SCADA system also by the RTUs. Conversely, supervisory control commands from the SCADA system prompt the RTUs to turn pumps on and off and open and close valves.

7.3.5 Prospects of Computer Programs for Integrated Hydrosystems Management

There is no doubt that the first computer programs developed to solve hydrosystems problems targeted specific problems such as catchment runoff simulation, streamflow characterization, and water quality monitoring. With the enhancement of computing efficiency and speed over the past several years, more sophisticated and user-friendly computer programs for hydrosystems problems have been developed. However, the objective of most of the computer programs was not to address the problems of integrated hydrosystems management inasmuch as a consensus exists as to the definition of integrated hydrosystems management given in Sec. 7.2.

More recently, computer programs that attempt to provide support for decision makers have been brought into the picture. Such computer programs, generally termed DSSs, have manifested themselves at this time as promising computer programs for integrated hydrosystems management. Section 7.4 discusses the DSS applications for integrated hydrosystems management.

7.4 DSSs AS TOOLS FOR INTEGRATED HYDROSYSTEMS MANAGEMENT

7.4.1 Definition of DSSs

Decision support systems, as might be inferred from the name, do not refer to a specific area of specialty. It is not easy to connote a specific definition to DSSs based on their uses. Reitsma et al. (1996) point out that although some consensus exists as to the purpose of DSSs, "a single, clear, and unambiguous definition is lacking." Generally, however, DSSs give pieces of information, sometimes real-time information, that help make better decisions. Sprague and Carlson (1982) defined a DSS as an interactive computer-based support system that helps decision makers utilize data and computer programs to solve unstructured problems.

7.4.2 Basic Structure of DSSs

A DSS generally consists of three main components: (1) state representation, (2) state transition, and (3) plan evaluation (Reitsma et al., 1996). State representation consists of information about the system in such forms as databases. State transition takes place through modeling such as simulation. Plan evaluation consists of evaluation tools such as multicriteria evaluation, visualization, and status checking (Reitsma, 1996). The three main components comprise the database management subsystem, model base management subsystem, and dialog generation and management subsystem, respectively. Figure 7.5 depicts these subsystems including their specific purposes and functions. Some examples of DSSs for different integrated hydrosystems management are presented later in this section.

Jamieson and Fedra (1996) elaborated on the basic structure of the WaterWare DSS (Fig. 7.6). As shown in Fig. 7.6, each subsystem is made up of different components. The data management subsystem can use different tools such as geographic information systems (GIS) as well as other simplistic data. The model

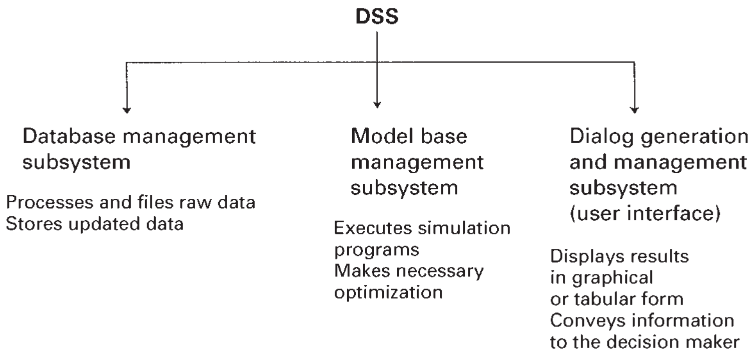


FIGURE 7.5 Basic components of a typical DSS.

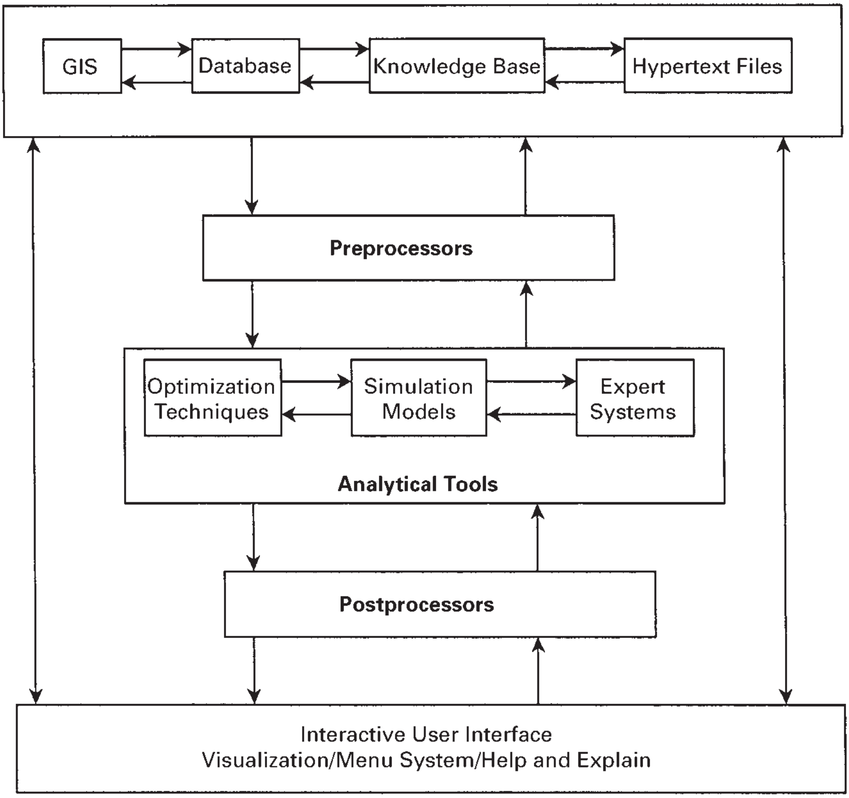


FIGURE 7.6 Basic structure of the WaterWare DSS. (After Jamieson and Fedra, 1996)

base subsystem consists of simple simulation computer programs, optimization techniques, and expert systems (also sometimes known as rule-based simulation computer programs). The dialog generation and management subsystem helps in visualization and making decisions through interactive user interface.

The structure of DSSs discussed above has, perhaps, made them the best structured and most promising computer programs for integrated resource management. These programs are believed to contribute largely to this objective. Reitsma et al. (1996) pointed out that "…the next few years will be most interesting" for DSSs. This stems from the fact that DSSs are promising computer programs for integrated hydrosystems management and the advance in the computing and information technology is remarkable.

7.4.3 Examples of DSSs for Integrated Hydrosystems Management

DSS for Trinity River Basin, Texas. One of the integrated DSSs in regional hydrosystems management was developed for the Trinity River in Texas (Ford and Killen, 1995). This DSS has the capability of integrating three major hydrosystems problems. Accordingly, it has three components which perform the following tasks: (1) retrieve, process, and file rainfall and streamflow data; (2) estimate basin average rainfall and forecast runoff; and (3) simulate reservoir operation in order to forecast regulated flows basinwide. Each of the tasks is done by the DSS subsystems, which use existing computer programs. The first subsystem, the *data-retrieval, processing, and filing subsystem,* retrieves data that are collected from an existing precipitation and streamflow gauge network and stores the data using a time series *database management system* (DBMS) designated as HEC-DSS. The second subsystem, the *rainfall estimating and runoff forecasting subsystem*, uses the following computer programs: (1) PRECIP to compute catchment areal-average rainfall and (2) HEC-1F for forecasting runoff. The third subsystem, the *reservoir simulation subsystem*, uses HEC-5 that is customized and fitted to basin conditions.

Figure 7.7 shows different components of this DSS that are used for forecasting streamflow. The Trinity River Advanced Computing Environment (TRACE) is the forecaster's interface of the DSS. It executes programs PRECIP, HEC-1F, and HEC-5 with the proper input. It also serves as a file manager, input processor, and DBMS interface. Furthermore, it executes, behind the scenes, programs PREFOR and PREOP to complete the HEC-1F and HEC-5 files, respectively. The DBMS-interface component of TRACE executes program EXTRCT to create working copies of data records, program DISPLAY to graph data, and program DWINDO to tabulate and edit data (Ford and Killen, 1995).

The size of the Trinity River basin for which this DSS was developed is approximately 4.6 million hectares (17,800 mi^2). Seven multipurpose major reservoirs having a total capacity of approximately 13.63 billion m^3 (11,080,000 acre·ft) are found in the basin.

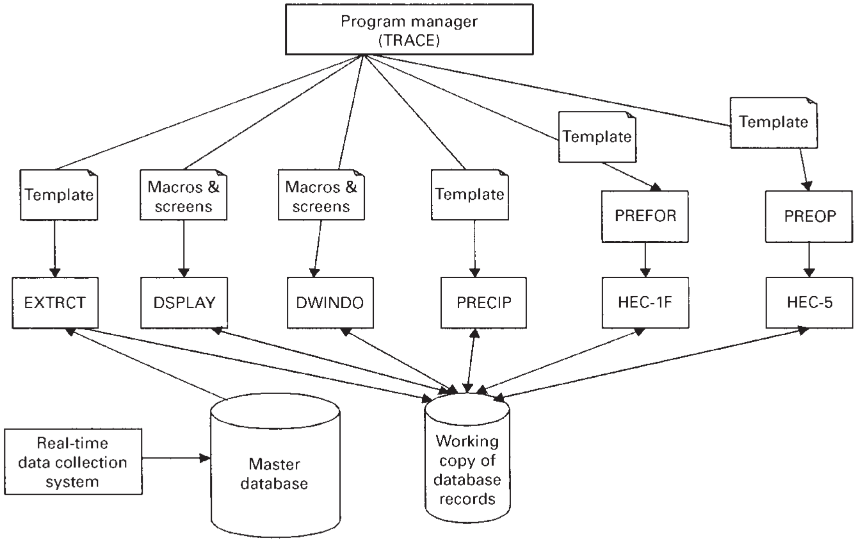


FIGURE 7.7 Components of forecasting software, Trinity River basin, Texas. (Ford and Killen, 1995)

TERRA (TVA Environment and River Resource Aid). TERRA is a DSS developed for the Tennessee Valley Authority (TVA) and the Electric Power Research Institute (EPRI) (Reitsma et al., 1996). It was developed for the management of the TVA River, reservoir, and power resources. TERRA has the following characteristics:

1. Consists of georelational database
2. Serves as the central data-storage and retrieval system
3. Records the TERRA information flow
4. Supports interfacing specialized data management software
5. Has various visualization tools
6. Checks the data entering the database or data from both resident and nonresident computer programs against various sets of operational constraints (environmental, recreational, special/emergency, navigational, and so on)

TERRA consists of the three essential components of a DSS, namely, (1) management of the state information of the TVA river basin, (2) the computer programs for conducting simulations and optimizations, and (3) a comprehensive set of reporting and visualization tools for studying, analyzing, and evaluating current and forecast states of the river system.

PRSYM (Power and Reservoir System Model). This model is used for river, reservoir, and power systems. It provides a tool for scheduling, forecasting, and planning reservoir operations. It integrates the multiple purposes of reservoir systems such as flood control, navigation, recreation, water supply, and water quality, with power system economics by solving the problem based on pure simulation, rule-driven simulation, or a goal programming optimization (Zagona et al., 1995).

Shane et al. (1995) note that PRSYM represents a major advance in modeling flexibility, adaptability, and ease of use, which enables users to

1. Visually construct a model of their reservoir configuration using "icon programming" with icons representing reservoir objects, stream reach objects, diversions, etc.
2. Select appropriate engineering functions, standardized by the industry, to reflect object characteristics needed for schedule planning, e.g., reservoir and stream routing methods.
3. Replace outdated functions with improved versions developed by industry.
4. Develop and include functions that are unique to their systems.
5. Experiment with operating policies.
6. Use data display and analysis objects to customize data summary presentations.

CALSIM. The CALSIM computer program is a general-purpose planning program developed by the California Department of Water Resources (DWR) in collaboration with the U.S. Bureau of Reclamation (USBR). It is developed for the coordinated operation of the California State Water Project (SWP) and the Federal Central Valley Project (CVP). It replaces DWR's prior planning simulation program DWRSIM used to operate the SWP, as well as USBR's PROSIM and SANJASM programs used to operate the CVP. It utilizes optimization technique to route water through a network of nodes and links. The user can formulate the operation of water resources systems as a linear programming (LP) or a mixed integer linear programming (MILP) problem. The operational constraints are formulated using a programming language known as Water Resources Engineering Simulation Language (WRESL). Data can be managed using a relational database management system and/or the United States Army Corps of Engineers data storage system (dss, not to be confused with DSS). CALSIM uses an LP/MILP solver, known as XA solver, to determine the optimal set of decisions given a set of user-defined priorities or weights and a set of system constraints. Figure 7.8 gives the CALSIM program components and structure.

Conjunctive Stream-Aquifer Management. This DSS is used for conjunctive management of surface water and groundwater under the prior appropriation

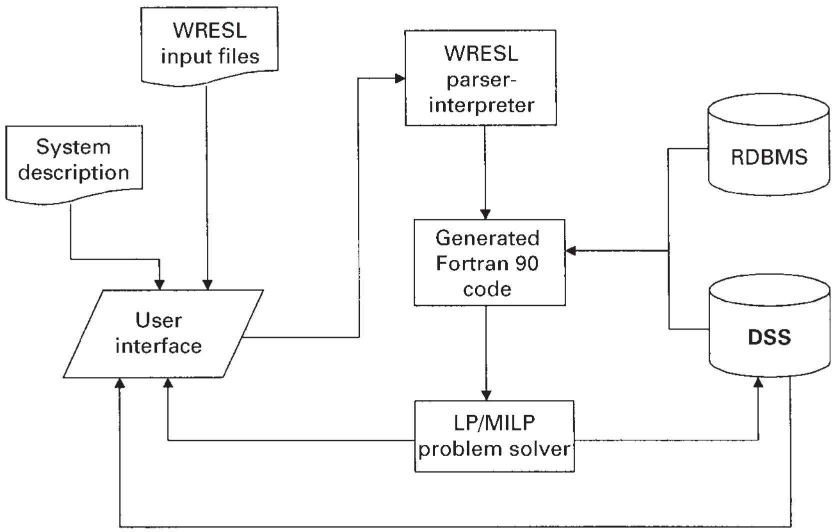


FIGURE 7.8 CALSIM model components and structure.

water right (Fredericks et al., 1998). It has the three components, which are typical of a DSS: database management subsystem, model base management subsystem, and a dialog generation and management subsystem or user interface. It is possible to prepare input data files for this DSS using GIS. The overlay of the GIS raster or grid database with other aquifer grid data enabled the finite groundwater model MODFLOW to readily read these data.

RiverWare. Developed by the Center for Advanced Decision Support for Water and Environmental Systems (CADSWES) at the University of Colorado, this DSS was designed for a general river basin modeling for a wide range of applications (Zagona, 1998). It has three fundamental solution methods: simple simulation, rule-based simulation, and optimization.

To abate the problems of complicated water policies, a different programming language called RiverWare Rule Language (RWRL) is used. Policy descriptions can be designed as a structured ruleset in RWRL. Once these policy descriptions are saved as ruleset files, a simulation may be guided by the ruleset. Furthermore, the policies can be modified between runs, without requiring the simulator to be changed or rebuilt (Wehrends and Reitsma, 1995).

Wehrends and Reitsma (1995) gave the following examples of how water policies can be formulated and interpreted.

```
IF Mead's elevation > 1229.0 THEN
   Mead's release = Mead's inflow
END IF
```

This approach gives a conditional water policy, which may be considered to be easy enough to be incorporated in a general simulation model.

```
IF Mead's elevation > value THEN
   Mead's release = mead's inflow
END IF
```

In this approach, the user has the choice of changing ***value*** at run-time without rebuilding the program. However, the policies expressed in this fashion may be still very specific.

A more comprehensive approach is to allow policies to be completely modifiable without requiring the underlying system to be rebuilt. As such, policies can be written in a rule language that interprets the policies and can be interfaced with the simulation computer programs. The policies are interpreted during run-time, which makes the running time of the program longer.

The general architecture of the RiverWare program employs the representation of a river basin by *objects*. The objects that are included in RiverWare include the following (Zagona et al., 1998):

- *Storage Reservoir.* Mass balance, evaporation, bank storage, spill
- *Level Power Reservoir. Storage Reservoir* plus hydropower, energy, tailwater, operating head
- *Sloped Power Reservoir. Level Power Reservoir* plus wedge storage for very long reservoirs
- *Pumped Storage Reservoir. Level Power Reservoir* plus pumped inflow from another reservoir
- *Reach.* Routing in a river reach, diversion and return flows
- *Aggregate Reach.* Many *Reach* objects aggregated to save space on the workspace
- *Confluence.* Brings together two inflows to a single outflow as in a river confluence
- *Canal.* Bidirectional flow in a canal between two reservoirs
- *Diversion.* Diversion structure with gravity or pumped diversion
- *Water User.* Depletion and return flow from a user of water
- *Aggregate Water User.* Multiple *Water Users* supplied by a diversion from a *Reach* or *Reservoir*
- *Aggregate Delivery Canal.* Generates demand and models supplies to off-line water users
- *Groundwater Storage Object.* Stores water from return flows
- *River Gage.* Specified flows imposed at a river node
- *Thermal Object.* Economics of thermal power system and value of hydropower
- *Data Object.* User-specified data: expression slots or data for policy statements

Table 7.3 shows user methods for selected objects in RiverWare.

AQUATOOL. Developed at the Universidad Politécnica de Valencia (UPV), Spain, as a result of a continuing research over a decade, AQUATOOL is a generalized DSS that has attracted several river basin agencies in Spain (Andreu et al., 1996). Andreu et al. (1996) also note that AQUATOOL has various capabilities that can be used in water resource systems to:

1. Screen design alternatives by means of an optimization module, obtaining criteria about the usefulness and performance of future water resource developments.

TABLE 7.3 Selected User Methods in RiverWare

Object type	User method category	User methods
Reservoirs	Evaporation and precipitation	No evaporation
		Pan and ice evaporation
		Daily evaporation
		Input evaporation
		CRSS evaporation
	Spill	Unregulated spill
		Regulated spill
		Unregulated plus regulated
		Regulated plus bypass
		Unregulated plus regulated plus bypass
Power reservoirs	Power	Plant power
		Unit generator power
		Peak base power
		LCR power
	Tailwater	Tailwater base value only
		Tailwater base value plus lookup table
		Tailwater storage flow lookup table
		Tailwater compare
		Hoover tailwater
Reaches	Routing	No routing
		Time lag routing
		Variable time lag routing
		SSARR
		Muskinghum
		Kinematic wave
		Muskingum-Cunge
		MacCormack
Water user (on Agg-Diversion)	Return flow	Fraction return flow
		Proportional storage
		Variable efficiency

Source: After Zagona et al. (1998).

2. Screen operational management alternatives by means of the optimization module, obtaining criteria from the analysis of the results.
3. Check and refine the screened alternatives by means of a simulation module.
4. Perform sensitivity analysis by comparing the results after changes in the design or in the operating rules.
5. Use different computer programs, once an alternative is implemented, as an aid in the operation of the water resource system, mainly for water allocation among conflicting demands and to study impacts of changes in the system.
6. Perform risk analysis for short- and medium-term operational management to decide, for instance, the appropriate time to apply restrictions and their extent.

AQUATOOL has been accepted by the Sagura and Tagus river basin agencies in Spain as a standard tool to develop their basin hydrologic plan and to manage the resource efficiently in the short to medium term (Andreu et al., 1996).

WaterWare. This DSS is a comprehensive model for integrated river basin planning. It has the capabilities of combining geographic information systems, database technology, modeling techniques, optimization procedures, and expert systems (Jamieson and Fedra, 1996). The aspects of integrated river basin management that this DSS incorporates are briefly as follows (Fedra and Jamieson, 1996):

1. *Groundwater pollution control.* Simulation of flow and contaminant transport, and reduction of the level of contaminant in the aquifer and/or protecting groundwater resources.
2. *Surface-water pollution control.* Estimation of the level of effluent treatment required to meet the river water quality objectives.
3. *Hydrologic processes.* Estimation of ungauged tributary for use in the water resources planning component (see no. 5); assessment of daily water balance for ungauged subcatchments, and the impact of land-use changes on runoff; and evaluation of the effects of conjunctive use of surface and groundwater.
4. *Demand forecasting.* Use of rule-based inference computer programs which use generic expert systems.
5. *Water resources planning.* Consists of
   - ***a.*** A model capable of simulating the dynamics of demand, supply, reservoir operations, and routing through the channel system
   - ***b.*** A module for reservoir site selection which assesses 10 problem classes:
     - (1) Landscape and archaeological or historical sites
     - (2) Land-use restrictions

(3) Drainage, soil, and microclimate
(4) Natural habitats and associated communities
(5) Water quality, aquatic biology and ecology
(6) Water resources and cost implications
(7) Reservoir construction
(8) Reservoir operations
(9) Socioeconomic effects of reservoir operations
(10) Recreational provisions

7.5 STATE OF PRACTICE AND PROSPECTS OF HYDROSYSTEMS COMPUTER PROGRAMS

7.5.1 State of Practice

The concept of integrated water resources management is a comprehensive representation of several components, each of which requires sufficient representation or modeling within the whole system. Modeling needs to be driven by coverage of all aspects of integrated hydrosystems management, not by the convenience or simplicity of the modeling of each aspect of the problem. Loucks (1996) clearly states that "an integrated view of water-resource systems can not be compartmentalized into either surface water or groundwater and either water quantity or water quality just because the respective time and space scales make the modeling or study of such divisions convenient."

On the contrary, as mentioned previously, computer programming generally started out with the simplification of calculations of analytical functions that required a very long time to solve by hand. Through time, the capability enhanced to the level of tackling complex hydrosystems problems. It is through improvements of the programming methodologies and new technological discoveries that more sophisticated hydrosystems computer programs have been developed. Therefore, hydrosystems computer programs have been approaching the essence of integrated hydrosystems management from the bottom up.

The important aspects of integrated hydrosystems problems which have been tackled using computer programs include simulation, database management, data collection, and storage. These efforts have reached a level of promising prospect and have diminished the gap between the concept of computer programs for integrated hydrosystems management and the actual programs. For instance, GIS generally provides facilities for storage and management of very large geo-information. It has been possible to represent the terrain of the entire United States as a database of the Digital Elevation Model (DEM). Automatic data collection systems such as SCADA and radar provide readily available input data for real-time analysis of integrated hydrosystems problems. By integrating together different computer programs, it has been possible to develop DSSs that have manifested to address these issues. A few of these systems have been designed not only to solve the problem, but also to attempt to interpret the result. Jamieson and Fedra (1996) point out that DSSs have the capabilities of predicting what may happen under a particular set of planning assumptions and of providing expert advice on the appropriate course of action.

Most of the available computer programs for hydrosystems problems, however, address only limited issues of the general concept of integrated hydrosystems management. While they may be satisfactory tools to solve the particular problem they are designed for, only a few DSSs currently available such as TERRA, RiverWare, AQUATOOL, and WaterWare are useful as stand-alone computer programs for integrated hydrosystems management. Therefore, it can be inferred that because of the availability of only a limited number of DSSs for integrated hydrosystems management, the state of practice of DSSs for integrated hydrosystems management is premature, yet evolving.

7.5.2 Prospects

Advances in software engineering appear to be promising for integrated hydrosystems management computer programs. They have enabled the development of computer programs that not only incorporate easy-to-use analytical capabilities, but also offer expert advice and intelligent interrogation facilities. With these types of computer programs, the artificial intelligence involved can be provided by a mixture of optimization techniques and expert systems that can evaluate, draw preliminary conclusions, and recommend appropriate actions. This stage of development of hydrosystems computer programs is the emergence of what has been referred to as the fifth generation of the hydroinformatics system (Jamieson and Fedra, 1996).

Even though the efforts made in the past to develop simulation computer programs have been tremendous, many hydrosystems computer programs were written to address specific hydrosystems problems such as reservoir operation, water distribution, urban drainage, and streamflow. The focus of the objectives of many of these computer programs is limited. The parts are out, but we are faced with the painstaking task of integrating these simple computer programs to create a more integrated hydrosystems computer program.

Some promising efforts in this regard have already been undertaken. The successful developments of TERRA, WaterWare, RiverWare, AQUATOOL, and so on are very good examples. The efforts made at the USACE Hydrologic Engineering Center to enhance the old computer programs to the new ones, generally known as the Next Generation (NexGen) computer programs, may form one of the strong cores of DSSs, and simulation computer programs.

Decision support systems in general are, perhaps, the most promising tool to integrate the simple computer programs to be used for integrated hydrosystems management. The three subsystems of DSSs—database management subsystem, model base management subsystem, and dialog generation and management subsystem—constitute a logical construct of the concept of integrated hydrosystems management. Figure 7.9 shows a representation of most of the possible components of a typical DSS. The dotted lines in the figure show the components that can be included in the DSS in the future or used to enhance its current proposed structure.

The database management subsystem provides the opportunity for easy collection, storage, and alteration of data, including on a real-time basis. GIS and SCADA, among others, are important systems for this purpose. The proliferation of simulation computer programs and the availability of some advanced optimization techniques provide valuable resources in dealing with different aspects of hydrosystems problems. The graphics-supported user-friendly interface environment also helps to draw appropriate conclusions and make necessary decisions that agree with predefined integrated hydrosystems management policies.

If there are challenges to overcome to use DSSs for integrated hydrosystems management problems, one of the most difficult challenges, perhaps, will be not having appropriate integrated hydrosystems management policies clearly defined. It may be noted that while it is possible to code policies in a computer program, no code may be written for a policy that does not exist. Likewise, it cannot be easy to write a clear computer code for an ambiguous or ill-defined policy.

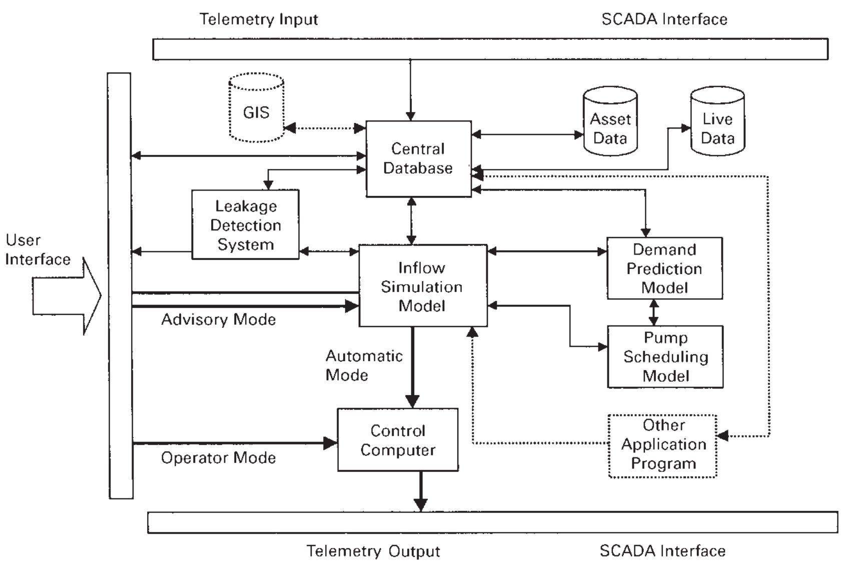


FIGURE 7.9 Proposed water management system for the Great Man-Made River Scheme, Libya. (Khalil, 1998)

7.6 SUMMARY AND CONCLUSIONS

Water as a limited resource poses the challenge of allocating a sufficient amount to all the competing users efficiently and effectively. An integrated hydrosystems management approach enables us to have knowledge in space and time of the purposes and amounts of water requirements, thereby allowing for balancing out between the competing needs. As discussed herein, an integrated hydrosystems management may be the most promising means to provide the water requirements of all the competing users, requiring the involvement of all parties concerned. This approach helps the formulation of viable water policies acceptable to most parties or satisfying most objectives. Nonetheless, although this concept is a strategy that is increasingly advocated in the literature, it is still relatively new. The challenge in this regard is yet to be fully overcome.

Because the concepts of integrated hydrosystems management can be best explained in terms of hydrosystems policies or rules and because such policies can be interpreted and coded in computer programs, it is very important to have these policies clearly defined for a given watershed. It may be noted that it is these policies that we begin with to deal with integrated hydrosystems management. Furthermore, the scope and areal coverage of integrated hydrosystems management that is mandated to an institution or water agency should be unambiguously defined. Implementing its concept would be more complete if the ideals of integrated hydrosystems management are clearly defined and understood, and if the policies can be easily interpreted and coded into computer programs. The authors agree with the watershed approach strategy already recommended by different institutions. This approach entails hydrosystems policies that transcend political boundaries for the purpose of integrated hydrosystems management and, therefore, it is necessary that this approach be acceptable by different parties so that the best overall result is obtained. To this effect, a river basin or watershed approach for regional coverage is a sound strategy.

Lack of efficient techniques in the past that could be used to code hydrosystems policies in computer programs might have had a negative impact on their development for integrated hydrosystems management. The advance in computing technology appears to be at a stage where it is capable of overcoming such problems. Today, a computer programming language specifically used for rulesets (a set of simulation rules) have been developed at CADSWES and therefore can be used for modeling integrated hydrosystems problems, should such languages become the requirement of the state of the art for this purpose.

Computer programs for integrated hydrosystems management can be very important tools for fast computations, easy data management, and drawing conclusions about certain water policies. Such computer programs, generally termed DSSs, have been introduced recently by different institutions. As computing speed and ease become more powerful, more complex yet more comprehensive computer programs are being developed. Such computer programs as CALSIM, TERRA, RiverWare, AQUATOOL, and WaterWare are examples of DSSs that can be used for integrated hydrosystems management. These DSSs are embodied with water policies in the form of rulesets (to use the term used in RiverWare) or expert systems (to use the term used in WaterWare). These computer programs have become successful as programs of integrated hydrosystems management by the incorporation of water policies that are formulated in a form understandable in the computation processes.

At the center of DSSs are found simulation and/or optimization computer programs. A tremendous amount of work has been done in the past to develop simulation and optimization computer programs that solve problems in the areas of hydrology, hydraulics, and water resources. An effort was also made to interface simulation and optimization computer programs to solve optimal control problems in water resources. In the process, water policy issues such as water rights, ecosystem sustainability, and amenity are easily incorporated in the form of rulesets or expert systems. In this regard, more effort is needed because rulesets or expert systems are relatively new approaches in modeling, and the concept of integrated hydrosystems management approach is yet to come to fruition.

In conclusion, some useful computer programs in the form of DSSs, which address integrated hydrosystems management problems, have been written. Several of these programs have proved the importance of the use of DSSs in integrated hydrosystems management problems. The availability of various hydrosystems computer programs that address specific hydrosystems problems and different optimization techniques, in conjunction with the advance in the information technology, provide a wealth of resources that are useful in designing DSSs. Thus, we may conclude that not enough work has been done to develop DSSs for integrated hydrosystems management. However, we have the technical resources—database management systems, simulation computer programs, optimization techniques, and advanced computing technology—and we are faced with the decision to make use of these resources to bring out more DSSs for integrated hydrosystems management.

7.7 REFERENCES

Adeli, H., and S. L. Hung (1995). *Machine Learning—Neural Networks, Genetic Algorithms, and Fuzzy Systems,* John Wiley & Sons, New York.

American Water Works Association Research Foundation (1996). Minutes of Seattle Workshop on Total Water Management, Denver, Colo.

Andreu, J., et al. (1996). AQUATOOL: A Generalized Decision Support System for Water-Resources Planning and Operational Management, *Journal of Hydrology* **177:**269–291.

Brion, L. M., and L. W. Mays (1989). Methodology for Optimal Operation of Pumping Stations in Water Distribution Systems, *Journal of Hydraulic Engineering,* ASCE, **117**(11):1551–1569.

Bulkley, J. W. (1995). Integrated Watershed Management: Past, Present and Future, *Water Resources Update,* Issue No. 100, Universities Council on Water Resources, Carbondale, Ill.

Carriaga, C. C., and L. W. Mays (1995a). Optimization Modeling for Simulation in Alluvial Rivers, *Journal of Water Resources Planning and Management,* ASCE, **121**(3):251–259.

Carriaga, C. C., and L. W. Mays (1995b). Optimal Control Approach for Sedimentation Control in Alluvial Rivers, *Journal of Water Resources Planning and Management,* ASCE, **121**(6):408–417.

Chambers, L. (1995). *Practical Handbook of Genetic Algorithms Applications,* Vol. 1, CRC Press.

Clement, D. P. (1996). SCADA System Using Packet Radios Helps to Lower Cincinnati's Telemetry Costs, *Water Engineering and Management* **134**(8):18–20.

Culver, T. B., and C. A. Shoemaker (1992). Dynamic Optimal Control for Groundwater Remediation with Flexible Management Periods, *Water Resources Research* **28**(3):629–641.

Fedra, K., and D. G. Jamieson (1996). The "WaterWare" Decision Support System for River-Basin Planning. 2. Planning Capability, *Journal of Hydrology* **177**:177–198.

Ford, D. T., and J. R. Killen (1995). PC-Based Decision-Support System for Trinity River, Texas, *Journal of Water Resources Planning and Management* **121**(5):375–381.

Fredericks, J. W., et al. (1998). Decision Support System for Conjunctive Stream-Aquifer Management, *Journal of Water Resources Planning and Management* **124**(2):69–78.

Goldman, F. E. (1998). The Application of Simulated Annealing for Optimal Operation of Water Distribution Systems, Ph.D. dissertation, Arizona State University, Tempe.

Goldman, F. E., and L. W. Mays (2002). Simulated Annealing Approach for Operation of Water Distribution Systems Considering Water Quality, ASCE (in review).

Grigg, N. S. (1998). Coordination: The Key to Integrated Water Management, *Water Resources Update,* Issue No. 111, Universities Council on Water Resources, Carbondale, Ill.

Hall, M. W. (1998). Extending the Resources: Integrating Water Quality Considerations into Water Resources Management, *Water Resources Update,* Issue No. 111, Universities Council on Water Resources, Carbondale, Ill.

Hooper, B. (1995). Towards More Effective Integrated Watershed Management in Australia: Results of a National Survey, and Integrated Implications for Urban Catchment Management, *Water Resources Update,* Issue No. 100, Universities Council on Water Resources, Carbondale, Ill.

Jamieson, D. G., and K. Fedra (1996). The "WaterWare" Decision Support System for River-Basin Planning. 1. Conceptual Design, *Journal of Hydrology* **177**:163–175.

Jeter, M. W. (1986). *Mathematical Programming: An Introduction to Optimization,* Marcel Dekker, New York.

Khalil, H. M. (1998). Proposed Water Management System for the Great Man-Made River Project, in W. R. Blain (ed.), *Hydraulic Engineering Software VII,* Wessex Institute of Technology Press, Boston, pp. 361–379.

Kool, J. B. and M. T. van Genuchten (1991). *HYDRUS: One-Dimensional Variable Saturated Flow and Transport Model, Including Hysteresis and Root Uptake*, U.S. Department of Agriculture, Agriculture Service, Riverside, Calif.

Lasdon, L. S., and A. D. Waren (1986). *GRG2 User's Guide,* Department of General Business, The University of Texas at Austin.

Li, G. L., and L. W. Mays (1995). Differential Dynamic Programming for Estuarine Management, *Water Resources Planning and Management* **121**(6):455–462.

Li, G., Z. Tang, L. W. Mays, and P. Fox (2000). New Methodology for Optimal Operation of Soil Aquifer Treatment Systems, *Water Resources Management* **14**(1):13–33.

Loucks, D. P. (1996). Surface Water Resource Systems, in L. W. Mays (ed.), *Water Resources Handbook,* McGraw-Hill, New York.

Mays, L. W. (1997). *Optimal Control of Hydrosystems,* Marcel Dekker, New York.

Mays, L. W., and Y. K. Tung (1992). *Hydrosystems Engineering and Management,* McGraw-Hill, New York.

Mitchell, B. (ed.) (1990). *Integrated Water Management: International Experiences and Perspectives,* Belhaven Press, London.

Murty, K. G. (1995). *Operations Research: Deterministic Optimization Models,* Prentice Hall, Englewood Cliffs, N.J.

Nicklow, J. W., and L. W. Mays (2000). Optimization of Multiple Reservoir Networks for Sedimentation Control, *Journal of Hydraulic Engineering* **126**(4):232–243.

Nicklow, J. W., and L. W. Mays (2001). Optimal Control of Reservoir Releases to Minimize Sedimentation in Rivers and Reservoirs, *Journal of the American Water Resources Association* **37**(1):197–211.

Reitsma, R. F., et al. (1996). Decision Support Systems (DSS) for Water-Resources Management, in L. W. Mays (ed.), *Water Resources Handbook,* McGraw-Hill, New York.

Rossman, L. A. (1994). EPANET Users Manual, Project Summary Report, Risk Reduction Engineering Laboratory, U. S. Environmental Protection Agency (EPA), Cincinnati, Ohio.

Sakarya, A. B., and L. W. Mays (2000). Optimal Operation of Water Distribution Pumps Considering Water Quality, *Journal of Water Resources Planning and Management* **126**(4):210–220.

Sakarya, A. B., et al. (1998). New Methodologies for Optimal Operation of Water Distribution Systems for Water Quality Purposes, in W. R. Blain (ed.), *Hydraulic Engineering Software VII,* Wessex Institute of Technology Press, Boston, pp. 101–110.

Schultz, G. A. (1998). A Change of Paradigm in Water Sciences at the Turn of the Century? *Water International,* **23**(1):37–44.

Shane, R. M., et al. (1995). The INTEGRAL Project: Overview in Computing in Civil Engineering, in *Proceedings of the Second Congress,* ASCE, June 5–8, Atlanta, Vol. 1, pp. 203–205.

Sprague, R. H., and E. D. Carlson (1982). *Building Effective Decision Support Systems,* Prentice-Hall, Englewood Cliffs, N.J.

Tang, A., and L. W. Mays (1998). Genetic Algorithms for Optimal Operation of Soil Aquifer Treatment Systems, *Water Resources Management,* **12**:375–396.

Topping, B. H. V, et al. (1993). Topological Design of Truss Structures Using Simulated Annealing, in B. H. V. Topping and A. I. Khan (eds.), *Neutral Networks and Combinatorial Optimization in Civil and Structural Engineering,* Civil-Comp Press, Edinburgh, Scotland, pp. 151–165.

U.S. General Accounting Office (1994), Ecosystem Management Additional Actions Needed to Adequately Test a Promising Approach, GAO/RCED-94-111.

Viessman, W., Jr. (1998). Water Policies for the Future: Bringing It All Together, *Water Resources Update,* Issue No. 111, Universities Council on Water Resources, Carbondale, Ill.

Vlachos, E. C. (1998). Practicing Hydrodiplomacy in the 21st Century, *Water Resources Update,* Issue No. 111, Universities Council on Water Resources, Carbondale, Ill.

Wada, R. N., et al. (1986). Honolulu's New SCADA System, *Journal of the American Water Works Association* **78**(8):43–48.

Wanakule, N., et al. (1986). Optimal Management of Large Scale Aquifers: Methodology and Applications, *Water Resources Research* **22**(4):447–465.

Wehrends, S. C., and R. F. Reitsma (1995). A Rule Language to Express Policy in a River Basin Simulator in Computing in Civil Engineering, in *Proceedings of the Second Congress,* ASCE, June 5–8, Atlanta, Vol. 1, pp. 392–395.

Wurbs, R. A. (1995). *Water Management Modelss: A Guide to Software,* Prentice-Hall PRT, Englewood Cliffs, N.J.

Zagona, E. A. (1995). The INTEGRAL Project: The PRYSM Reservoir Scheduling and Planning Tool in Computing in Civil Engineering, in *Proceedings of the Second Congress,* ASCE, June 5–8, Atlanta, Vol. 1.

Zagona, E. A. (1998). RiverWare: A General River and Reservoir Modeling Environment, in *Proceedings of the First Federal Interagency Hydrologic Modeling Conference,* April 19–23, Las Vegas.

CHAPTER 8
REGIONAL WATER SUPPLY PLANNING AND CAPACITY EXPANSION MODELS

Messele Z. Ejeta
California Department of Water Resources
Sacramento, California

Larry W. Mays
Department of Civil and Environmental Engineering
Arizona State University
Tempe, Arizona

8.1 INTRODUCTION

Water supply planning on a regional scale has attracted more interest in recent years. The major reason for this is that any effect on the quantity and quality of the water resources at a certain location in a region has a bearing on the quantity and quality of this resource at other locations in the region. To meet water supply needs in a given region, it becomes necessary to obtain water from different sources in the region or outside it. This entails not only blending different quality waters but also maintaining the quality of the water distributed within accepted standards, which is an important planning aspect for water supply agencies. Yeh et al. (2000) state that in a multisource, multiquality regional water distribution system, water agencies often find it necessary to impose blending requirements at certain control points in the system in order to secure the desired water quality downstream of the control points.

The desire for efficient water resources management has led to the development of various models for optimal management of water supply systems. Some such models that have been developed in the past consist of nonlinear programming (NLP) models with different sets of objective functions and constraints. These include combinations of nonlinear objective functions and constraints, nonlinear objective functions and a mix of linear and nonlinear constraints, nonlinear objective functions and linear constraints and so on. Ocanas and Mays (1980, 1981a, 1981b) developed a water reuse planning NLP model for determining the optimum allocation of water and reuse of wastewater on a regional basis for multiperiod planning, which was solved by using the large-scale generalized reduced gradient and successive linear programming methods. Schwartz and Mays (1983) also developed water reuse and wastewater treatment alternative planning models using a dynamic programming approach. Mehrez et al. (1992) developed a model for real-time control and operation of a multisource regional system with multiple customers, which involved both quality and quantity constraints.

A capacity expansion model for water energy systems was developed by Lall and Mays (1981) as a multisystem, multiperiod capacity expansion problem, which was solved using Bender's decomposition with a specialized algorithm. Matsumoto and Mays (1983) also developed a capacity expansion model for large-scale water energy systems as a multisystem, multiperiod capacity expansion problem. Bender's decomposition is applied to the original problem to decompose it into three subproblems: capacity, production, and distribution. Each subproblem was solved using a specialized algorithm (Matsumoto and Mays, 1983). Kim and Mays (1994) formulated the rehabilitation and replacement of water distribution system components as a mixed integer nonlinear programming (MINLP) problem and solved it by decomposing the MINLP problem into a mixed integer linear programming (MILP) master problem and an NLP subproblem. The NLP subproblem is solved using the generalized reduced gradient (GRG2) (Lasdon and Waren, 1986) computer code, and the MILP master problem is solved using ZOOM, a 0/1 mixed integer programming code.

Herein, two NLP models for regional water supply planning and an MINLP model for capacity expansion of water supply infrastructure are discussed. The NLP models are formulated with nonlinear objective functions to maximize net benefits. The first NLP model is a yearly static model, whereas the second one is a seasonal dynamic model. The revenue from water supplies from different sources, the cost of supplying the water from these sources, and the damage due to poor-quality water being delivered to users are considered in the objective functions of these models. These NLP models are applied to a regional water supply system located along the Rio Grande from Caballo Dam in New Mexico to El Paso County, Texas. The solution obtained for the seasonal model gives the "best" reservoir release policy for four seasons of a year and a maximum net benefit that is slightly higher than that obtained using the static NLP model. Furthermore, compared to the static model, this model generally maintained better-quality water in terms of total dissolved solids (TDS).

The MINLP model for the capacity expansion problem is formulated for water supply conveyance and treatment infrastructure over a long-term planning period. This model incorporates two sets of decision variables: (1) the optimum timing of the capacity expansion of water supply conveyance systems and water treatment plants and (2) the water allocation policy that maximizes the net benefit. The solution methodology interfaces a simulated annealing (SA) heuristic search algorithm with the GRG2 nonlinear optimization code. The model is applied to El Paso County's water supply system assuming three water supply regions, two potential water sources (surface water and groundwater) for each region, and a planning period of 10 years. The optimum capacity expansion schedules for the water supply infrastructure components for optimum water supply allocations are given.

8.2 MODEL FORMULATIONS

8.2.1 Regional Water Supply Models

A general problem on a regional (watershed) scale is considered in which all the water users in the region are first identified. The general water supply customers considered include municipality, industry, hydropower, irrigation, ranching, recreation, aquaculture, environmental protection, and remediation, whereas general water sources considered include in-stream flow, groundwater pumping, precipitation, interbasin transfer, and return flow or drainage. The various equations used in the models are discussed in the following subsections.

Continuity Equations. A watershed is divided into a total of I reaches where the delineation of a reach is based on whether there is one or more diversions and/or one or more return flows at only one point in the reach (see Fig. 8.1). In each reach, a total of J activities are considered. This approach enables one to model the entire watershed on a reach-by-reach approach.

To develop the continuity equations, which are, in general, mass conservation equations, an arbitrary reach i in Fig. 8.1 is chosen. In Fig. 8.1, all the possible flows associated with reach i are indicated. To derive the first continuity equation, the point of diversion, or return, in reach i is considered. The equation is expressed as

$$\sum_{k=(i-n)}^{i-1} \sum_{j=1}^{J} (Q_{r_{kji}} - Q_{s_{kji}}) + Q_{ds_i} - \sum_{j=1}^{J} Q_{d_{ij}} - Q_{us\,(i+1)} = 0 \qquad (8.1)$$

where $Q_{r_{kji}}$ = return flow from demand point j in reach k to reach i ($k \neq i$)
$Q_{s_{kji}}$ = seepage loss in return flow from demand point j in reach k to reach i ($k \neq i$)

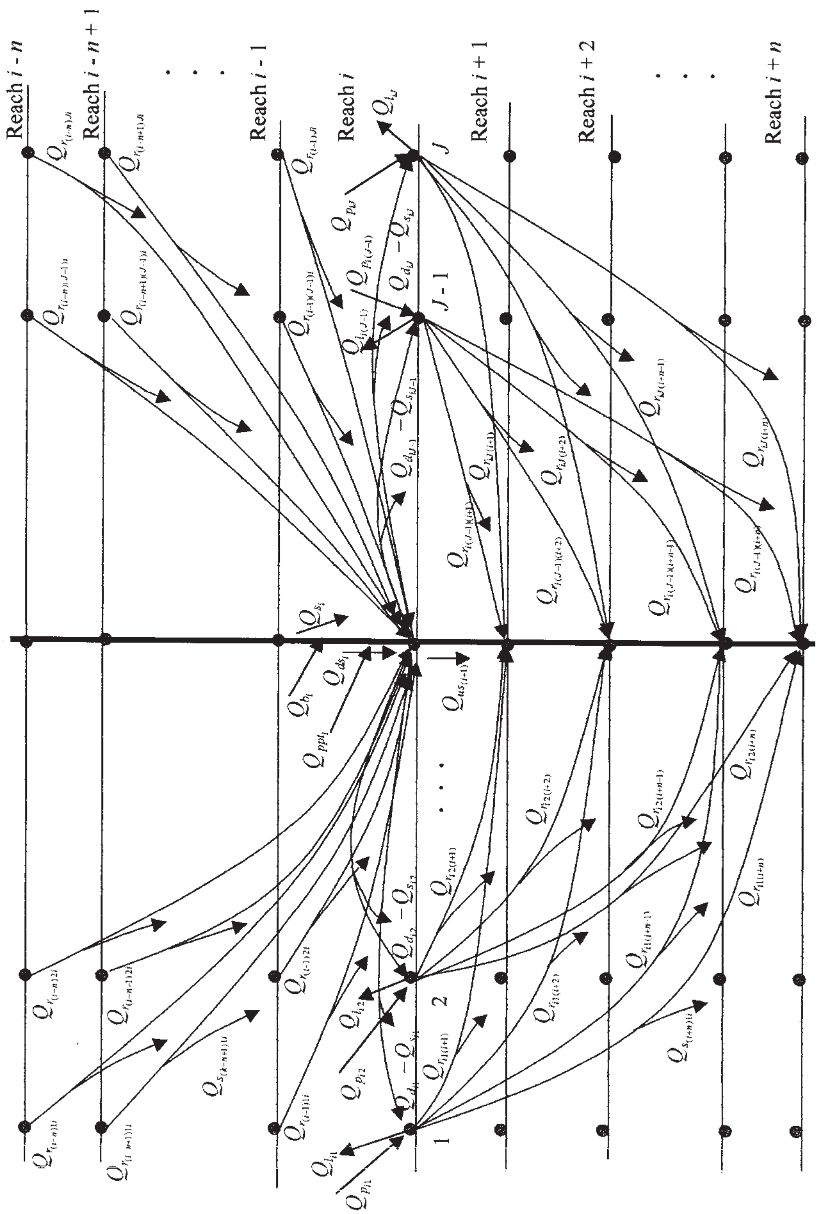


FIGURE 8.1 Schematic representation of the hydraulics of a typical watershed.

$Q_{d_{ij}}$ = diverted flow from reach i for demand point j, where j represents any kind of demand for water ($i = 1, 2, \ldots, I; j = 1, 2, \ldots, J$)
Q_{ds_i} = flow at downstream end of reach i, that is, in-stream flow into point of diversion in reach i, referred to herein as a *node*
Q_{us_i} = flow at upstream end of reach i
n = maximum number of reaches a return flow travels before draining into the main stream

The same equation holds true for the entire reaches with boundary conditions for the return flows near the upstream and the downstream ends of the watershed. Also, for each demand point, the following continuity equations can be derived.

$$(Q_{d_{ij}} - Q_{s_{ij}}) + Q_{p_{ij}} - Q_{l_{ij}} - \sum_{k=(i+1)}^{i+n} Q_{r_{ijk}} = 0 \qquad \text{for all } i, j \tag{8.2}$$

where $Q_{s_{ij}}$ = seepage loss in reach i on way to demand point j
$Q_{p_{ij}}$ = pumped flow from reach i for demand point j
$Q_{l_{ij}}$ = consumptive use in reach i at demand point j

Furthermore, within each reach, the following continuity equation is written.

$$Q_{us_i} + Q_{b_i} + Q_{ppt_i} - Q_{ds_i} - Q_{s_i} = 0 \qquad \text{for all } i \tag{8.3}$$

where Q_{b_i} = interbasin transfer to reach i
Q_{ppt_i} = flow into reach i accountable to net precipitation contribution
Q_{s_i} = in-stream seepage loss in the stream in reach i

Water Quality Equations. A typical quality parameter for water distribution on a watershed scale is the amount of TDS in the water supplied. This water quality parameter is taken as an example to develop the equations for water quality management as part of water supply system planning on a watershed scale. Generally, however, multiple water quality parameters may be considered since the quality equations developed will be similar.

The constraints for the point of diversion in reach i can be expressed as

$$\sum_{k=(i-n)}^{i=1} \sum_{j=1}^{J} C_{r_{kji}}(Q_{r_{kji}} - Q_{s_{kji}}) + C_{ds_i} Q_{ds_i} - C_{w_i}\left(\sum_{j=1}^{J} Q_{d_{ij}} + Q_{us\ (i+1)}\right) = 0 \tag{8.4}$$

where Q_{ij} = flow from node i to point j
C_{ij} = quality parameter in Q_{ij}
C_w = weighted quality parameter in Q_{ij}

The flow out of the nodes will have the same quality assuming "perfect" mixing conditions (Yang et al., 1999). Thus, the following equalities also hold true at the diversion point in reach *i*.

$$C_{w_i} = C_{d_{ij}} = C_{us_{(i+1)}} \tag{8.5}$$

Defining $C_{w_{ij}}$ as the weighted TDS concentration at any demand point *j* in reach *i*, the constraints for demand point *j* can be expressed as

$$C_{d_{ij}}(Q_{d_{ij}} - Q_{s_{ij}}) + C_{p_{ij}} Q_{p_{ij}} - C_{w_{ij}} \left(Q_{l_{ij}} + \sum_{k=(i+1)}^{i+n} Q_{r_{ijk}} \right) = 0 \qquad \text{for all } i, j \tag{8.6}$$

Similar equations can be also derived for the in-stream flow in each reach as

$$C_{us_i} Q_{us_i} + C_{b_i} Q_{b_i} + C_{ppt_i} Q_{ppt_i} - C_{ds_i} Q_{ds_i} - \frac{C_{us_i} + C_{ds_i}}{2} Q_{s_i} = 0 \qquad \text{for all } i \tag{8.7}$$

The average TDS concentration is used for the seepage loss in reach *i* assuming that the change in the concentration in the reach is approximately linear.

Flow Regulation. In Fig. 8.1, the flow system represents the general scenario in which diversions and pumpages in any reach and return flows from any reach to any other reach downstream are possible. However, this may not be the case in practice since, for instance, the number of demand points in one reach may not be equal to those in other reaches. Introducing coefficients with values of 0 or 1 for each of the arcs can regulate these flows, where 0 is used if there is no flow along the arc and 1 is used if there is a flow along the arc. The only exceptions that do not require the 0 or 1 coefficients are the in-stream flows $Q_{u_{si}}$ and $Q_{d_{si}}$, since the coefficient of these flows is always 1. The use of such coefficients has advantages in sensitivity analysis to the optimization model developed, which can be done by turning on and off all uncertain activities.

Let $K_{d_{ij}}$ be the coefficient of flow for $Q_{d_{ij}}$, $K_{p_{ij}}$ be the coefficient of flow for $Q_{p_{ij}}$ and so on. Then, the above continuity and TDS equations are modified by introducing these coefficients.

$$\sum_{k=(i-n)}^{i=1} \sum_{j=1}^{J} (K_{r_{kji}} Q_{r_{kji}} - K_{s_{kji}} + Q_{s_{kji}}) + Q_{ds_i} - \sum_{j=1}^{J} K_{d_{ij}} Q_{d_{ij}} - Q_{us\,(i+1)} = 0 \qquad \text{for all } i \tag{8.8}$$

$$(K_{d_{ij}} Q_{d_{ij}} - K_{s_{ij}} Q_{s_{ij}}) + K_{p_{ij}} Q_{p_{ij}} - K_{l_{ij}} Q_{l_{ij}} - \sum_{k=(i+1)}^{i+n} K_{r_{ijk}} Q_{r_{ijk}} = 0 \qquad \text{for all } i, j \tag{8.9}$$

$$Q_{us_i} + K_{b_i}Q_{b_i} + K_{ppt_i}Q_{ppt_i} - Q_{ds_i} - K_{s_i}Q_{s_i} = 0 \qquad \text{for all } i \tag{8.10}$$

$$\sum_{k=(i-n)}^{i-1} \sum_{j=1}^{J} C_{r_{kji}}(K_{r_{kji}}Q_{r_{kji}} - K_{s_{kji}}Q_{s_{kji}}) + C_{ds_i}Q_{ds_i} - C_{w_i}\left(\sum_{j=1}^{J} K_{d_{ij}}Q_{d_{ij}} + Q_{us\,(i+1)}\right) = 0 \qquad \text{for all } i \tag{8.11}$$

$$C_{d_{ij}}(K_{d_{ij}}Q_{d_{ij}} - K_{s_{ij}}Q_{s_{ij}}) + C_{d_{ij}}(K_{p_{ij}}Q_{p_{ij}}) - C_{w_{ij}}(K_{l_{ij}}Q_{l_{ij}} + \sum_{k=(i+1)}^{i+n} K_{r_{ijk}}Q_{r_{ijk}}) = 0 \qquad \text{for all } i, j \tag{8.12}$$

$$C_{us_i}Q_{us_i} + C_{b_i}(K_{b_i}Q_{b_i}) + C_{ppt_i}(K_{ppt_i}Q_{ppt_i}) - C_{ds_i}Q_{ds_i} - \frac{C_{us_i} + C_{ds_i}}{2} K_{s_i}Q_{s_i} = 0 \qquad \text{for all } i \tag{8.13}$$

$$C_{w_i} - C_{d_{ij}} = 0 \qquad \text{for all } i, j \tag{8.14}$$

$$C_{w_i} - C_{us\,(i+1)} = 0 \qquad \text{for all } i \tag{8.15}$$

$$C_{w\,(i-1)} - C_{ds_i} = 0 \qquad \text{for all } i \tag{8.16}$$

Additional Constraints. Apart from the mass balance constraints, some other physical and resource constraints may be prevalent. The capital expenditure available for pumping the water and the amount that can be pumped may be limited. Environmental regulations may require that the TDS level at any point be kept below some acceptable level. Such constraints generally can be given as follows.

$$K_{p_{ij}}Q_{p_{ij}} \le A \tag{8.17}$$

$$C(K_{p_{ij}}Q_{p_{ij}}) \le B \tag{8.18}$$

$$K_{w_{ij}}C_{w_{ij}} \le C \tag{8.19}$$

where A = constant
B = constant or function of amount of pumped water
C = constant or function of quality parameter in supplied water
$C\,(K_{p_{ij}}Q_{p_{ij}})$ = cost of any pumping at demand point j in reach i

Objective Function. The objective function can be expressed in terms of the profits from the water resources allocated at each of the demand points, the costs associated with each allocation, and the damages caused due to an undesirable water quality level (such as unacceptable TDS level) at each user point. Defining $P(Q_{d_{ij}}, Q_{p_{ij}})$ as the profit as a function of the allocated supplies (both surface water and groundwater), $C(Q_{d_{ij}}, Q_{p_{ij}})$ as the cost of allocation, and $D(C_{d_{ij}}, C_{p_{ij}}) = D(C_{w_{ij}})$ as the damage, the objective is to maximize the net profit Z given as

$$Z = P(Q_{d_{ij}}, Q_{p_{ij}}) - C(Q_{d_{ij}}, Q_{p_{ij}}) - D(C_{w_{ij}}) \tag{8.20}$$

subject to constraint equations (8.8) to (8.19).

8.2.2 Capacity Expansion Model

Modeling Approach. In this subsection, a water supply system capacity expansion problem is formulated as an MINLP model. Costs for construction, operation, and expansion of the infrastructure are considered as continuous functions of water demands, and the times of expansion are modeled using 0 or 1 integer variables. Thus, for any given unit of the infrastructure, there exists $P - 2$ possible alternatives of expansion of the unit in any P time units of a planning period, such as 20 or 30 years. No expansion during the first or the last time unit in the planning period may be practically feasible, and, thus, we obtain $P - 2 + 1 = P - 1$ possible alternatives for P time units of a planning period. One of the alternatives that doesn't involve any expansion is to initially construct a capacity that is sufficient for the demand during the last time unit of the planning period. Figure 8.2 shows a pictorial representation of such alternatives for any single unit. The objective is therefore to determine the size of the water supply infrastructures and/or when to expand them within the time framework of the planning period at a minimum cost.

The number of possible combinations of construction times grows exponentially with the number of the components to be expanded. In general, for c components that are expandable during any one time unit from a total of p time units, the total possible alternatives will be p^c. For instance, for a certain water supply system with four supply ends, with each end requiring a supply line and a treatment facility, the total number of possible combinations for a planning period of 50 years becomes $(50 - 1)^8 \approx 3.32 \times 10^{13}$. While a problem with a small number of combinations could be solved by the total enumeration approach, a problem of this size is practically not feasible to solve by this approach. Alternative approaches involving some heuristic search techniques may be employed to obtain a reasonable optimum solution.

For a mathematical formulation, the water supply system is divided into I reaches with J activities (diversion, pumpage, reservoir, etc.) in each reach (see Fig. 8.3). The optimal capacity expansion problem is formulated, in general, as a problem to minimize the sum of costs including construction costs, operation costs, and capacity expansion costs subject to constraints on water demand, capacity, mass conservation, and capital availability. The objective function is to minimize the sum of construction, operation, and expansion costs of water conveyance and treatment systems.

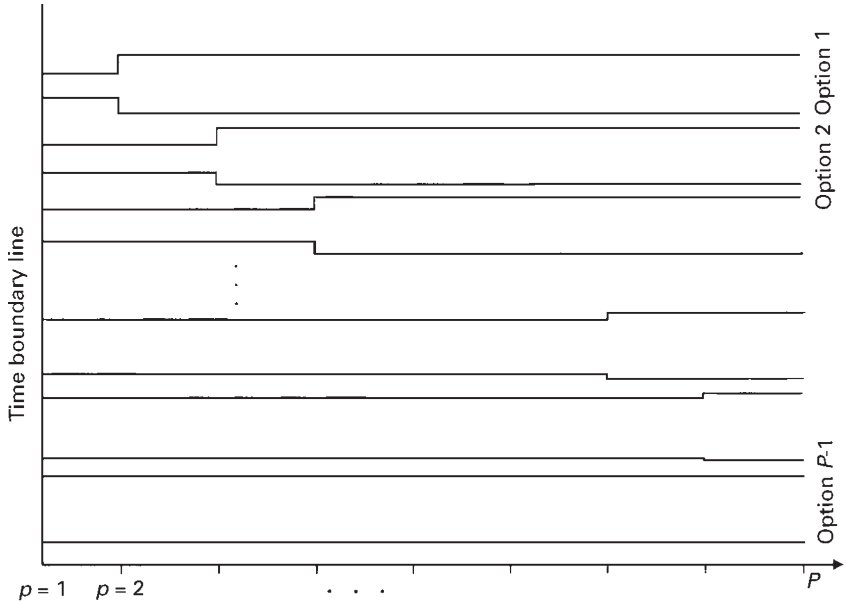


FIGURE 8.2 Different alternative times of capacity expansion of a single water supply unit.

Construction Costs. The total construction costs for water supply lines for all activity points j in all reaches i for a supply quantity of Q can be written as

$$\mathrm{CC}_Q = \sum_{i=1}^{I} \sum_{j=1}^{J} \left[\sum_{p=2}^{P-1} X_{ijp} f_c(Q_{ijp}) + X_{ijP} f_c(Q_{ijP}) \right] \tag{8.21}$$

where CC_Q = total cost of construction of supply lines
Q_{ijp} = supply to activity point j in reach i during time p of planning period
$f_c(Q_{ijp})$ = construction cost function for supply line of Q_{ijp}
X_{ijp} = 0 or 1 integer associated with expansion time of capacity of supply line to activity point j in reach i during time p of planning period.

Similarly, the total construction cost for all activity points j in all reaches i can be written as

$$\mathrm{CC}_{\mathrm{CAP}} = \sum_{i=1}^{I} \sum_{j=1}^{J} \left[\sum_{p=2}^{P-1} Y_{ijp} f_c(\mathrm{CAP}_{ijp}) + Y_{ijP} f_c(\mathrm{CAP}_{ijP}) \right] \tag{8.22}$$

where $\mathrm{CC}_{\mathrm{CAP}}$ = cost of construction of treatment facilities
CAP_{ijp} = capacity of treatment facility at activity point j in reach i during time p of planning period
$f_c(\mathrm{CAP}_{ijp})$ = construction cost for water treatment facility of capacity CAP_{ijp}

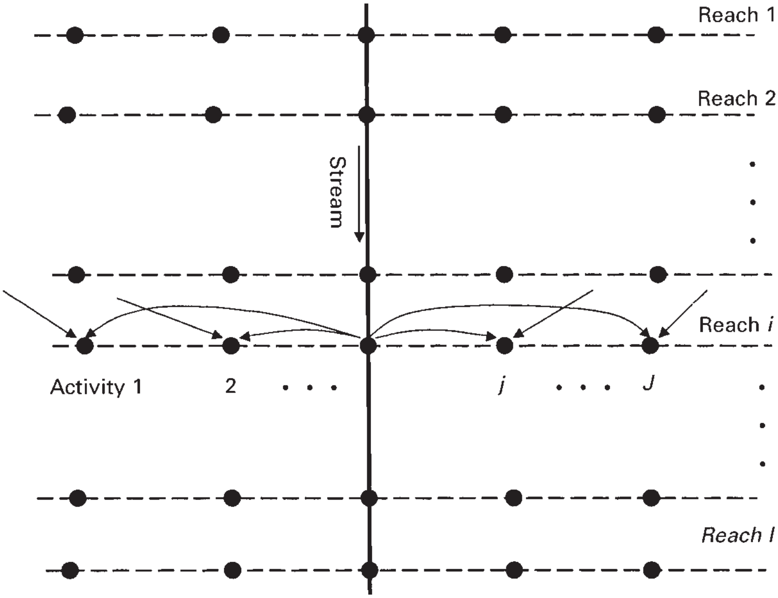


FIGURE 8.3 A water supply system divided into I reaches and J activities with both surface and groundwater supplies.

Y_{ijp} = 0 or 1 integer associated with expansion time of capacity of water treatment facility at activity point j in reach i during time p of planning period.

In both Eqs. (8.21) and (8.22), the last alternative, which is the "expansion" during time P, is written outside the summation over time to indicate that this case is practically different from the other cases in the sense that there is no capacity expansion involved if this alternative is selected. For such a case, the unit or facility that is constructed at the beginning of the planning period will be sufficient for all the time units in the planning period, that is, no capacity expansion is involved (see Fig. 8.2).

Operation Costs. The total operation cost function can be written for both supply lines and water treatment facilities as

$$\mathrm{OC}_Q = \sum_{i=1}^{I} \sum_{j=1}^{J} \left\{ \sum_{p=2}^{P} X_{ijp} \left[f_o(Q_{ij\,(p-1)}) \left(\frac{(1+i)^{p-1} - 1}{i\,(1+i)^{p-1}} \right) + f_o(Q_{ijP}) \left(\frac{(1+i)^{P-p} - 1}{i\,(1+i)^{P}} \right) \right] \right\} \tag{8.23}$$

$$\mathrm{OC_{CAP}} = \sum_{i=1}^{I} \sum_{j=1}^{J} \left\{ \sum_{p=2}^{P} Y_{ijp} \left[f_o(\mathrm{CAP}_{ij\,(p-1)}) \left(\frac{(1+i)^{p-1} - 1}{i\,(1+i)^{p-1}} \right) + f_o(\mathrm{CAP}_{ijP}) \left(\frac{(1+i)^{P-p} - 1}{i\,(1+i)^{P}} \right) \right] \right\} \tag{8.24}$$

where OC_Q = total cost of operation of supply lines
$\mathrm{OC_{CAP}}$ = total cost of operation of treatment facilities
f_o = annual operation function
$\left(\frac{(1+i)^{p-1}}{i\,(1+\mathrm{i})^{p-1}} \right)$ and $\left(\frac{(1+i)^{P-p} - 1}{i\,(1+i)^{P}} \right)$
= discount factors to present-worth values

In both Eqs. (8.23) and (8.24), the first term in the sum over time is taken up to time $p-1$ to avoid any overlap between the operation of the old and the expanded units. In other words, if the expanded unit is ready for time unit p, the operation cost during this time unit should be the operation cost of the expanded unit, not that of the old unit; the operation of the old unit must have ceased at the end of time unit $p-1$.

Expansion Costs. The total expansion cost function can be written for both supply lines and water treatment facilities, respectively, as

$$\mathrm{EC}_Q = \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{p=2}^{P-1} X_{ijp} \frac{f_e(Q_{ijP} - Q_{ijp})}{(1+i)^{p}} \tag{8.25}$$

$$\mathrm{EC_{CAP}} = \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{p=2}^{P-1} Y_{ijp} \frac{f_e(\mathrm{CAP}_{ijP} - \mathrm{CAP}_{ijp})}{(1+i)^{p}} \tag{8.26}$$

where EC_Q = total cost of expansion of supply lines
$\mathrm{EC_{CAP}}$ = total cost of expansion of treatment facilities
$\frac{f_e(Q_{ijP} - Q_{ijp})}{(1+i)^{p}}$
= unit expansion cost discounted to present-worth value

Constraints. The constraints include water demand constraints, capacity constraints, mass conservation constraints, and capital availability constraints. For all *i, j,* and *p,* the water demand constraints and the capacity constraints can be given as

$$Q_{ijp} \geq Q_{ijp}^{\min} \tag{8.27}$$

and

$$\mathrm{CAP}_{ijp} \geq \mathrm{CAP}_{ijp}^{\min} \tag{8.28}$$

where $Q_{ijp}^{\min}$ is the minimum supply required for activity point j in reach i during time p of the planning period and $\mathrm{CAP}_{ijp}^{\min}$ is the capacity required for minimum supply requirement at activity point j in reach i during time p of the planning period. The mass conservation constraints can be given as

$$\sum_{j=1}^{J} \sum_{i=1}^{I} Q_{ijp} \leq Q_p^{\mathrm{tot}} \tag{8.29}$$

for all p, where Q_p^{tot} is the total available water supply during time p of the planning period.

Defining TCC, TOC, and TEC, as the total capital available for construction, operation, and expansion, respectively, the capital availability constraints can be written for each category as

$$\mathrm{CC}_Q + \mathrm{CC}_{\mathrm{CAP}} \leq \mathrm{TCC} \tag{8.30}$$

$$\mathrm{OC}_Q + \mathrm{OC}_{\mathrm{CAP}} \leq \mathrm{TOC} \tag{8.31}$$

$$\mathrm{EC}_Q + \mathrm{EC}_{\mathrm{CAP}} \leq \mathrm{TEC} \tag{8.32}$$

In Eqs. (8.30) to (8.32), each of the cost terms may be substituted by the right-hand side expressions of Eqs. (8.21) to (8.26).

In addition to the above constraints, each of the integer variables for each activity and in each reach must add to 1 since any unit is expanded once. Thus,

$$\sum_{p=2}^{P} X_{ijp} = 1 \tag{8.33}$$

and

$$\sum_{p=2}^{P} Y_{ijp} = 1 \tag{8.34}$$

for all i and j.

Therefore, the general mathematical formulation of the MINLP capacity expansion problem can be formulated as follows. Minimize:

$$\mathrm{CC}_{\mathrm{Q}} + \mathrm{CC}_{\mathrm{CAP}} + \mathrm{OC}_Q + \mathrm{OC}_{\mathrm{CAP}} + \mathrm{EC}_Q + \mathrm{EC}_{\mathrm{CAP}} \tag{8.35}$$

subject to constraint equations (8.27) to (8.34).

8.3 APPLICATIONS TO THE RIO GRANDE PROJECT AND THE CITY OF EL PASO WATER SUPPLY

8.3.1 Overview of the Rio Grande Project

The Rio Grande project studied stretches from Caballo Dam in New Mexico to the United States–Mexico boundary below El Paso, Texas, on the Rio Grande (see Fig. 8.4). The U.S. Bureau of Reclamation completed the construction of the project in 1916. The primary source of water is the San Juan Mountains of southern Colorado that drain to the Rio Grande, which flows south through New Mexico and Texas to the Gulf of Mexico (Engineering-Science, Inc., 1991). The project's potential water demands include

1. Municipal water supplies for Hatch, Las Cruces, Anthony, and El Paso
2. Irrigation water supplies for the irrigation areas at Percha, Leasburg, Mesilla, and El Paso District No. 1, which consists of two locations
3. Industrial water supplies for Las Cruces and El Paso
4. A 60,000-acre·ft annual water supply to Ciudad Juarez, a city in Mexico, under an international treaty

The competition for the demand of the Rio Grande project water is increasing to the extent where the total water of the project will fall short of the requirements of the different parties, which will have practically little alternative sources. The irrigation districts are currently the primary beneficiaries of the Rio Grande project water. Return flows from drainage from water diverted upstream increases the salinity of the streamflow downstream thereby increasing the damage from irrigation, municipal, and industrial water supplies downstream due to poor water quality. Regulations by the U.S. Environmental Protection Agency (EPA) and local agencies require that the TDS level be maintained below 1000 parts per million (ppm).

The sources of project water are the Caballo Reservoir, which supplies an annual average of 790,000 acre·ft of water, and groundwater pumping at several locations. The growth of the population of the cities and towns and the maximum limit regulation on the amount of groundwater pumped pose a serious water shortage problem. In addition, the delivery of poor-quality water is also a problem since a portion of the water supply returns to the river, carrying high levels of TDS that affect the water quality of the mainstream flow system downstream. The drainage water first filtrates into the subsurface and then flows to the mainstream with a fairly uniform TDS content approximated at about 1500 ppm.

The five diversion dams for irrigation purposes include Percha, Leasburg, Mesilla, American, and Riverside. The first three diversion dams serve the Elephant Butte Irrigation District (EBID) which consists of the Percha, Leasburg, and Mesilla irrigation areas. The last two diversion dams serve the El Paso County

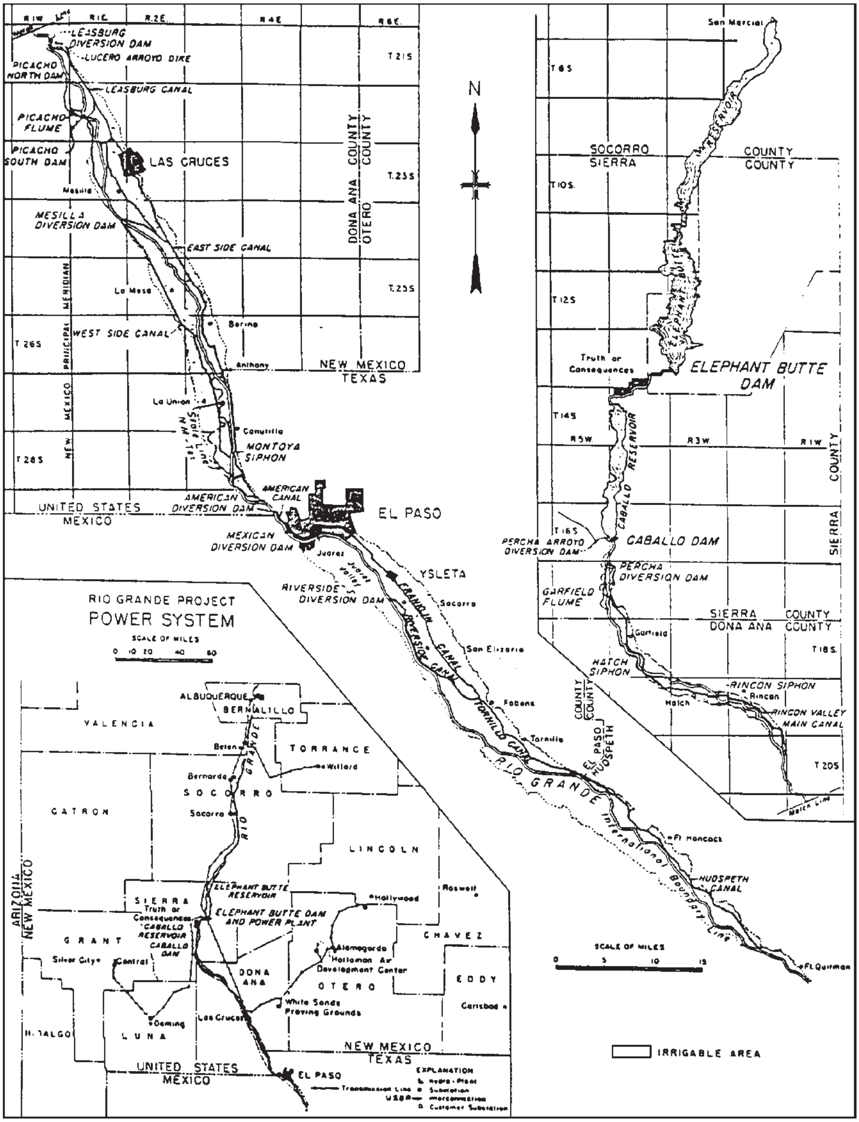


FIGURE 8.4 The Rio Grande project. (Adapted from USBR project data, 1981)

Water Improvement District (EPCWID) No. 1. Percha Diversion Dam is located 2 mi downstream of Caballo Dam and diverts water to the Rincon Valley Main Canal to serve irrigation in the Rincon Valley. Leasburg Diversion Dam is located 62 mi north of El Paso at the head of the Mesilla Valley and diverts water into the Leasburg Canal to serve irrigation in the upper Mesilla Valley. Mesilla Diversion Dam is located 40 mi north of El Paso at the head of the Mesilla Valley and diverts water into the East Side Canal and West Side Canal for irrigation in the lower Mesilla Valley. The American Diversion Dam is located 2 mi northwest of El Paso, immediately above the location where the river becomes the international boundary line between the United States and Mexico. It diverts water into the American Canal for irrigation in the upper portion of the El Paso Valley. The Riverside Diversion Dam is located 15 mi southeast of El Paso and diverts water into the Riverside Canal to serve irrigation in the lower portion of the El Paso Valley.

The current number of estimated households in El Paso is 120,000 whereas that of Hatch, Las Cruces, and Anthony are 1000, 19,000 and 2000, respectively. Hatch, Las Cruces, and El Paso are located in the Rincon, Mesilla, and El Paso valleys, on the opposite sides of the irrigation areas of Percha, Leasburg, and El Paso, respectively. Anthony is located in the Mesilla Valley, below the Mesilla irrigation area.

Several studies have been completed on the Rio Grande project by different agencies and consulting firms, including those by Engineering-Science, Inc. (1991) and jointly by Boyle Engineering Corp. and Parsons Engineering Science, Inc. (1998). The general approach utilized in these studies was the identification and assessment of different management alternatives for the project using available data so that the alternative that resulted in the best management would be selected.

The study by Engineering-Science, Inc. (1991) assessed various opportunities for water storage facilities and the possibilities of additional water supplies by transmountain diversion and by salvage of present losses so that water shortages in the project region would be overcome. The study by Boyle Engineering Corp. and Parsons Engineering Science, Inc. (1998) was aimed at drain mitigation strategies so that acceptable water quality would be maintained for the downstream users in the project area. Based on the results of the simulation study, almost all the selected strategies reduced the problem of quality issues to some extent. However, none of these studies considered the formulation of the problem as a single optimal problem for which the best solution may be obtained.

8.3.2 The City of El Paso Water Supply System

The city of El Paso, Texas, is located along the Rio Grande at the boundary with Mexico (see Fig. 8.4). Most of the population in El Paso County is located in El Paso, which is one of the fastest-growing cities in the nation. The population of El Paso is estimated to become 1.16 million by 2040, whereas that of El Paso County is estimated to become 1.37 million (Boyle Engineering Corporation, 1991), as compared to the 1990 estimates of 529,723 and 752,188, respectively. By 2040, El Paso County's municipal and industrial water requirements are expected to equal 300,000 acre·ft per year, which corresponds to an average consumption rate of 196 gallons per capita per day (gcd). The city is facing a very high water supply shortage because of the apparent increase in demand with time and because its current primary source of groundwater supply is being withdrawn to its potential. It has a limited share of the water of the Rio Grande project. In 1990, the city owned approximately 2000 acres per year of water rights land from the Rio Grande project (Boyle Engineering Corporation, 1991).

The county's water supply system consists of seven major supply regions: Northwest, Northeast, Central, Lower Valley, East, Fort Bliss, and Hueco. Figure 8.5 shows these seven planning areas, and Table 8.1 gives the population projections for these planning areas for the time from 1980 to 2040, in 10-year intervals. Potential water supply sources and the potential quantity that can be obtained from each of these sources, which include alternative surface water supplies, alternative groundwater supplies, and other alternative sources, have been identified (Boyle Engineering Corporation, 1991). Boyle Engineering Corporation (1991) recommended attempting to overcome future water shortage problems by a more balanced approach of conserving and utilizing both surface water and groundwater supplies to meet the demands.

8.3.3 Static NLP Model Application

The NLP model developed is applied to the Rio Grande project by considering yearly and seasonal time frames. Groundwater and Caballo reservoir's average annual storage of 790,000 acre·ft are considered as the water sources, whereas water requirements for irrigation and municipal purposes are considered as the primary demands. The practical significance of this application can be summerized as follows.

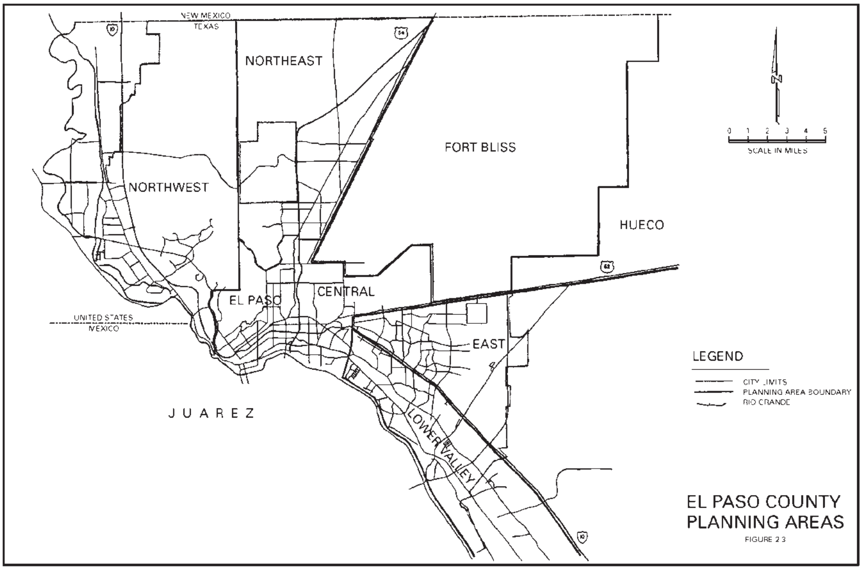


FIGURE 8.5 Major water supply regions in El Paso County. (Boyle, 1991)

TABLE 8.1 Population Projections for the Seven Planning Areas of the City of El Paso

Planning area	Year 1980	1990	2000	2010	2020	2030	2040
Northwest	48,938	90,111	135,031	176,800	231,371	280,907	304,634
Northeast	75,398	88,940	106,866	123,696	138,897	154,365	159,162
Central	141,533	140,694	143,184	145,744	145,648	146,184	146,471
Lower Valley	115,807	152,177	192,046	244,025	305,063	370,283	406,870
East	72,023	110,610	141,711	179,014	220,213	267,535	301,026
Fort Bliss	26,200	26,661	26,700	26,700	26,700	26,700	26,700
Hueco			6,650	9,816	13,872	18,731	23,053
Total	479,899	609,193	752,188	905,795	1,081,764	1,264,705	1,367,916

Source: Boyle Engineering Corporation, 1991a,b,c.

1. All the concerned parties get a fair share of the Rio Grande water.
2. The water requirements under the international treaty are satisfied.
3. The minimum requirements for industrial purposes are satisfied.
4. The net profit from the irrigation and municipal water supplies is maximized.
5. The TDS damage is minimized.

The potential income from the irrigation water supply is dependent on the location of two major irrigation districts within the project region. These include the EBID and the EPCWID No. 1. The average yearly return from the former irrigation district is estimated to be about \$266 per acre of land, which needs a total of 5 acre·ft of water per year. Of this water requirement, 3 acre·ft of water should be obtained from diverted surface water from the Rio Grande and 2 acre·ft of water should be pumped in the vicinity of the irrigation area. For EPCWID No. 1, the average return per acre of land irrigated is estimated to be about \$187 per year. All the water for irrigation purposes for this district, also 5 acre·ft per year per acre of land, should be obtained from surface water. Because of its adverse salinity effect on crop growth, groundwater use is not to be opted for in this irrigation district. Table 8.2 gives the incomes from different crops and the weighted average return per acre of irrigated land.

The return from domestic water supply can be derived using the concepts of econometrics. In general, the return is given as

$$R = aQ - \left(\frac{b}{N}\right) Q^2 \tag{8.36}$$

where R = return, \$/yr
Q = yearly domestic water supply, acre·ft
N = number of households in city or town
a and b = constants

Table 8.3 gives the values of a and b for different towns and cities in the project area.

The cost of supplying surface water from the river to the towns and cities is estimated at \$503.7 per acre·ft per year, whereas the cost of supplying water from the same source to irrigation districts is estimated at \$15 per acre·ft per year. The cost of supplying groundwater to the towns and cities is estimated at \$325 per acre·ft per year, and the cost of supplying groundwater to the irrigation districts is estimated at \$10 per acre·ft per year.

The economic damage that results from supplying saline water to domestic uses and irrigation districts is generally small, but appreciable. The damage results from the reduced lives of household appliances such as dishwashers, water heaters, food waste disposers, water softeners, and evaporative coolers. In the towns and cities that are present in the study area, the damage due to salinity is estimated at 3 cents per household per 1 ppm of TDS per year. Therefore, the TDS damage generally can be expressed as

TABLE 8.2 Return from Agricultural Activity

Irrigation district	Grains	Forage	Cotton	Chile	Pecans	Average
EBID						
Acreage	4,550	24,240	17,048	17,459	13,767	
Benefit, \$/acre·ft	60	350	126	278	300	266
EPCWID						
Acreage	6,676	11,331	22,404	1,464	5,047	
Benefit, \$/acre·ft	65	319	133	253	187	18

TABLE 8.3 Values of a and b in Eq. (8.36) for Different Cities in the Project Area (1999/2000 Estimates)

S. No.	Town/city name	a value	b value	No. of households	b/N value
1	Hatch	8,948	7,480	1,000	−7.479
2	Las Cruces	21,604	18,050	19,000	−0.950
3	Anthony	8,948	7,480	2,000	−3.739
4	El Paso	8,948	7,480	120,000	−0.062

$$D = 0.03NC(Q) \tag{8.37}$$

where D = damage, \$/yr
N = number of households for which Q is supplied
$C(Q)$ = TDS concentration in Q

On the other hand, the damage to crops because of the salinity of the irrigation water occurs in the form of reduced yield of irrigated crops due to toxic and osmotic effects. In the project area, this is reflected by the reduced income per acre of irrigated land in EPCWID No. 1. The salinity damage to irrigated crops is estimated as the average difference between the income from agriculture in EBID and that in EPCWID's irrigation district, which is presumed to be due to the increase in the salinity in the latter irrigation district. Thus the damage is estimated to be 2.6 cents per the increased TDS over the base value of 450 ppm per acre·ft of water supplied per year.

As shown in Fig. 8.6, the entire project area is subdivided into a total of 17 reaches. In these reaches, the points of interest considered include diversion node, return node, two potential irrigation areas (one on each side), and municipal and industrial points of use. These points of use are currently taken for this project to be the maximum possible points of interest in each reach. Possible return flows to this node or reach from as far as four reaches upstream are considered, which is the maximum distance currently observed to be traveled by the drainage system in the project region (see Fig. 8.7).

At each demand point in a reach, water supply may be obtained from diverted water or from pumpage except for agricultural purposes in reaches 12 and 15 where no pumpage is desirable. Possible return flows from each demand point to the stream as far downstream as four reaches below the reach, in retrospect of the continuity at the node, are considered (see Fig. 8.8). The municipal and industrial pumpages are from deeper groundwater storage, which has a significantly lower TDS level of approximately 250 ppm. However, the pumped waters for irrigation purposes for the Percha, Leasburg, and Mesilla irrigation areas are from shallow groundwater that has a high TDS level, approximately 1500 ppm, and are assumed to be essentially drawn from the river. Before the diverted water reaches the point of use for irrigation purposes, a significant amount of seepage loss occurs. Although this water is a loss to the immediate purpose for which it is diverted, it will find its way back to the river as a drainage flow or percolate deep into the groundwater storage. For this project, it is assumed that about 40 percent of the water diverted for irrigation purposes will be lost and will become seepage flow, which will eventually drain back to the stream flow. The return flows from the municipalities take place through conduits and thus there is negligible seepage loss.

At each of these demand points, the consumptive use is accounted for by determining the amount of water that is consumed. At the agricultural areas, this constitutes the evapotranspiration, whereas at the municipal points of use, the

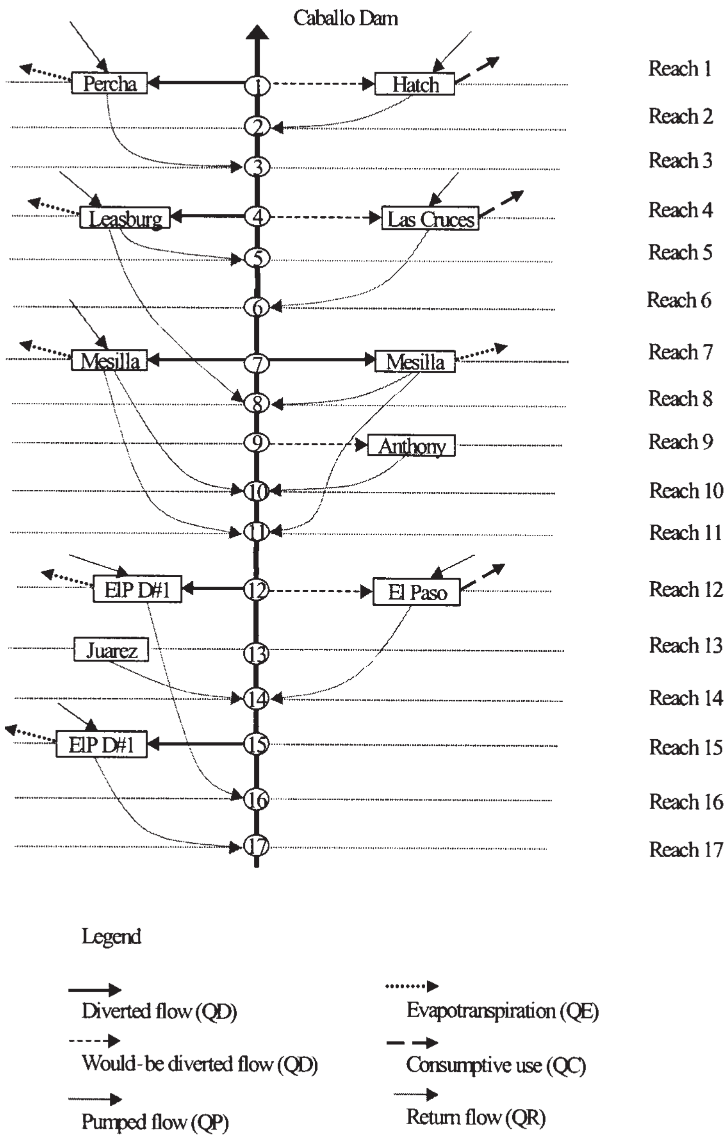


FIGURE 8.6 Schematic diagram of the Rio Grande project system.

consumptive losses constitute the consumptive uses for purposes such as drinking and cooking. At the industrial points of use, the consumptive losses constitute the amount that is used in the industrial by-product. To account for these effects, it is assumed that about 50 percent of the water used for irrigation and municipal purposes and about 70 percent of the water used for industrial purposes are “used up.”

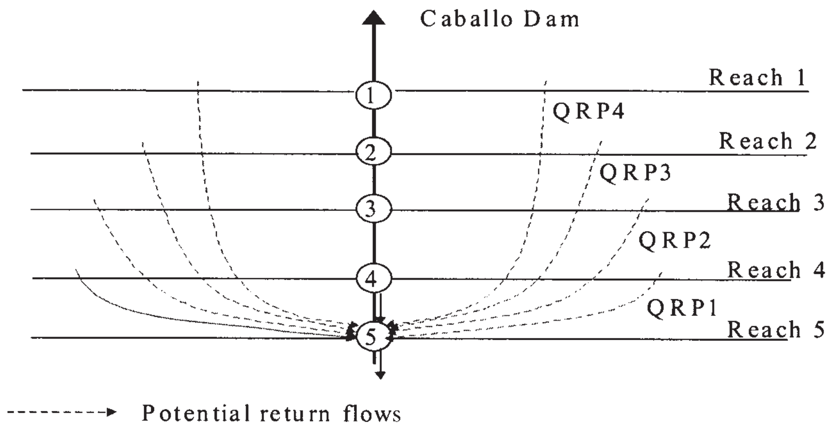


FIGURE 8.7 Schematic diagram showing continuity at a node.

The remaining water from each point of use is assumed to find its way back into the stream flow in the form of drainage.

8.3.4 Seasonal NLP Model Application

The seasonality aspect of the management of the Rio Grande project results from the fact that some of the activities occur only during certain seasons of the year. The irrigation activities occur during the spring, summer, and fall seasons. In addition, the magnitude of the urban water demands depends on the season. For each city and town, this demand is maximum during the summer season and minimum during the winter season. Because of these reasons, consideration of seasonal effects may prove to be more appropriate for this project. Therefore, the planning year is divided into four seasons and the problem is reformulated as a more complex, but more realistic, NLP problem. The seasons considered include December to February as season 1, March to May as season 2, June to August as season 3, and September to November as season 4. The approximate demand proportions by the different activities are given in Table 8.4. This approach surpasses the original problem formulation for a full year that was presented earlier. However, this approach tends to transform the prevailing management practice into a more realistic optimal model.

The general form of the objective function for the urban water demands is essentially the same as the one used in the static model except that the annual urban water demand is taken as the sum of the seasonal demands. For the cities and towns, the modified income (benefit) from the supply of demand is expressed as

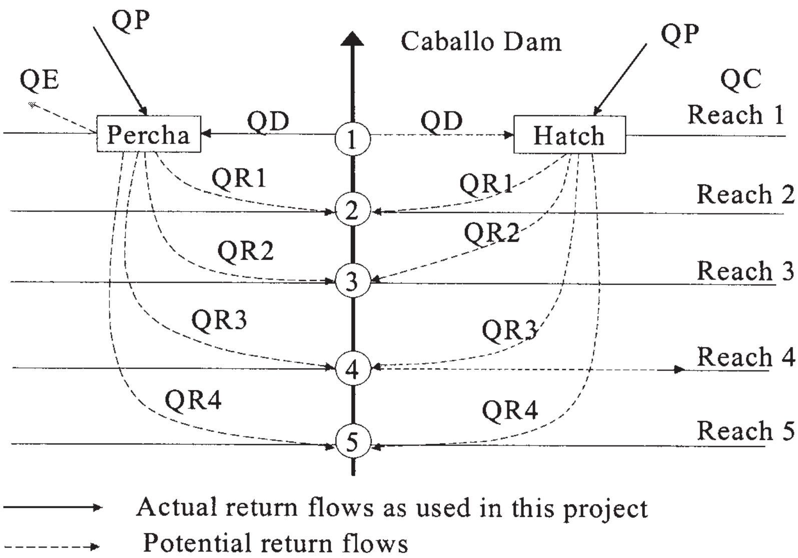


FIGURE 8.8 Schematic diagram showing continuity at a point of use.

TABLE 8.4 Approximate Demand Proportions by Different Activities at the Rio Grande Project

Activity	Dec–Feb	Mar–May	Jun–Aug	Sep–Nov
Municipal (fractions)	0.15	0.25	0.35	0.25
Agricultural (unit proportions)				
Diversion	—	0.30	2.20	0.50
Pumpage	—	1.50	0.50	—
Subtotal	—	1.80	2.70	0.50

$$a[Q(S_1) + Q(S_2) + Q(S_3) + Q(S_4)] - \frac{b}{N}[Q(S_1) + Q(S_2) + Q(S_3) + Q(S_4)]^2 \tag{8.38}$$

where $Q(S_k)$ denotes the demand during season k where $k = 1, 2, 3,$ or 4. The other parameters are as defined earlier. It may be noted further that Eq. (8.38) is essentially the same as Eq. (8.36) except that in this case the total yearly demand is substituted by the sum of the four seasonal demands.

For the agricultural activities, it is assumed that the total income will be generated only during the harvesting season of September through November, that is, season 4. The total income taken is the same as that used in the static model, but the water demands for this purpose prevail during the last three seasons of the year to generate this income. Thus, the total income per acre of land is taken as \$266 in the agricultural areas in reaches 1, 4, and 7 and as \$187 in the agricultural areas in reaches 12 and 15. Similarly, the water supply costs and the TDS damages used in the objective function are taken as the sum of the seasonal costs and damages. The costs of supplying surface water and groundwater to the towns and cities are taken as \$503.7 and \$325 per acre·ft, respectively, for the four seasons of the year. These figures were taken in the static model as the annual costs of surface water and groundwater supplies, respectively. The cost of supplying water from the river to the irrigation districts is estimated at \$15 per acre·ft for the last three seasons of the year during which diversions to the irrigation districts prevail. Similarly, the cost of supplying groundwater to the irrigation districts is estimated at \$10 per acre·ft for the two seasons of the year during which pumpage to the irrigation districts is allowed. Again, these figures were taken in the static model as the annual costs of surface water and groundwater supplies for irrigation, respectively.

The TDS damages are also broken down into seasons so that the seasonality effects are reflected. The seasonal TDS damages that result from both the urban and the irrigation supplies are approximated by taking the same proportionate values as the seasonal demands (see Table 8.4). Thus, the seasonal urban TDS damages per 1 ppm are 0.45, 0.75, 1.05, and 0.75 cents for the first, second, third, and fourth seasons, respectively. Similarly, the estimated irrigation TDS damages per ppm for the last three seasons are 0.936, 1.404, and 0.26 cents, respectively.

In the static model, it was assumed that the total drainage flow occurs during the same year of diversion for all agricultural, municipal, and industrial activities. However, in this model, it is assumed that the total drainage flow from the agricultural activities is spread over two seasons, with about 75 percent of the drainage occurring during the same season as the diversion season and the remaining 25 percent of the drainage occurring during the subsequent season. No diversion or pumping activity for irrigation exists during the first season. However, there is carryover drainage during this season from the diversion during the fourth season of the previous year. This is accounted for by assuming the same 75 percent return flow during the same season and the remaining 25 percent return flow during the first season of the following year. For practical purposes, the quantity of this drainage is taken to be the same as the previous year's drainage during the same season, that is, the first season. Since the drainage from the municipal and industrial activities takes place in conduits, it is assumed to occur during the same season of diversion and pumping.

To write the above distribution of drainage flows over two seasons in equation form, let $Q_{lm(kij)}$ be the return flow from reach *i,* that was diverted for purpose *j,* to *l* reaches downstream from the water diverted during season *m* and returning during season *k.* Then the two general drainage flow distributions over two seasons can be given as

$$Q_{lm\,(kij)} = 0.75\,(Q_{lm\,(kij)} + Q_{l\,(m+1)\,(kij)}) \tag{8.39}$$

$$Q_{l\,(m+1)\,(kij)} = 0.25\,(Q_{lm\,(kij)} + Q_{l\,(m+1)\,(kij)}) \tag{8.40}$$

To better illustrate the notations used here consider, for instance, $Q_{22(\mathrm{S2,R1,A1})}$, which is the drainage flow during season 2 that is returning to two reaches downstream of reach 1 (R1) from the water supplied during season 2 (S2) to agricultural activity 1 (A1). Figure 8.9 shows the general scheme of the seasonal flow system of the Rio Grande project for any given typical season. In general, there is one such scheme for every season.

8.3.5 Capacity Expansion Model Application

Solution Methodology. The MINLP model formulated consists of two sets of variables: discrete time variables and continuous cost variables associated with water demand and treatment, thus making analytical solution of the problem difficult. However, if the discrete binary integer variables are set to specific values, then the problem at that discrete point in time reduces to an NLP problem. The resulting NLP problem can be solved by an appropriate NLP computer program. Nonetheless, such a solution can be obtained only after the values of the integer variables have been fixed.

For small problems, it may be possible to fix the values of the binary discrete time variables and solve the reduced NLP at each of the discrete points, one at a time, using an NLP computer program. In such a case, the best solution will be obtained from these optimal solutions. For a large number of discrete points, in which case it would become very tedious to use this approach, the use of some form of heuristic search algorithm, such as SA, becomes a necessity. One such algorithm may be the use of an automated computer program that is capable of performing the following tasks:

1. Generate the values of the discrete time variables one at a time.
2. Use this information to reduce the MINLP to an NLP.
3. Obtain the optimum solution of the reduced NLP at each of the discrete points.
4. Return back the best optimum solution among those obtained for all the reduced NLPs.

This solution strategy is pursued in two phases. In the first phase the simulated annealing heuristic search algorithm is implemented to solve a combinatorial problem. In the second phase, the SA code is interfaced with GRG2 to solve the MINLP problem.

The general SA algorithm is given below (Wolsey, 1998, with modifications in the notations).

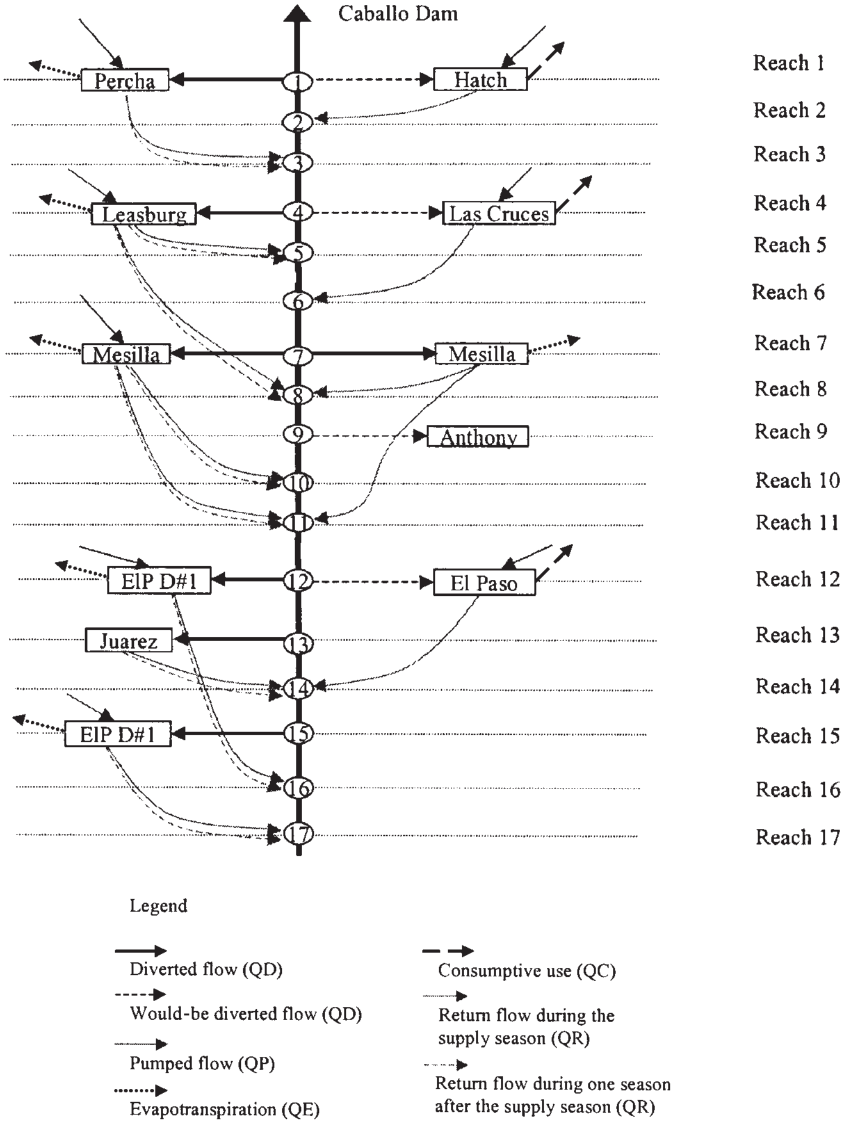


FIGURE 8.9 Schematic diagram of the Rio Grande project system showing the drainage for any given season in which diversion to agriculture exists.

1. Get an initial solution $\mathbf{X}_o$.
2. Get an initial temperature T and a reduction factor α with $0 < \alpha < 1$.
3. While not yet frozen, do the following:
  3.1 Perform the following loop k times:
    3.1.1 Pick a random neighbor $\mathbf{Y}$ of $\mathbf{X}$.
    3.1.2 Let $\delta = f(\mathbf{Y}) - f(\mathbf{X})$.
    3.1.3 If $\delta \leq 0$, set $\mathbf{X}_{i+1} = \mathbf{Y}$.
    3.1.4 If $\delta > 0$, set $\mathbf{X}_{i+1} = \mathbf{Y}$ with probability $e^{-\delta/T}$; otherwise, set $\mathbf{X}_{i+1} = \mathbf{X}_i$.
  3.2 Set $T \leftarrow \alpha T$ (reduce the temperature).
4. Return the best solution found.

The implementation of the SA heuristic search algorithm requires that a neighborhood and a search strategy be defined first. The starting point with a given set of time units is chosen randomly, and its neighbors are selected. For any given discrete point, the number of neighbors considered is a function of the total number of units to be expanded.

In the course of the neighborhood search, if the current point is at or adjacent to a boundary, that is, where one or more of the time lines are at their limits which include the beginning and the end of the planning period along each time line, then special attention is paid to the search mechanism. This is tackled by the search algorithm in such a way that whether the randomly selected neighbor is at or adjacent to any boundary line is checked before any move. If it is already at the boundary, then no move is allowed, that is, the new neighbor coincides with the center point. If it is already one unit away from the boundary and if the selected neighbor were to move two time units to the boundary, then no two time unit move is allowed; instead, it moves only one time unit to reach the boundary.

The Generalized Reduced Gradient (GRG2). GRG2 solves nonlinear optimization problems with the general form

$$\text{Minimize} \quad f(\mathbf{x}) \tag{8.41a}$$

$$\text{Subject to} \quad \mathbf{g}(\mathbf{x}) = \mathbf{0} \tag{8.41b}$$

$$\underline{\mathbf{x}} \leq \mathbf{x} \leq \overline{\mathbf{x}} \tag{8.41c}$$

where $\mathbf{x}$ = n-dimensional column of vector variables
$f(\mathbf{x})$ = objective as a function of $\mathbf{x}$
$\mathbf{g}(\mathbf{x})$ = m constraints
$\underline{\mathbf{x}}$ = lower bound on $\mathbf{x}$
$\overline{\mathbf{x}}$ = upper bound on $\mathbf{x}$

The fundamental idea of the generalized reduced gradient is to express m variables, called basic variables, in terms of $n - m$ variables, called nonbasic variables. The objective function is thus expressed as a function of the nonbasic variables, which are bound between their upper and lower limits. The resulting formulation, called the reduced problem, is solved using the unconstrained NLP solution technique with modifications to account for the bounds for the nonbasic variables. Solving a sequence of reduced problems solves the original NLP problem.

Interfacing the Simulated Annealing Algorithm with GRG2. The SA search algorithm code was interfaced with the GRG2 code to solve the MINLP problem. Figure 8.10 shows the general flowchart of the computer program used to solve the MINLP problem. To use the GRG2 code, a user-supplied GCOMP subroutine was written. The main elements of this subroutine are the definitions of the objective function and the constraint functions, which are given as a list of functions whereby the index of the objective function is specified in the main program. Furthermore, two options exist for the user in selecting the methodology to calculate the gradient. GRG2's default option is to use the forward differencing procedure. Alternatively, partial derivatives of the functions given in the GCOMP subroutine can be determined by the user and another subroutine called PARSH can be provided to the model. For this application, GRG2's default option was used.

General Construction, Operation, and Expansion Cost Equations. In general, many capital facilities are usually constructed with a capacity that will satisfy the requirements over some years to come, instead of a capacity that satisfies immediate requirements. The main reason for this lies in the economies of scale available with a large plant that may be achieved in investment cost or operating cost (Hinomoto, 1974). Chenery (1952) proposed the following power function for an industrial facility.

$$C = \alpha K^{\beta} \tag{8.42}$$

where C = investment cost
K = design capacity
α and β = positive parameters determined by observed data

In this equation, if $K = 1$, then C equals α, which can be defined as the investment cost of a system with a capacity of one unit. Parameter β determines the manner in which investment cost changes capacity. The investment cost increases with capacity at an increasing or decreasing rate depending on whether β is greater than or less than 1, respectively.

The general equation for the construction of a water supply unit, including setup costs, K_c, can be expressed as

$$\mathrm{CC}(Q_{ijp}) = K_c + \alpha_c Q_{ijp}{}^{\beta c} \tag{8.43}$$

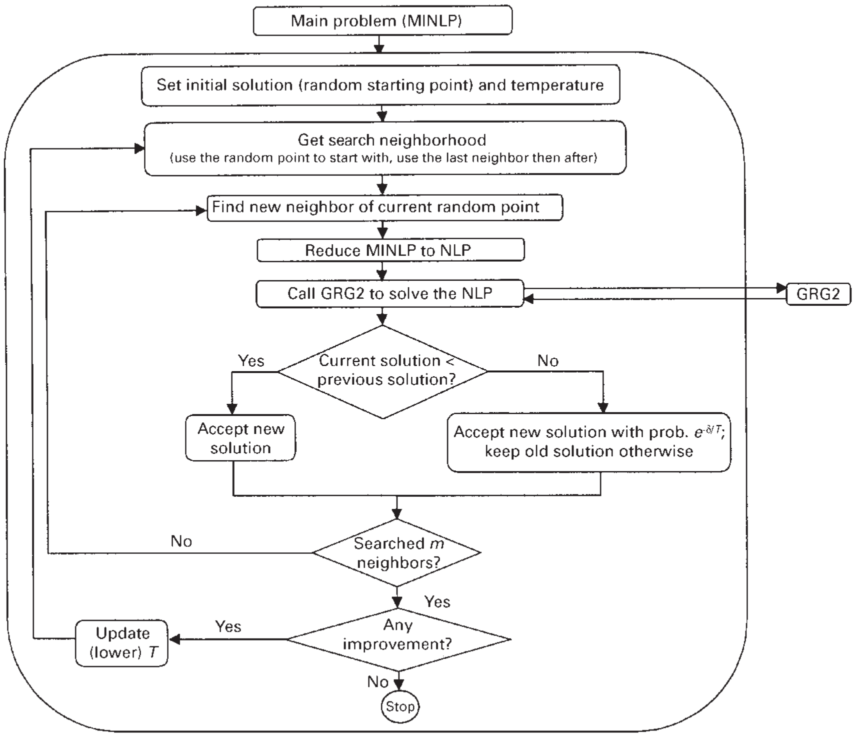


FIGURE 8.10 A flowchart showing the procedures used by the model.

where α_c and β_c are constants. Using this general form, the operation cost for a water supply facility during a planning period p is expressed as

$$\mathrm{OC}(Q_{ijp}) = K_o + \alpha_o Q_{ijp}{}^{\beta_o} \tag{8.44}$$

and for expansion of the water supply facility as

$$\mathrm{EC}(Q_{ijp}) = K_e + \alpha_e (Q_{ij\,(p+1)} - Q_{ijp})^{\beta_c} \tag{8.45}$$

where K_c, K_o, K_e α_o, β_o, α_e, and β_e are constants. The construction, operation, and expansion costs of a water treatment facility are expressed, respectively, as

$$\mathrm{CC}(\mathrm{CAP}_{ijp}) = C_c + \theta_c (\mathrm{CAP}_{ijp})^{\phi_c} \tag{8.46}$$

$$\mathrm{OC}(\mathrm{CAP}_{ijp}) = C_o + \theta_o (\mathrm{CAP}_{ijp})^{\phi_o} \tag{8.47}$$

$$\mathrm{EC}(\mathrm{CAP}_{ijp}) = C_e + \theta_e (\mathrm{CAP}_{ij\,(p+1)} - \mathrm{CAP}_{ijp})^{\phi_e} \tag{8.48}$$

where C, θ, and ϕ are constants with the subscripts c, o, and e denoting construction, operation, and expansion, respectively.

The EPA has developed empirical functions for the estimation of the three types of costs for water and wastewater conveyance systems and water and wastewater treatment plants. Clark et al. (2002a,b) developed different cost equations for the construction of water supply distribution system components; for water supply distribution system rehabilitation; and for the construction of pumps, tanks, and reservoirs. The cost functions for water supply distribution system components for different pipe systems include:

1. Base installed pipe costs
2. Trenching and excavation costs
3. Embedment costs
4. Backfill and compaction costs
5. Valve fitting and hydrant costs
6. Dewatering costs
7. Sheeting and shoring costs
8. Horizontal boring costs
9. Pavement removal and replacement costs
10. Utility interference costs
11. Traffic control costs
12. Household service connection costs

The general cost equation for each of these different costs is given as

$$y = a + bx^c + du^e + fxu \tag{8.49}$$

where y = cost of a particular component, \$/ft
x = design parameter (for example, pipe diameter)
u = indicator variable

The variables a, b, c, d, e, and f are estimated using regression techniques. Table 8.5 gives the values of these parameters for selected pipe types.

The aggregated cost of construction can be determined in an additive manner by considering the construction cost components involved (Clark et al., 2002a,b). Thus, considering the cost of using a ductile iron pipe (base installed, with push-on joints) and the cost of the major construction activities involved, including trenching and excavation, embedment, backfill, and compaction, and valve fitting and hydrant placement, the total aggregated cost can be estimated by

TABLE 8.5 Parameters for Distribution System Component Cost Equations for Selected Pipe Types

Type of activity	Conditions	Pipe diameter range, in	Indicator variable u value	Parameter values						
				a	b	c	d	e	f	R^2
Ductile iron pipe	Base installed cost, with push-on joints	4–144	50, 52	−44.33	0.33	1.72	2.87	0.74	0.0	0.99
Trenching and excavation	Sandy gravel soil with 1:1 side slope	1. 4–8	4, 6, 8, 10 ,12	−24.18	0.32	0.67	16.66	0.38	0.0	0.99
		2. 8–144	4, 6, 8, 10, 12	2.91	0.0018	1.90	0.13	1.77	0.0	0.98
Embedment	Concrete arch	4–144	—	7.11	0.26	1.46	0.0	0.0	0.0	0.99
Backfill and compaction	Sandy native soil with a 1:1 side slope	4–144	4, 6, 8, 10, 12	−0.094	−0.062	0.73	0.18	2.03	0.02	0.99
Valve fitting and hydrant	Medium frequency using ductile iron fitting	12–72	—	9.83	0.02	1.79	0.0	0.0	0.0	0.99
Horizontal boring	—	4–60	—	503.67	1.99	1.52	0.0	0.0	0.0	0.98

Source: Clarke et al. (2002a).

$$y_{\text{total}} = 50.74 + 0.33x^{1.72} + 0.32x^{0.67} + 0.26x^{1.46} - 0.062x^{0.73} + 0.02x^{1.79} + 0.16x \tag{8.50}$$

where y_{total} is the aggregated construction cost (\$/ft) and x is the pipe diameter (in). To use Eq. (8.50) in the model, the diameter is expressed in terms of the flow rate in the pipe since the flow rate, rather than the pipe diameter, is used in the model. Also, the total distance of construction must be determined. Using the equation of continuity and an average design velocity of 5 ft/s, the diameter is expressed in terms of Q as $x = 6.06Q^{0.5}$ from which Eq. (8.51) is obtained.

$$y_{\text{total}} = 50.74 + 7.32Q^{0.860} + 1.07Q^{0.335} + 3.61Q^{0.730} - 0.23Q^{0.365} + 0.50Q^{.895} + 0.97Q^{0.5} \tag{8.51}$$

Therefore, if L_{ij} is the length of construction from the diversion node in reach i to activity point j, the total construction cost $\text{CC}(Q_{ijp})$ for a water supply line in reach i for activity j during planning period p can be given as

$$\text{CC}(Q_{ijp}) = L_{ij}(50.74 + 7.32Q_{ijp}{}^{0.86} + 1.07Q_{ijp}{}^{0.335} + 3.61Q_{ijp}{}^{0.73} - 0.23Q_{ijp}{}^{0.365} + 0.50Q_{ijp}{}^{0.895} + 0.97Q_{ijp}{}^{0.5}) \tag{8.52}$$

Although Eq. (8.52) can be used as an estimate for construction costs, such costs may change with time. In fact, the original equations given in the sum of equations in Eq. (8.50) were derived based on the Producer Price Index (PPI) for 1997. However, such costs can be updated by the use of standard indices. For example, the PPI for 1997 is 358.5, whereas that for 2000 is 388.5, which yields a ratio of 1.08. Therefore, the above costs can be updated to 2000 by multiplying the 1997 costs by 1.08 (Clark et al., 2002a,b).

The costs for pump, reservoir, and tank construction and operation can be given by an equation that is essentially similar to Eq (8.24) as (Clark et al., 2002a):

$$y = a + bx^c + du^e + fxu \tag{8.53}$$

where y = construction cost (\$)
x = design variable
u = indicator variable

Parameters $a, b, c, d, e,$ and f are determined from regression analysis. Table 8.6 gives values of these parameters for some selected activities.

Using the parameters given in Table 8.6, aggregated cost equations for construction, operation, and maintenance of pump stations and reservoirs are determined. Therefore, if a horizontal split case pump with suction and discharge piping with 200 ft total dynamic head is selected, the aggregated construction cost can be computed as the sum of the base construction cost, the pump station site

TABLE 8.6 Parameters for Costs of Selected Pumps, Tanks, and Reservoirs

Type of activity	Conditions	Independent variable description and range	Indicator variable *u* values
Horizontal split case pumps	Includes suction discharge piping	Firm pumping capacity (0.5–200 mgd)	100, 200, or 300 ft of total dynamic head
Pump station site work	—	Firm pumping capacity (0.5–200 mgd) for extreme conditions	1 for average conditions and 2
Building installation for horizontal centrifrigal pump	—	Firm pumping capacity (0.5—200 mgd)	2 for a simple slab
Construction of electrical and instrumentation	—	Firm pumping capacity (0.5–200 mgd)	100, 200, or 300 ft of total dynamic head
Expansion for horizontal split case pump	Includes suction and discharge pumping	Additional pumping capacity (0.5–200 mgd)	100, 200, or 300 ft of total dynamic head
Expansion of site work	—	Additional pumping capacity (0.25–33 mgd) conditions	1 for average conditions and 2 for extreme
Expansion of electrical and instrumentation	—	Additional pumping capacity (0.25–12 mgd)	100, 200, or 300 ft of total dynamic head
Annual operation and maintenance for horizontal centrifugal pump installations	—	Firm pumping capacity (0.5–200 mgd)	100, 200 or 300 ft of total dynamic head

Source: Clarke et al. 2002b.

TABLE 8.6 Parameters for Costs of Selected Pumps, Tanks, and Reservoirs (*Continued*)

Parameter values						
a	*b*	*c*	*d*	*e*	*f*	R^2
33,246.43	15,615.82	0.99	0.0	—	22.78	0.99
−14,695.02	25,947.34	0.42	0.28	1.00	0.0	0.99
225,500.9	8,629.99	0.84	−116,177.1	0.88	−1,449.77	0.99
24,953.27	40,018.32	0.62	0.072	2.20	33.93	0.91
12,766.78	23,136.77	1.06	0.18	1.81	0.0	0.82
−26,101.60	31,414.44	0.04	159.61	0.0	0.0	0.98
−31,272.00	7,902.90	0.93	9,500.48	0.34	9.92	0.90
10,094.42	1,369.70	0.97	0.08	1.69	135.00	0.99

work, the building cost, and the construction cost for electricity and instrumentation. Considering these costs, we obtain the following equation:

$$\mathrm{CC}(Q_{ijp}) = 63{,}506.35 + 23{,}639.68Q_{ijp}^{0.95} + 37{,}931.97Q_{ijp}^{0.87} + 12{,}451.92Q_{ijp}^{0.84} + 52{,}454.49Q_{ijp}^{0.62} + 13{,}062.42Q_{ijp} \tag{8.54}$$

The annual operation and maintenance cost for the selected type of horizontal centrifugal pump installations can also be determined by referring to Table 8.6. For a pump station of firm capacity of 0.5 to 200 mgd and total dynamic head of 200 ft, this cost is obtained as

$$\mathrm{OC}(Q_{ijp}) = 10{,}713.62 + 2091.67Q_{ijp}^{0.97} + 41{,}775.18Q_{ijp} \tag{8.55}$$

The expansion cost for the selected type of pump can be determined as the aggregate of the expansion costs of the pump, the site work, the building, and the extension of electrical power and instrumentation. The parameters of these costs are obtained from Table 8.6.

$$\mathrm{EC}(Q_{ijp}) = 27{,}151.70 + 36{,}747.74Q_{ijp}^{1.06} + 31{,}967.71Q_{ijp}^{0.04} + 11{,}064.73Q_{ijp}^{1.07} + 11{,}859.66Q_{ijp}^{0.93} - 7935.95Q_{ijp} \tag{8.56}$$

where $\mathrm{EC}(Q_{ijp})$ is the total construction cost for a supply line in reach i for activity j during time p of the planning period. Expansion costs for pipelines depend on the approach used in the process, which include pipe bursting, microtunneling, and horizontal directional drilling. For this example, pipe bursting, a method for replacing pipe by bursting from within while simultaneously pulling in a new pipe (Selvakumar et al., 2000), is selected. The liner pipe can be the same size or as much as two pipe sizes larger than the existing pipe (Selvakumar et al., 2000). Boyce and Bried (1998) estimate a range of \$7 to \$9 per inch diameter for installed costs for pipe replacement. Thus, an approximate cost equation for pipe replacement is obtained as

$$\mathrm{EC}(Q_{ijp}) = 8(12)\,(6.06Q^{0.5})L_{ij} = 581.76Q_{ijp}^{0.5}L_{ij} \tag{8.57}$$

8.4 MODEL RESULTS

8.4.1 Static Model

The formulation given is prepared as the input code for the GAMS/MINOS solver (Brook et al., 1992), and the results shown in Tables 8.7 and 8.8 are obtained, which give a net objective value of \$434.12 million. A summary of the major breakdown of the costs and benefits of this project in the major water-use districts is given in Table 8.8, whereas Fig. 8.11 gives its pictorial illustration. The results clearly show that the damage from the TDS is more pronounced in the irrigation districts than in the cites and towns. The TDS damages to the cities and towns can be considered practically negligible when compared to the gross income from the cities and towns.

The solution of the problem depends on the initial values of some of the parameters in the model. As a result, different local maximum points were obtained depending on the initial conditions used. The most important parameter among these appeared to be the initial value of the river flow. Using different values of the initial river flow (from 400,000 to 850,000), the result in Table 8.9 was obtained. A comparison of the calculated TDS values at each of the nodes from 1 to 13 with 30-year available data at these nodes is given in Table 8.10 and Fig. 8.12.

The sensitivity analysis to the most important factor in this project, the average annual water supply, shows two important trends (see Fig. 8.13). First, even when there is more than enough water in the river, the TDS mixing dynamics make the total damage higher when the surplus flow in the river is low. This is due to the fact that the drainage water from the agricultural areas always carries higher TDS than the water flow in the river. The marginal increase in the objective function versus the annual available flow in the river, shown roughly on the right half of Fig. 8.13, illustrates the dilution effect of the river water on the drainage water. The other half of the figure shows the dilution effects plus the marginal increase due to agricultural activity as a function of the marginal increase in the acreage.

TABLE 8.7 "Best" Optimal Solution Obtained

Objective function value, million $	Water supply to the irrigation districts, acre·ft/yr			Water supply to the cities/towns, acre·ft/yr		
	Area name	Diversion	Pumpage	City/town	Diversion	Pumpage
434.12	Percha	75,000	30,000	Hatch	0	576.682
	Leasburg	100,000	40,000	Las Cruces	0	11,201.686
	Mesilla (right)	75,000	30,000	Anthony	0	1,153.389
	Mesilla (left)	75,000	30,000			
	El Paso (reach 12)	83,333.333	N/A	El Paso	13,028.762	55,000.000
	El Paso (reach 15)	416,666.667	N/A			

Note: N/A = not applicable.

TABLE 8.8 Breakdown of the Benefits and the Damages for the "Best" Optimal Solution Obtained

Benefit/cost type	EBID	EPCWID No. 1	NM cities/towns*	El Paso
Gross benefit, $	17,290,000	11,319,000	130,815,109	321,790,790
Water cost, $	4,225,000	4,500,000	4,202,821	24,437,587
TDS damage, $	4,120,457	3,416,011	150,000	2,036,815
Net benefit, $	8,944,543	3,402,989	126,462,288	295,316,388
Total net benefit, $	434,126,208			

*These include Hatch, Las Cruces, and Anthony.

8.4.2 Seasonal Model

The seasonal model formulation was also prepared as a GAMS/MINOS input code, which was solved using the latest release (version 2.50) of GAMS/MINOS software. The optimal solution obtained is given in Table 8.11. The objective function value obtained for this solution is $435.43 million. The optimal seasonal releases from Caballo Reservoir are obtained as 19,188, 128,648, 555,022, and 87,141 acre·ft for the first, second, third, and fourth seasons, respectively. The TDS concentrations observed in the reaches did not show much disparity between the seasons.

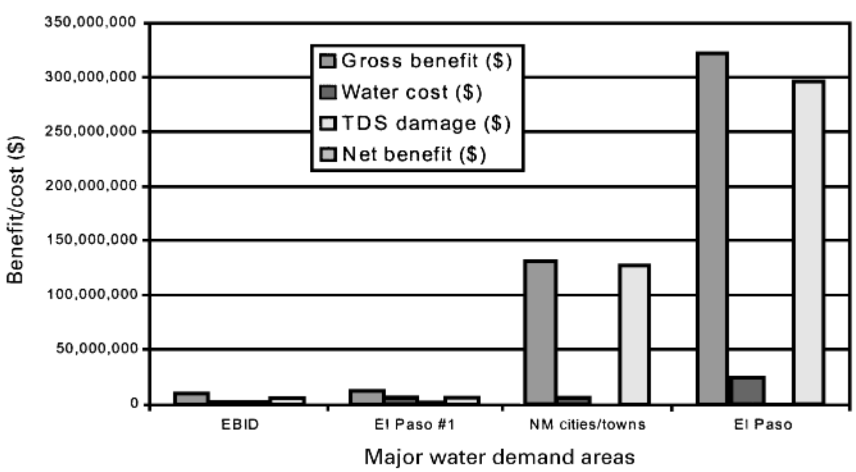


FIGURE 8.11 Comparison of the major benefits and costs for the "best" optimal solution.

TABLE 8.9 Sensitivity Analysis of the Net Benefit Result to Annual Available Flow

Reservoir release, acre·ft/yr	Benefit, million $	Reservoir release, acre·ft/yr	Benefit, million $	Reservoir release, acre·ft/yr	Benefit, million $
400,000	429.78	560,000	431.91	690,000	433.47
410,000	429.91	570,000	432.05	700,000	433.54
430,000	430.17	580,000	432.18	710,000	433.61
450,000	430.44	590,000	432.32	720,000	433.68
460,000	430.57	600,000	432.45	730,000	433.75
470,000	430.71	610,000	432.59	740,000	433.81
480,000	430.84	620,000	432.72	750,000	433.88
490,000	430.97	630,000	432.84	760,000	433.94
500,000	431.11	640,000	432.97	770,000	434.00
510,000	431.24	650,000	433.09	790,000	434.12
520,000	431.38	660,000	433.22	810,000	434.23
540,000	431.64	670,000	433.31	830,000	434.34
550,000	431.78	680,000	433.39	850,000	434.45

TABLE 8.10 Observed 30-Year Average TDS Data and TDS Data Determined Using the Model

Reach	Calculated TDS, ppm	30-year average TDS, ppm
1	450	487
2	450	534
3	494	554
4	544	581
5	544	589
6	636	589
7	631	592
8	631	724
9	787	750
10	787	873
11	809	1037
12	850	1036
13	850	1020

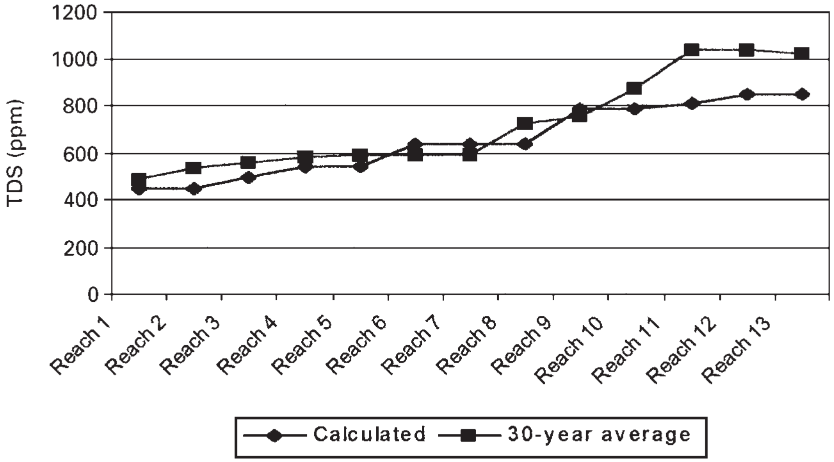


FIGURE 8.12 Graph showing the calculated and 30-year average observed TDS values.

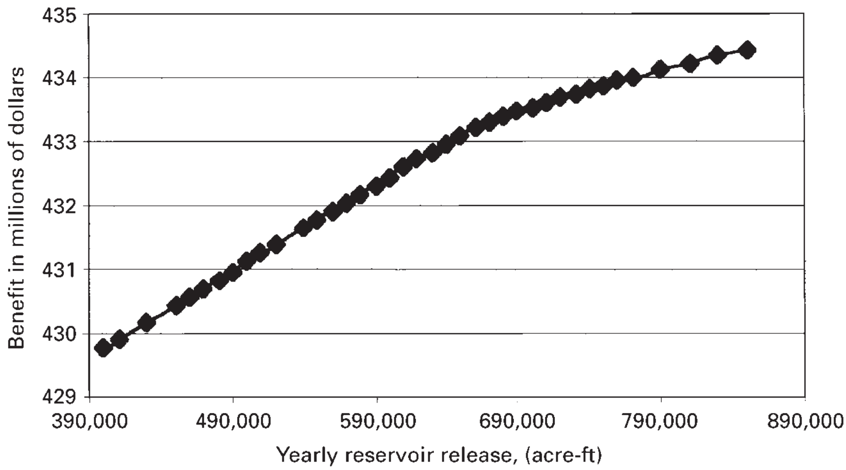


FIGURE 8.13 Net benefit as a function of the annual available flow.

8.4.3 Capacity Expansion Model

The seven major water demand areas of El Paso County's water supply system were combined together to form a total of three major water demand regions. Herein, the Northwest and Northeast areas are combined to form region 1; the Central, Fort Bliss, and Hueco areas are combined to form region 2; and Lower Valley and the East areas are combined to form region 3. The estimates of the population and the households in each region are given in Tables 8.12 and 8.13,

TABLE 8.11 The "Best" Solution Obtained for the Seasonal Model

		Agriculture (left)		Agriculture (right)		Municipal		Industrial	
Reach	Season	Div.	Pump	Div.	Pump	Div.	Pump	Div.	Pump
1	1					0	86		
	2	7,500	22,500			0	143		
	3	54,750	7,500			50	150		
	4	12,750				0	143		
4	1					0	1,680	0	1,625
	2	0	0			0	2,800	0	1,625
	3	0	0			0	3,920	1,045	580
	4	0				0	2,800	0	1,625
7	1								
	2	7,500	22,500	7,500	22,500				
	3	54,750	7,500	54,750	7,500				
	4	12,750		12,750					
9	1					0	172		
	2					0	286		
	3					100	300		
	4					0	286		
12	1					0	10,225	8,513	3,525
	2	8,333				3,292	13,750	12,038	0
	3	60,833				10,109	13,750	12,038	0
	4	14,167				3,292	13,750	12,038	0
15	1								
	2	41,667							
	3	304,167							
	4	70,833							

respectively. For the purpose of this application, a planning period of 10 years is considered. Therefore, the total number of possible combinations of expansion times equals $(10 - 1)^6 = 531{,}441$. The problem envisages minimizing the total cost over the 10-year planning period.

Using Eqs. (8.51), (8.52), and (8.54) to (8.57) in the problem formulation given by Eq. (8.35), the MINLP problem is solved. The minimum yearly water demands for El Paso for the years 2001 to 2010, which are used in the constraints, were determined from a different study by considering supply–benefit analysis. Table 8.14 gives the optimum (minimum) demand by the customers. The maximum yearly groundwater volume that can be pumped per year for each region was limited to 55,000 acre·ft. In the formulation, the length of the pipelines from the diversion points to each of the regions was assumed to be 10 mi each.

The optimal capacity expansion schedule was obtained by solving the above MINLP model. The computed expansion times for the pipelines and the pumping

TABLE 8.12 Estimated Population of Each of the Regions

Year	Region 1	Region 2	Region 3
2001	247,757	177,107	342,685
2002	253,617	177,679	351,613
2003	259,477	178,252	360,542
2004	265,337	178,824	369,470
2005	271,197	179,397	378,398
2006	277,056	179,970	387,326
2007	282,916	180,542	396,254
2008	288,776	181,115	405,183
2009	294,636	181,687	414,111
2010	300,496	182,260	423,039

TABLE 8.13 Estimated Number of Households in Each Region

Year	Region 1	Region 2	Region 3
2001	82,586	59,036	114,228
2002	84,539	59,226	117,204
2003	86,492	59,417	120,181
2004	88,446	59,608	123,157
2005	90,399	59,799	126,133
2006	92,352	59,990	129,109
2007	94,305	60,181	132,085
2008	96,259	60,372	135,061
2009	98,212	60,562	138,037
2010	100,165	60,753	141,013

stations to regions 1, 2, and 3 were found to be 9, 8, 9 and 2, 4, 10, respectively. It may be noted that the value 10 for the pumping station to region 3 indicates that there is no expansion necessary and the initial capacity should be built so that it serves during the entire planning period. The total construction, operation, and expansion costs were found to be \$136.20, \$48.04, and \$498.26 million, respectively, with the overall total minimum cost of \$682.50 million. In obtaining the above result, the termination message report by GRG2 was "Termination Criterion Met. Kuhn-Tucker Conditions Satisfied to Within [EPSTOP] at Current Point."

This is the best termination message because it ensures that the necessary conditions have been satisfied at the current point and thus the current point may be a local optimum point.

8.5 REFERENCES

Boyce, G. M., and E. M. Bried (1998). Social Cost Accounting for Trenchless Projects, *Proceedings of the Conference entitled North American No-Dig' 98,* Albuquerque, N.M.

Boyle Engineering Corp. (1991a). El Paso Water Resources Management Plan, Phase I Completion Report, July 1991.

Boyle Engineering Corp. (1991b). El Paso Water Resources Management Plan, Phase II Completion Report, August 1991.

Boyle Engineering Corp. (1991c). El Paso Water Resources Management Plan, Phase III Completion Report, February 1992.

Boyle Engineering Corp. and Parsons Engineering Science, Inc. (1998). El Paso–Las Cruces Regional Sustainable Water Project, Drain Mitigation Strategies Report, a final report prepared for New Mexico–Texas Water Commission, July 1998.

Brook, A., et al. (1992). *GAMS: A User's Guide,* Release 2.25, The Scientific Press, Redwood City, Calif.

Chenery, H. B. (1952). Overcapacity and the Acceleration Principle, *Econometrica,* No. 20, pp. 1–28.

Clark, R. M., et al. (2000a). Costing Models for Pumps, Tanks and Reservoirs, *Journal of Water Resources Planning and Management* (to be published).

Clark, R. M., et al. (20002b). Costing Models for Water Supply Distribution System Components, *Journal of Water Resources Planning and Management* (to be published).

Engineering-Science, Inc. (1991). Surface Water Supply Alternatives for the City of El Paso and Southern New Mexico Users, Final Report Prepared for the New Mexico Entities, Las Cruces, N.M.

Goldman, F. E., and L. W. Mays (2002). Simulated Annealing Approach for Operation of Water Distribution Systems Considering Water Quality, *Journal of Water Resources Planning and Management* (to be published).

Hinomoto, H. (1974). Determination of Investment Cost Functions of Water Treatment Plants, Report No. 87, University of Illinois at Urbana-Champaign.

Kim, J. H., and Larry W. Mays (1994). Optimal Rehabilitation Model for Water Distribution Systems, *Journal of Water Resources Planning and Management,* Vol. 120, No. 5, pp. 674–692.

Lall, U., and L. W. Mays (1981). Model for Planning Water-Energy Systems, *Water Resources Research,* Vol. 17, No. 4, pp. 853–865.

Lasdon, L. S., and A. D. Waren (1986). *GRG2 User's Guide,* Department of General Business, University of Texas at Austin.

Matsumoto, J., and L. W. Mays (1983). Capacity Expansion Model for Large-Scale Water Energy Systems, *Water Resources Research,* Vol. 19, No. 3, pp. 593–607.

Mays, L. W. (1997). *Optimal Control of Hydrosystems,* Marcel Dekker, New York.

Mays, L. W., and Y. K. Tung (1992). *Hydrosystems Engineering and Management,* McGraw-Hill, New York.

Mehrez, A., et al. (1992). Optimal Operation of a Multisource and Multiquality Regional Water System, *Water Resources Research,* Vol. 28, No. 5, pp. 1199–1206.

Niemi, E., and T. McGuckin (1997). Water Management Study: Upper Rio Grande Basin, Final Report to the Western Water Policy Review Advisory Commission, July 1997.

Ocanas, G., and L. W. Mays (1980). Models for Water Reuse Planning, Final Report to Texas Department of Water Resources, CRWR-173, Center for Research in Water Resources, University of Texas at Austin.

Ocanas, G., and L. W. Mays (1981a). A Model for Water Reuse Planning, *Water Resources Research,* Vol. 17, No. 1, pp. 25–32.

Ocanas, G., and L. W. Mays (1981b). Water Reuse Planning Models: Extensions and Applications, *Water Resources Research,* Vol. 17, No. 5, pp. 1311–1327.

Schwartz, M., and L. W. Mays (1983). Models for Water Reuse and Wastewater Planning, *Journal of Environmental Engineering,* Vol. 109, No. 5, pp. 1128–1147.

Selvakumar, A., et al. (2002). Costing Models for Water Supply Distribution System Rehabilitation, *Journal of Water Resources Planning and Management* (to be published).

Wolsey, L. A. (1998). *Integer Programming,* John Wiley and Sons, New York.

Yang, S. L., et al. (1999). A Multicommodity Formulation for Multisource Water Distribution Network, *Proceedings of the 26th Annual Conference of the Water Resources Planning and Management Division of ASCE,* Tempe, Ariz., June 6–9, 1999.

Yeh, W. W-G., et al. (2000). Optimization of Regional Water Distribution System with Blending Requirements, *Journal of Water Resources Planning and Management,* Vol. 126, No. 4, pp. 229–235.

P • A • R • T • 3

PERFORMANCE, RELIABILITY, GIS, OPERATION, AND MAINTENANCE

CHAPTER 9
PERFORMANCE INDICATORS AS A MANAGEMENT SUPPORT TOOL

Helena Alegre
National Civil Engineering Laboratory
Lisbon, Portugal

9.1 INTRODUCTION

The urban water supply combines three main characteristics that places it in a rather peculiar situation as an industry: it is an essential service for the health and welfare of the populations, a natural monopoly, and a relevant economic activity. The former characteristic, which is self-explanatory, provides great political and social importance; conversely, it requires that the service provided fits the population needs on a sustainable basis. It is a natural monopoly because it requires the construction and use of expensive transport infrastructures that cannot be used by other service providers, either of the same or of a different nature; subsequently, market competition rules are not applicable. The latter characteristic—*economic relevance*—derives from the fact that it requires high capital investments and produces significant management and operating profits; it has transformed this activity into an attractive business.

These three characteristics are sometimes conflicting. On the one hand, the sustainable development of the urban water supply services requires an efficient management of the undertakings. The systems and technologies have grown in complexity in recent years, and demand and expectations on the service are ever-increasing. On the other hand, the level of development of the urban water supply activity tends to be a step behind other activities with much lower economic relevance, due to the lack of market competition stimulation derived from the monopolistic nature of the service. With the growth of the private participation, the management objectives move toward the shareholders medium- and long-term recovery of their investments. This may sometimes clash with the consumers' interests.

In summary, there is today a strong call for efficiency and good quality of service. The improvement of performance in terms of organizational and management procedures, aiming to satisfy both public needs and expectations and the environmental policy goals, is therefore a major medium-term target. The development and use of performance indicators is nowadays a hot topic of the water industry agenda worldwide. In fact, the use of standardized performance assessment procedures is a key leverage factor for an enhanced performance, able to identify what activities can be improved, to introduce some artificial competition, and to help in establishing contractual agreements that protect the consumers' interests.

This chapter aims to provide responses for the following questions about performance indicators (PIs):

- What is a PI?
- What are the potential uses and benefits of using a PI system?
- What is the current international state of the art of performance assessment?
- Are there PI systems for water supply services publicly available?
- How is a PI system implemented?

An example is presented based on the PI system recommended by the International Water Association.

9.2 CONCEPT OF PERFORMANCE INDICATOR, CONTEXT INFORMATION, AND UTILITY INFORMATION

Performance indicators are measures of the efficiency and effectiveness of the water utilities with regard to specific aspects of the utility's activity and of the system's behavior (Deb and Cesario, 1997). *Efficiency* is a measure of the extent to which the resources of a water utility are utilized optimally to produce the service, while *effectiveness* is a measure of the extent to which the targeted objectives (specifically and realistically defined) are achieved. Each performance indicator expresses the level of actual performance achieved in a certain area and during a given period of time, allowing for a clear-cut comparison with targeted objectives and simplifying an otherwise complex analysis.

According to the definition recommended by the International Water Association, any performance indicator shall be defined as a ratio between variables of the same nature (e.g., %) or of different natures (e.g., $\$/m^3$ or liters per service connection). In any case, the denominator shall represent a characteristic of the system dimension (e.g., number of service connections or total mains length) to allow for comparisons along time, even though the size of the system evolves, or between systems of different sizes. Variables that may vary substantially from one year to the other, particularly if not under the control of the undertaking, shall be avoided as denominators (e.g., annual consumption, which may be affected by weather or other external reasons), unless the numerator varies in the same proportion.

A global system of performance indicators must comply with the following requirements. It must

- Represent all the relevant aspects of the water utility performance, allowing for a global representation of the system by a reduced number of indicators.
- Be suitable for representing those aspects in a true and unbiased way.
- Reflect the results of the managing activity of the undertaking.
- Be clearly defined, with a concise meaning and a unique interpretation for each indicator.
- Include only nonoverlapping performance indicators.
- Require only measuring equipment that targeted utilities can afford; the requirement of sophisticated and expensive equipment should be avoided.
- Be verifiable, which is especially important when the performance indicators are to be used by regulating entities that may need to check the results reported.
- Be easy to understand, even by nonspecialists—particularly by consumers.
- Refer to a certain period of time (1 year is the basic assessment period of time recommended, although in some cases other periods are appropriate).
- Refer to a well-limited geographic area.
- Be applicable to utilities with different characteristics and stages of development.
- Be as few as possible, avoiding the inclusion of nonessential aspects.

However, the performance indicators cannot be adequately interpreted without taking into account a number of factors that do not depend on managing performance but will affect it. For instance, the cost to produce good-quality drinking water depends on the availability and quality of the raw water, characteristics that are generally beyond the undertaking's control, although something managers must deal with. This context information includes external characteristics (e.g., geographic, demographic, economic, climatic), regarding the region, and internal ones, regarding the water utility and the system on focus (Fig. 9.1). This is particularly important when comparing results with reference numbers from the literature or with other undertakings.

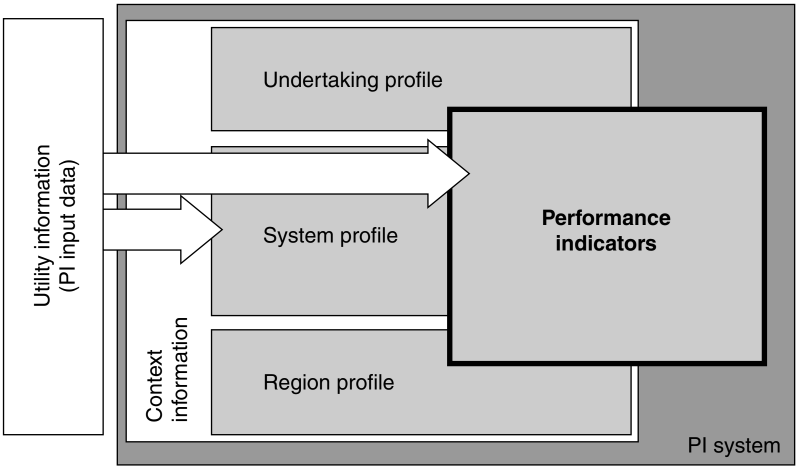


FIGURE 9.1 Performance indicators and context information.

The assessment of PIs requires the availability of reliable utility information. This information can be originated from multiple sources within the organization and be generated in many different ways (e.g., accounting system, metering equipment, information systems). As a general principle, the assessment of the PIs will require only data relevant for the decision making at the various levels within the undertaking, and therefore usually available regardless of its use as PI input. Common characteristics of PI utility data are that they should

- Be absolute values
- Refer to the same period of time and geographic area as the PIs they will be used for
- Be as reliable and accurate as required by the decisions made based on them
- Fit the definition of the PIs they are used for

The utility information used as PI input variables partially overlaps with the context information, particularly with the system profile information.

9.3 USERS, BENEFITS, AND SCOPE OF APPLICATION OF PERFORMANCE INDICATORS

For the managers of water undertakings, demands for higher efficiency and effectiveness can be internal or come from the users of the services (direct users), from the politicians and administrators in the City Hall (indirect users), and from the state regulators and policy-making departments (proactive users). Last, but not least, there is continued and intense awareness from the media and from environmental nongovernmental organizations (NGOs) and other pressure groups (proactive users).

A well-devised system of performance indicators can be useful to any of these entities, having the following potential benefits and uses:

For *water utilities:*

- Facilitates better quality and more timely response from managers
- Allows for an easier monitoring of the effects of management decisions
- Provides key information that supports a pro-active approach to management, with less reliance on apparent system malfunctions (reactive approach)
- Highlights strengths and weaknesses of departments, identifying the need for corrective measures to improve productivity, procedures, and routines
- Assists with implementation of a Total Quality Management regime, as a way of emphasizing all-round quality and efficiency throughout the organization
- Facilitates the implementation of benchmarking routines, both internally, for comparing the performance at different locations or systems, and externally, for comparison with other similar entities, thus promoting performance improvements
- Provides a sound technical basis for auditing the organization's workings and predicting the effect of any recommendations made as a result of an audit

For the national or regional *policy-making bodies:*

- Provides an objective means to monitor the undertaking's performance with regard to existing legislation
- Provides a common basis for comparing the performance of water services and identifying possible corrective measures
- Supports the formulation of policies for the water sector, within the integrated management of water resources, including resource allocations, investments, and the development of new regulating tools
- Provides a common reference language to be adopted in the scope of national or international statistic data banks

For *regulatory agencies:*

- Provides key monitoring tools to help safeguard user interests in a monopoly service situation and monitor compliance with contracted goals

For *financing bodies:*

- Provides assistance in assessing investment priorities, project selection, and follow-up

For *quality certifying entities:*

- Provides key monitoring tools

For *auditors*

- Provides a sound technical basis for auditing the organization's workings and can be used to specify some of the recommendations

For *direct users* and *indirect and proactive stakeholders:*

- Provides the means of translating complex processes into simple-to-understand information and of transmitting a measure of the quality of service provided

For *supranational organizations:*

- Provides a very appropriate language for identifying the main asymmetries between regions of the world, their causes and evolution, thereby assisting in defining strategies

The main forms of using performance indicators are

- *Exclusively within the undertaking.* Managers may use PIs to monitor the evolution of performance, comparing the results achieved at a given time with those achieved in previous periods; another form of use is comparing the results actually achieved with preset targets or with reference values from other organizations that have been published. For the two former types of use managers, the undertaking can afford to define and use its own indicators; for the latter, it has to adopt standardized definitions to prevent comparing "apples with pears."
- *In the framework of benchmarking initiatives.* The internal use of PIs can be complemented with metric benchmarking initiatives; these can be held within a small group of undertakings that can agree among them upon a set of indicators, based on a common set of definitions, to be adopted by all the group members. The analysis, interpretation, and eventual publication of the results must always take into account these definitions.
- *As part of a regulatory framework.* The worldwide growth of the private participation in the management of water supply systems demands added responsibility from regulators. The adoption of PI systems to promote artificial competition among water services (comparative regulation or yardstick competition) is gaining new supporters. Also in this case the use of standardized definitions is essential. Normally the PIs used for regulation focus on the quality of service. The increase of efficiency is of the shareholders own interest, and therefore regulators do not have the need to control the internal processes but rather the results.
- *As part of contractual agreements.* Traditionally, the contract agreements between the municipalities and private operators are rather detailed in terms of the financial aspects, but they tend to be less specific in terms of the quality of

TABLE 9.1 Scope of Application of PI Systems

	Exclusively within the undertaking	In the framework of benchmarking initiatives	As part of a regulatory framework	As part of contractual agreements	In the scope of quality certification systems	Guaranteed standard schemes	In the scope of statistic reports publicly available
Water utilities	✓	✓	✓	✓	✓	✓	✓
Policy-making bodies		✓					✓
Regulatory agencies		✓	✓				
Financing bodies		✓		✓			
Quality certifying entities					✓		
Auditors	✓	✓	✓		✓		
Direct users and indirect and proactive stakeholders		✓				✓	✓
Supranational organizations		✓					✓

service to be provided to the consumers. The use of PIs allows scheduled targets to be specified in an objective and auditable way, protecting the consumers rights. Consequently, PIs are starting to be used in this context. The use of custom PIs is feasible and sometimes necessary, but the use of standardized PIs provides a better assistance for the negotiations, which can be supported on the basis of comparisons with others.

- *As part of quality certification systems.* The monitoring of the organization's internal processes can be undertaken by the use of PIs. As in the previous case, the use of custom PIs is feasible, but the use of standardized PIs is preferable.
- *As part of guaranteed standard schemes.* An increasing number of undertakings have established contracts with their customers that clearly establish their right and duties, based on the so-called guaranteed standards scheme (GSS), which include at least minimum service pressure at the delivery point, maximum time to get a new connection and to repair an existing one, maximum time of written responses, and appointment times to attend customers' premises. In this case the undertakings generally report back to the customers the application of the GSS by monitoring corresponding PIs.
- *In the scope of statistic reports publicly available.* Another form of using PIs is in statistics published by waterworks associations or by regional or international organizations (e.g., World Health Organization, Asian Development Bank, World Bank). Traditionally, this type of performance monitoring was limited to statistical data and economical indicators, but recently reports have started to incorporate other types of indicators. The practice shows that incoherent information is frequently found in these reports, due to the lack of precise definitions. The standardization of definitions is therefore very important in these cases as well.

Table 9.1 defines the main forms of use of a PI system for each type of user.

9.4 STATE OF THE ART OF PERFORMANCE ASSESSMENT

9.4.1 Overview

In the early nineties, the International Water Supply Association (IWSA) selected the topic Performance Indicators for one of its world congresses. No abstracts were submitted under this topic, and it had to be cancelled. The subject did not seem to raise much interest. However, just 3 or 4 years later, the response to an inquiry held in the scope of the IWSA to about 150 senior members of water utilities from all over the world clearly showed that PIs and unaccounted-for-water are by far the two topics of greatest interest in the scope of the water transmission and distribution systems. Such rapid evolution deserves some thought.

Independently from their nature (private, public, or combined) and geographic extent, all water utilities comply with a managing logic whose main philosophy may be stated as follows: *greater satisfaction of a greater number of consumers and concerned entities, with the best use of the available resources* (Faria and Alegre, 1996). In terms of the utilities, this is equivalent to stating *greater efficiency and effectiveness of the management.*

However, the need to improve the efficiency and effectiveness is not new and does not explain by itself the current interest in the assessment of PIs. A very brief analysis of some key events in the world can help to understand this trend.

9.4.2 Influence of the Growing Private Participation in Undertakings Management

It is not the aim of this section to discuss the comparative advantages and disadvantages of the public and private water services. However, it does aim to present the reasons why the use of PIs is becoming more and more important nowadays.

Most water services in the world are public. The main reason for this is the nature of the service: human life requires the consumption of water, and the health of the population is greatly affected by the quality and availability of this water. It is therefore a public service. However, many countries came to the conclusion that the management of public companies is not efficient, as these companies are often underfunded, and there is a need to change the situation. Enhancing the private sector participation was the solution pointed out in many of these cases. An extreme situation of full privatization occurred in England and Wales. Although there are no parallels to this extreme case, examples of this trend can be found in many countries all over the world. The increasing number of systems run by French companies in Europe, America, the Far East, and Africa is a clear demonstration of this trend.

The requirement to report systematically the performance achieved by means of a framework of PIs is not a general practice in most existing cases. The most relevant exception is what happens in England and Wales, where the utilities must assess and publish their levels of service. The contracts focus mainly on financial aspects and on the physical conditions of the assets, and the private companies are subject to the same legislation as the public companies in terms of service requirements. This option goes in line with the tradition in France, where private companies have for many years been responsible for the management of important water supply systems, in coexistence with public companies. Only very recently have these companies started to demonstrate some interest in the use of PIs. However, it has to be taken into account that the French privatization process was very slow, and there was time enough to stabilize the procedures. Nowadays, these processes tend to be much more rapid, there are normally several companies competing when there is a call for tenders for the management of water and sewerage systems, and there is a consensus of the benefits arising from clear rules and objectives. This is another very important reason that explains the increasing interest for PIs.

Last, but not least, many managers of public undertakings want to demonstrate that they are not less efficient and effective than the private companies. The systematic use of PIs is a good means of reaching this target.

9.4.3 Objective-Oriented Management and Benchmarking

Another factor that has greatly influenced this trend is closely related with the evolution of management procedures in most industrial sectors. Management techniques have changed all over the world, and nowadays there seems to be a general agreement that the implementation of objective-oriented management procedures is an indispensable step for the success of most companies. This approach requires the establishment of clear objectives to be achieved within given deadlines, the comparison between targets and results, and the correction of the causes for the deviations, so that the company's efficiency may improve. Performance indicators are a rather powerful tool in this context, as they allow for clear and quantified comparative measures. Some companies have realized that if they compare themselves to the best ones, and correctly identify the reasons for any discrepancies, they can improve their performance significantly. This is how benchmarking appeared and was successfully used in many industrial sectors (for example, photocopying companies, and car manufacturers). Benchmarking is starting to become popular in the water industry as well, and it is obvious that the comparison between different companies requires the use of standardized PIs.

9.4.4 Lessons Arising from the U.K. Privatization Process

It was mentioned earlier that the privatization process held in England and Wales is a rather special one. It has advantages and disadvantages; it has supporters, but also detractors. One of the criticisms often pointed out about the system is the fact that the British companies spend enormous amounts of money in publicity campaigns to improve their external image and keep the customer happy, instead of reinforcing the investment for improving their physical assets. However, it has to be recognized that the history of water services in England and Wales since the 1970s has had a major impact on the water industry in the world. For this reason, its analysis is fundamental for any country interested in enhancing the private sector partnership.

In 1970 the British government decided to merge several hundreds of medium- and small-size water and wastewater companies into 10 water authorities. The scale effect thus achieved allowed the British water industry to undergo a tremendous technical evolution. The new bodies could afford the human expertise and equipment that the previous companies could not reach due to their small size. During this phase many modern management procedures were implemented, and the efficiency and effectiveness achieved were assessed by the use of PIs.

However, many of the existing systems were rather old, heavy investments were required, and the water authorities could not solve the difficulties without external financial support. It was within this context that the government decided to privatize all water and wastewater services in England and Wales in 1990. Regardless of its advantages and disadvantages, this was an unparalleled initiative worldwide, at least to the author's knowledge. It was a major challenge, particularly due to the monopolistic nature of the industry. There had been some unsuccessful experiences with private companies in the English water sector, and the British government, under criticism from many sides, tried to protect itself by creating a legal framework that would ensure an efficient control of the situation (the 1990 Water Act is the key legal instrument of that time). Before the private players started to perform, all the "rules of the game" were defined. New regulatory bodies were created, and the new water service companies had to report regularly the levels of service achieved to the Office of the Water Services (OFWAT). Therefore, a clear and coherent way to assess those levels of service had to be defined. The previous work already developed by some water authorities inspired the new assessment system. Three main alternatives were available:

- Allow the private companies to act just as the previous public authorities, subject only to the existing legislation
- Create mechanisms to control the activity developed inside the company
- Create mechanisms to control the output of the company

The first option, adopted by other countries in privatization processes, was not acceptable due to the reasons previously presented. The second alternative had some major disadvantages:

- The private companies did not accept easily continuous interference in their activities and therefore would not support this option.
- It is a mistake to support any control mechanisms on data that cannot be easily verifiable; in order to control all the internal procedures in a reliable way, the regulators would need to have huge teams, and the system would hardly be effective.

The third alternative, selected by the British government, has some important advantages:

- The companies can freely keep their "business secrets," provided that the service actually delivered is good enough.
- The consumers, together with the media, help to control the whole process, as actual managing partners.

- The number of variables to monitor and control is much smaller and these variables are verifiable by any audit, whenever appropriate.

The decision to limit the reporting requirements to the *output indicators,* in terms of the actual service delivered, is one of the key reasons why the British experience is so relevant to other countries. This means that the companies were free to adopt the management procedures they considered most appropriate, provided that the tariffs were fair and the service provided to the consumers was good enough. This apparently simple decision represents a major milestone in the history of privatization of water supply systems. Nothing similar had been implemented before. This decision was based on the principle that the regulator should focus its attention on the effectiveness of the utility. It is a target of any private agency to improve its efficiency, in order to increase its profits. Therefore, the regulator does not need to care too much about this side of performance, particularly because the physical assets belong to the private managing body. When this is not the case, complementary measures are necessary, such as the need for periodic independent audits of the physical conditions of the assets.

The presentation of the British experience within this context does not mean it is necessarily an example to follow. On the contrary, it is a singular experience, implemented due to a number of circumstances that are unlikely to happen in any other country. Therefore, it cannot be directly exported elsewhere. However, there are some lessons that any country should learn from it, particularly if the country is going through the process of enhancing the private sector participation in the water industry. In the author's view, the major input of the British experience to the outside world is

- The implementation of a clear regulatory framework applicable to all the water service companies, that ensures that the key needs of the population are safeguarded
- The systematic publication of the results achieved with regard to the quality of service provided to the population by each company, contributing to the creation of a new culture of "open doors" and transparency
- The implementation of competition mechanisms in a sector of natural monopoly
- The implementation of procedures for data quality control and the publication of the degree of confidence associated with each result reported
- The decision of controlling only the output performance, leaving companies with a reasonable degree of freedom to implement their own management strategies
- The capacity for focusing the control mechanisms in a relatively small number of PIs
- The recognition of the increasing role and rights of the consumer

The information published yearly by OFWAT is divided into two main reports: Report on Levels of Service for the Water Industry in England and Wales and Report on the Cost of Water Delivered and Sewage Collected. The former report also receives inputs from two other regulatory bodies, the Drinking Water Inspectorate, which controls drinking water quality, and the Environment Agency, which enforces environmental quality standards and reports on compliance, particularly with regard to the sewerage companies. These reports focus on the following PIs, designated by levels of service:

- Properties at risk of low pressure
- Properties subject to unplanned supply interruptions of 12 h or more
- Population subject to flooding incidents
- Properties at risk of flooding (sewerage services)
- Billing contacts not responded to within 5 working days
- Written complaints not responded to within 10 working days
- Bills not based on meter reading[1]

The process held in England and Wales had a great leverage effect worldwide. Most of the other existing initiatives related to performance assessment and regulation got its influence in one way or another.

9.4.5 The IWA PI System

The International Water Association (IWA)—the largest international association in the water and wastewater field, representing about 130 countries—recently developed a system of PIs for water services, which is currently becoming a reference in the industry. This system is a powerful and timely management tool for water services undertakings, independent from their level of development and climatic, demographic, and cultural characteristics of the regions. It aims to cover the full range of management, water resources, personnel, physical, operational, quality of service, and financial PIs. It aims to be a reference PI language covering the full range of management.

The IWA PI system includes water resources, personnel, physical, operational, quality of service, and financial indicators. It also includes the definition of the information required to establish the utility profile, system profile, and region profile. While this system is being implemented and field-tested by almost 70 undertakings worldwide, a similar system is under development for the wastewater services.

The IWA project on PIs began as long ago as 1997 with establishment of an IWSA[2] task force under the operations and maintenance committee (Alegre et al., 1997). The work has since progressed through more than 20 scientific and technical meetings in Europe, South America, and Africa and with the benefit of helpful written comments and suggestions from over 100 experienced managers, practitioners, and researchers in more than 50 countries from the five continents. In July 2000 this task force produced *Performance Indicators for Water Supply Services* in the IWA Manual of Best Practices Series (Alegre et al., 2000). The primary aim is that the manual should be a useful management tool in water undertakings at any stage of development and regardless of local demographic, climatic, or cultural conditions.

Water undertakings are the main users of the manual, but the position of the industry as a monopoly essential service supplier capable of making significant impacts on the natural environment means that the manual is equally capable of use by national and regional policy makers, environmental and financial regulators, watchdog organizations, and the general public as a means of gauging individual aspects of water supplier performance. The manual has been constructed with these needs in mind and consists of a 160-page textbook and a CD-ROM containing the SIGMA Lite software.

The IWA-PI system is currently becoming a reference in the industry (Merkel, 2001). Many undertakings are adopting it directly as an internal management tool. In other cases it was the starting point for performance assessment approaches, as reported by Denmark, the Czech Republic, Australia, Germany, South Africa, Sweden, and Portugal (Merkel, 2001), often in modified and adapted versions. This system is publicly available and is the basic reference for the following sections of this chapter.

9.4.6 The World Bank Benchmarking Toolkit

It was noted previously that PIs are a powerful tool for financial agencies, as a means for assessing investment priorities, project selection, and subsequent investment follow-up. In fact, agencies such as the World Bank and the Asian Bank seem to be fully aware of this fact and started the development and use of PIs long ago. *Water and Wastewater Utilities Indicators* (Yepes and Dianderas, 1996) is a World Bank publication focusing on water and wastewater services indicators and is one of the former relevant publications on this topic.

More recently the World Bank published a new system of indicators aiming to support benchmarking initiatives. The *Benchmarking Water and Sanitation Utilities* (BWSU), designed primarily for developing regions, aims at facilitating the sharing of cost and performance information between utilities and between countries by creating a network of linked web sites, through a global partnership effort. The BWSU includes 27 indicators on water consumption and production, unaccounted-for water, metering practices, pipe network performance, cost and staffing, quality of service, billings and collection, financial performance, and capital investment. In order to assist users in comparing PI values, three explanatory factors—utility size, range of service, and extent of private sector involvement—are also provided (*www.worldbank.org/html/fpd/water/topics/uom_bench.html*). Currently, the project counts around 130 undertakings from Africa, North and South America, Australia, and Europe (Merkel, 2001).

9.4.7 The Asian Development Bank Data Book

In 1997 the Asian Development Bank (McIntosh and Iñiguez, 1997) compared in a well-structured way the performance of 50 utilities of the Asian and Pacific regions. The report is divided into three parts: part I: sector profile; part II: regional profiles; and part III: water utility and city profiles.

A multidimensional analysis procedure was developed: each indicator was analyzed individually for the set of utilities, in terms of trend analysis, and each company and city is presented and analyzed through the whole range of indicators. Information includes both explanatory information and indicators. The indicators used to summarize the results achieved are given in Table 9.2.

TABLE 9.2 Indicators Used by the Asian Development Bank

Indicator	Unit	Indicator	Unit
Private sector participation	Short description	Metering	%
Production/population	m^3/d/c	Operating ratio	—
Coverage	%	Staff/1000 connections ratio	—
Water availability	Hours	Management salary	US$
Consumption	Liters/connection/day	New connection	US$
Unaccounted-for water	%	Accounts receivable	Months
Nonrevenue water	%	Grant financing	%
Average tariff	US$/$m^3$	Commercial financing	%
Water bill	US$/month	Local bond financing	%
Power/water bill ratio	—	Capital expenditure/connection	US$
Public taps	Yes/no	Annual report	None, type script, or glossy covered report

9.4.8 The Water Utility Partnership for Capacity Building in Africa

The *Water Utility Partnership for Capacity Building in Africa* (WUP) is a joint program initiated by the *Union of African Water Suppliers* (*www.uade-wup.org/*)—Abidjan, Côte d'Ivoire, the Regional Center for Low Cost Water and Sanitation (CREPA)—Ouagadougou, Burkina Faso, and the *Center for Training, Research and Networking for Development* (TREND)—Kumasi, Ghana. Though established in 1995, the program was launched in 1996 with the support of the World Bank. The European Union, France, Sweden, and the United Kingdom currently fund WUP as well. The basic idea leading to the creation of the WUP was to build a partnership among *African Water Supply and Sanitation Utilities* (WSSUs) and other key sector institutions to create opportunities for sharing of experiences and capacity building. The WUP intends to achieve its objectives through five specific but very closely linked projects, one of them titled *Service Providers' Performance Indicators and Benchmarking Network* (SPBNET).

The purpose of this project is to provide utilities with sustainable arrangements for compiling and sharing performance data and to develop an understanding of how the data can be used for benchmarking. Twenty-one utilities from 15 countries provided their performance information, and the results have been published by the WUP. The project is now being extended to cover the rest of the utilities in Africa (about 100 committed to participate). Specific activities include

- Development of a software package for use in performance data analysis
- Collection of performance data from utilities
- Production of a data bank on the performance of utilities in Africa
- Training of utilities personnel on benchmarking and its application to the sector

The IWA PI system and the World Bank benchmarking toolkit were the two basic references for the system developed. A spreadsheet application has been developed for data collection and PI assessment.

9.4.9 The Six Scandinavian Cities Group

A similar initiative is under way in the Nordic countries (Adamsson, 1997; Adamsson et al., 2000), involving in a first stage a group of six cities: Stockholm, Gothenburg, Malmö, Copenhagen, Oslo, and Helsinki. Water supply and wastewater services in Scandinavia are normally provided by the local authorities that also operate and own the assets. In small communities in Denmark, Finland, and Norway the water supply is often managed by consumer cooperatives, while the municipality provides sewerage. There is a growing interest in increasing cooperation between neighboring municipalities.

The cooperation among the six cities started in the late seventies between the planning departments in the utilities. It turned out to be very efficient by exchange of experiences and resulted in effective master plans. Later on, overall management issues were discussed and the managing directors were involved and started yearly 2-day meetings with a well-prepared agenda.

In 1995 the six-cities group decided to start a joint cooperation project with the aim to develop PIs that would facilitate comparisons between the cities and give a better base for discussions with the politicians and give impulses for development of the utilities. Their PI system covers

- *Customer satisfaction*—PIs and measuring methods reflecting the customers' expectations and appraisal of the water services
- *Quality*—Quality-related PIs complementing the economical PIs and the customer satisfaction PIs
- *Availability*—PIs describing the reliability in operation of the entire system
- *Environment*—PIs illustrating the utilities environmental achievements
- *Organization/personnel*—PIs describing efficiency and the relation between in-house work and external services
- *Economy*—PIs comparing costs on an overall level

The selected PI structure was based on the objectives, targets, and PIs used in the six cities. The group discussed which PIs could be of interest on the management level and added some new PIs that were considered to be of interest for the overall assessment of the long-term development of the water and wastewater services.

The selected PI set has been tested yearly on the six cities' activities since 1996. The test showed that despite country boundaries, different languages, and different currencies, it was possible to compare the results. The tests have shown that

- Definitions of data and PIs are most important.
- Many of the PIs should include a description of the trend during a 5- to 10-year period.
- Local conditions and accounting principles have a great influence on the PI values.
- During the process of testing, the existing PIs were improved and new PIs developed.
- The interest for working with PIs increased in the utilities.
- The PIs can be used as a supplement to the annual accounts.

This project encouraged the Swedish Water and Wastewater Association to initiate a study with the aim to survey the performance assessment of its members. This system is based on the IWA PI system.

9.4.10 The Dutch Contact Club for Water Companies

A rather interesting example can be found in the Netherlands (van der Willigan, 1997). In 1992, the 13 larger Dutch water companies created a "contact club for water companies" and a common standard to be used for benchmarking has been introduced. These companies agreed to define indicators to assess their own performance and compare the results achieved once a year. There are no goals set or yearly targets. According to van der Willigen (1997), the following main categories of performance indicators are being considered:

- *General*—Includes connections, personnel, and length of main pipes
- *Production*—Includes internal production, external production, external delivery, delivery to distribution, unaccounted-for water, net sales, sales in volume, and sales per connection
- *Costs*—Includes production, distribution, sales, general, total, income, and result and is broken down into costs total, per m^3 sales, and per connection
- *Personnel*—Includes connection per person, production, distribution, sales, general, salary per person, absenteeism, short-term absenteeism, and long-term absenteeism

9.4.11 An Engineering Approach for Performance Assessment

Coelho (1997) and Cardoso et al. (2000) address the engineering side of the problem by using a flexible framework based on an array of performance indices. The performance evaluation system they developed is a technical analysis tool, designed to shift the focus of management and operations of water systems to a wider, more rigorous, performance-oriented view. It is based on a system of penalty curves, flexible enough to accommodate individual sensitivities and interpretations. The methodology allows for a rapid gain in sensitivity to the network's behavior, providing a standardized means of diagnosis, and is a valuable aid for planning, design, operation, and rehabilitation of water distribution systems.

Similarly to the other type of performance indicators, these also contemplate quantitative indicators based on the analysis, from specific viewpoints, of the network's characteristics and behavior. However, the former type of performance analysis is mainly based on sound records, such as damage rate or maintenance costs, or on data obtained from systematic or ad hoc field surveys. This latter type requires the use of hydraulic and water quality dynamic simulation tools, in order to infer the system's behavior if a given remedial solution was implemented.

The performance indicators developed by the author are classified under three generic terms, although the methodology is applicable to other viewpoints:

- *Hydraulic performance*—Focuses on pressure head and pressure head variation
- *Water quality performance*—Focuses on concentrations and travel times
- *Reliability performance*—Focuses on network topological reliability

The limit values for every one of these performance indicators must be defined case by case by the undertaker, according the local characteristics, legal constraints, and its management strategy.

The performance evaluation can be accessed on an empirical and nonsystematic basis or in a more systematic way. In this latter situation, three types of entities, established for each aspect to be analyzed, can define the framework:

- A relevant *state variable,* that is, the quantity that translates the network behavior or properties at the network element level, from the point of view taken into consideration. The network element is the node or the pipe, as conventionally used in network analysis, from which all modeling assumptions are inherited. Examples of such state variables are nodal head, to check compliance with the pressure requirements; damage rate, to be compared to reference values defined for each type of material and class of diameter; or nodal concentration of a particular constituent as compared to the water quality guidelines.
- A *penalty function,* which scores the values of the state variable against a scale of index values (6—optimum through 0—no service). The penalty curves are designed by the user and translate as much as possible to a commonsense grading of performance in the addressed domain.
- A *generalizing function,* used for extending the element-level calculation across the network, producing zonal or networkwide diagnoses. The indices are intended to have both local and networkwide meaning. In the case, for example, of a pressure index, the proposed generalizing function is the weighted average across the network. Other types of operators may be used, such as those that focus on maximum or minimum values.

The technique is applied both to extended period operational scenarios and to a range of load factors, displaying appropriate *dispersion bands* (25 percent stepped percentiles).

The above considerations about the nature of the indices imply the need for a careful selection of the variables that may be eligible for such a treatment. In hydraulic terms, for instance, the variables used to assess a system's energy performance are total pumping energy, total energy dissipated, and the difference between the actual potential energy and the minimum potential energy required for supplying every node with the minimum allowable pressure.

9.4.12 Status of Performance Assessment in the United States

From the applied research point of view, there are two projects launched by the American Water Works Research Foundation that became an international reference. The former, a report entitled Distribution System Performance Evaluation, defines distribution system performance measures, develops measurement procedures and techniques, and establishes national target levels for these measures. It also develops guidelines for utility managers to routinely assess the overall condition of their distribution system in order to identify system and investment needs (AWWARF, 1995). The latter, a report entitled Performance Benchmarking for Water Utilities, determines areas of water utility operations and management that would be appropriate for developing benchmarks to evaluate utility performance. It includes suggestions that water utility managers would be able to use to evaluate and improve the performance of their individual operations (AWWARF, 1996).

From the practical point of view, and according to Paralez (2001), 53 percent of the 900 utilities who replied to a questionnaire held by the Government Accounting Standards Board (GASB survey) in 1998–1999 use PI measures. However, there is a clear gap among those who assess performance measures and those who use them as part of the decision-making process. In fact, only about half of those who have PIs report them to elected officials, and only 25 percent report outcome measures to internal management. A more recent custom survey (5 percent of respondents) confirms the same trend. A good number of undertakings state that their measurement efforts do not enjoy practical use or are not accepted by the staff. Parelez (2001) considers that the following issues need to be actively addressed in order to overcome the shortcomings identified:

- Those that gather data should be those using the data.
- Field staff (labor) should be involved in both gathering and using the data; i.e., maintenance, repair, refurbishing, decision making.
- The distribution and understanding of performance data should be enhanced.
- An added effort is needed for the integration of information systems.
- Maximizing data availability and use should be emphasized.
- Appropriate quality control should be applied to all data.
- The acceptability and practical use of information should be encouraged.

9.4.13 Other Initiatives

Many water utilities are already using their own PIs. However, the most interesting experiences to be analyzed from the conceptual viewpoint are the ones regarding groups of utilities. Beyond the benchmarking processes previously referred to, such as in the Nordic countries and in North America (involving the United States and Canada), others are currently starting to be developed and implemented in other areas of the world. In Africa, for instance, a benchmarking process conducted by Umgeni Water, from South Africa, has already produced its results.

A study that cannot go without reference is a report published in 1996 by the *Malaysian Water Association* (MWA, 1996). It is an excellent document that suggests that the PIs are organized under three generic terms, as follows:

- *Physical*—Gives an indication of the size and coverage of the physical aspect of the water supply and the extent of the underground assets and is useful for future planning.
- *Service*—Focuses on consumer services and productivity in an environment where consumers' expectations are increasing; they measure the utility's output against staff resources, assessment of level of service such as staff ratios, interruptions to mains supply, quality of water, water losses, and consumers' complaints on an annual basis.
- *Financial*—Assesses the financial, human, and physical aspect of efficiency, covering costs and the economic utilization of assets, work force, plant, and equipment.

The MWA continued its activity within this topic, publishing a report in 2001 (MWA, 2001). The MWA is also participating in an IWA PI field-testing project with the case study of Penang.

Also from the scientific point of view the topic of PIs has been attracting the attention of researchers in different places of the world. Apart from the studies previously mentioned, several Ph.D. theses have already been published on the topic. This is, for instance, the case of Ancarani (1999), who focused on the economic aspects of the performance assessment, of Guérin-Schneider (2001), a very good work on the use of PIs from the regulatory point of view, and of Cabrera Jr. (2001), who designed a system of performance assessment embracing a methodology for metric and process benchmarking. The two latter studies have a very close link to the IWA PI system.

One aspect of performance assessment and benchmarking is communication. The Water Research Centre, in the United Kingdom, publishes periodically the on-line newsletter *Watermarque* (*Watermarque Magazine: Your Guide to Water and Wastewater Industry Benchmarking*). It is issued bimonthly via e-mail, provides updates on global regional, national, and company initiatives; reports on case studies and on benchmarking awards; provides conference details; and includes an on-line searchable archive of back issues. It is published six times per year for subscribers. Subscription is free at *www.wrcplc.com/pbngroup/pbngroup.nsf/htmlmedia/index.html.*

9.4.14 Summary of Performance Assessment Projects for Water Supply (WS) and Wastewater (WW) Services

Table 9.3 summarizes the main international and national initiatives related to performance assessment. It includes not only the initiatives reported in the previous section but also other ongoing initiatives. It is partially based on the information provided by Merkel (2001).

TABLE 9.3 Summary of Performance Assessment Projects for Water Supply (WS) and Wastewater (WW) Services

Organization/country	Scope	Projects/products
		International initiatives
IWA	WS	*Performance Indicators for Water Supply Services,* Manual of Best Practices Series (published in July 2000).
		http://www.iwapublishing.com/template.cfm?name = isbn1900222272
		http://www.sigmalite.com
		International field test of the IWA PI system (2000–2003).
IWA	WW	*Performance Indicators for Water Supply Services,* Manual of Best Practices Series (in preparation; to be concluded in Dec. 2002).
Asian Develop. Bank	WS	Report: *Second Water Utilities Data Book–Asian and Pacific Region (1997)*
The World Bank	WS+WW	Project: "Benchmarking Water and Sanitation Utilities."
		http://www.worldbank.org/html/fpd/water/topics/uom_bench.html
WUP	WS+WW	Service Providers' Performance Indicators and Benchmarking Network (SPBNET): Questionnaire in MS Excel format with 184 data entries for the assessment of PI indicators of about 100 African utilities (*http://www.wupafrica.org/*).
World Health Organization	WS+WW	2000 Water and Sanitation Assessment: the WHO and UNICEF Joint Monitoring Program for Water Supply and Sanitation (JMP) provides a snapshot of water supply and sanitation worldwide at the turn of the millennium using information available from different sources. From 2001 the JMP database—for both historic data and future projections—will be periodically updated.
		http://www.who.int/water_sanitation_health/Globassessment/GlobalTOC.htm
ISO	WS+WW	The International Standardization Organization established in late 2001 is a work group for the preparation of standards for performance assessment of water supply and wastewater services.

Six Scandinavian cities group	WS+WW	Six-cities PI group: Copenhagen, Gothenburg, Helsinki, Malmö, Oslo, and Stockholm established a systematic routine for PI comparison: in January the group meets and discusses the work plan for the year; in February the process starts with sending out forms for collection of data for the previous year; during March/April, data is collected and the new PI values are assessed and analyzed; in May a preliminary report is produced; in June the reports are discussed within the PI group; at the yearly six-city meeting in September/October, the report and proposals for continued work are presented for the managing directors of each utility.
		National initiatives
Argentina	WS+WW	Ongoing project for the establishment of a common platform of PI for WS and WW to be adopted by the Argentinean regulators, as a basis for yardstick competition; other countries of the region are starting to cooperate as well.
		Participation in the IWA-PI field testing (ETOSS, La Plata, Trelew, and Tucumán).
		Participation in the core team of the IWA-PI-WW project.
Australia	WS+WW	Annual performance reports from WSAA, NMU (AWA), NSW, Vicwater, Queensland Performance Monitoring System, ANCID (e.g., AWA, 2000).
Brazil	WS+WW	ABES Quality National Award: this award, created in August 1995, aims to stimulate the practice of managerial models compatible with the world trends, acknowledgment of successful experiments which utilize the methodology and promote the exchange of the best practices, making possible the quality improvement in the life of Brazilian populations through the development of the sanitation sector (*http://www.pnqs.com.br/english/index.html*).
		Participation in the IWA-PI project (SABESP—São Paulo, COPASA—Belo Horizonte, and CAESB—Brasília).
Czech Republic	WS	Case study in 1997–1999 based on a preliminary version of the IWA PI system (WS).
Germany	WS+WW	Three annual performance assessment projects (WS): 30 large-scale, 270 medium or scale, 20 bulk water suppliers; IWA field test (WS) with 13 utilities, Bavarian WS performance assessment project with around 100 utilities, several WW benchmarking projects (e.g., AV+EG/LV, VkY, and others).
Denmark	WS+WW	Participation in the six-Scandinavian cities group, in the seven-Danish cities group (WS), 18 Danish wastewater plants (WW), PA project of DWSA and DWWA.

TABLE 9.3 Summary of Performance Assessment Projects for Water Supply (WS) and Wastewater (WW) Services (*Continued*)

Organization/country	Scope	Projects/products
		National initiatives
France	WS+WW	Quality management projects (WS); participation in the IWA-PI project (SAGEP-Paris); workshops on performance assessment; Ph.D. thesis on the use of PI for regulation in France (Guérin-Schneider, 2001); development of national standards (AFNOR) on performance and quality of service.
Finland	WS+WW	Participation in the six Scandinavian cities group and in the 10 Finnish Water and Wastewater Utilities.
Italy	WS	Benchmarking initiative of Italian waterworks under FEDERGASAQUA; participation in the IWA PI project (Turin, SIDRA, AGAC, and SOGEAS).
Malaysia	WS	*Performance Indicators for Water Supply* (MWA, 1996) and *Malaysia Water Industry Guide 2001* (MWA, 2001).
Mexico	WS	Project of the IMTA, Mexican Institute of Water Technology: the objective is the definition of a short number of indicators and of applicability subsets, depending on the size and level of development of the utility.
Netherlands	WS+WW	Reports produced by all Dutch WS companies (annual reports include only economic indicators; triannual reports include water quality, service, and environmental and economic indicators); first pilot projects on performance assessment for WW services held in 2001.
Portugal	WS+WW	PI-Waters project (National Civil Engineering Laboratory, LNEC): the objectives are: (i) to reinforce the implementation of the IWA PI system at a national level; (ii) to test the PI application within different national contexts; (iii) to share experiences at the national level; (iv) to promote a broader application of the PI system by the other Portuguese water utilities. Products: Portuguese versions of the IWA PI manual and of SIGMA software.
		Participation in the core team of the IWA PI project.
		Participation in the IWA PI project with 17 undertakings from Algarve, Cávado river catchment, Douro region, Águas de Portugal, S.A. (public holding), Barreiro, Beja, Castelo Branco, Esposende,

		Figueira da Foz, greater Lisboa region, Loures, Luságua (private holding), Oeiras, Amadora, Sintra, Vila Nova de Gaia, and IRAR, the national regulator.
Poland	WS+WW	Quality management projects, private “leadership” projects (e.g., SNG).
Romania	WS+WW	A set of 19 PI is being assessed systematically for some Romanian cities (Cluj, Constanta, Arad, Oradea, Bucharest).
Spain	WS+WW	Development of the software SIGMA for the assessment of the IWA PIs.
		Design of System for the Management Assessment of Urban Water Supply Systems (Cabrera, 2001); participation in the core team of the IWA PI project.
		Participation in the IWA PI project with 12 undertakings (AEAS VI, Alicante, Barcelona, Castellón, Córdoba, Huelva, Madrid, Murcia, Sevilla, Valencia, Vigo, and Zaragoza).
South Africa	WS + WW	Annual reports from the Department of Water Affairs (WS + WW), ZA Assn. of Water Boards (WS), Water Research Commission (Guidelines for Benchmarking Water and Sanitation Services for Local Authorities).
		Participation in the IWA PI project with Rand Water (Johannesburg)
Sweden	WS + WW	Six Scandinavian cities group, DRIVA, VASAM, VA-Plan 2050.
		Participation in the IWA PI project with Stockholm and Gothenburg.
United Kingdom	WS + WW	OFWAT performance assessment projects and periodic reports (e.g., OFWAT, 2000).
		Participation in the IWA PI project with Bristol, Hatfield, North East and East of Scotland.
United States	WS + WW	Several projects sponsored by the AWWARF (AWWARF, 1995 and 1996; Deb and Cesario 1997) and WERF, initiatives of utilities (e.g., WRWBG). Initiatives to improve effectiveness and efficiency (Westerhoff et al., 1998).

9.5 THE IWA PI SYSTEM FOR WATER SUPPLY SERVICES

9.5.1 Highlights of the IWA PI System

In order to meet the criterion that the PI system be independent of the state of development or internal organizational structure of any particular undertaking, the indicators have been framed in terms of the principal management objectives common to all undertakings. These are structured in six separate categories of performance: 2 water resources indicators, 22 personnel indicators, 12 physical indicators, 36 operational indicators, 25 quality-of-service indicators, and 36 financial indicators.

In recognition of the differing states of development and management sophistication that may apply in different undertakings, the manual includes a tentative grading of level of priority of implementation, to be improved during the ongoing field test. However, this is just a general guidance, and it is up to each undertaking to define the importance and applicability of every indicator for the organization. Any subsets can easily be selected, according to the utility's needs and objectives. Conversely, if the degree of detail in considered insufficient, users are naturally free to split the existing indicators in subcategories or add their own indicators, provided that they keep in mind that these new PIs are not standard and therefore are not suitable for future comparisons among undertakings. However, as noted in Sec. 9.2, the interpretation of an undertaking's performance cannot be assessed without taking its own context into account, as well as the most relevant characteristics of the system and the region. This is why the IWA PI system includes the definition of context information, structured within the profiles of the undertaking, system, and region.

These profiles contain the figures that many good managers know by heart and frequently use to present their companies. The undertaking profile outlines the framework of the organization. The system profile focuses mainly on the water volumes managed, on the physical assets, on the technological means used, and on the customers. The region profile will be relevant for comparisons between undertakings, allowing for a better understanding of the demographic, economic, geographic, and environmental context.

9.5.2 Listing of the IWA PIs and Guidance on Their Relative Importance

The current stage of development aims to include the PIs relevant at the top-management level of a water supply undertaking. Those aim to incorporate all the relevant aspects required to express management objectives and results in terms of an undertaking's performance.

The tentative grading of level of priority of implementation referred to in Sec. 9.4 includes three categories, plus one category of complementary indicators. Complementary indicators may be needed only at the departmental level and are generally much more organization dependent and therefore are not included in the IWA PI system.

- *Level 1 (tagged "L1").* A first layer of indicators that provide a general management overview of the efficiency and effectiveness of the water undertaking.
- *Level 2 (tagged "L2").* Additional indicators, which provide a better insight than the level 1 indicators for users who need to go further in depth.
- *Level 3 (tagged "L3").* Indicators that provide the greatest amount of specific detail but are still relevant at the top-management level.
- *Complementary (not included in this report).* Further indicators, which provide a greater amount of specific detail than level 3 indicators or which are for specific use at the departmental level (which tend to be undertaking dependent).

Indicators that conceptually have a given level of importance but that are not easily assessed in a reliable way are classified in a lower-level grade. This full listing of indicators and of the respective preliminary guidance of level of implementation priority is shown in Tables 9.4 to 9.9.

9.5.3 Data Reliability and Accuracy

The methodology proposed by the IWA was adopted first in England and Wales by the OFWAT, where economic regulation of the water supply companies is highly structured and detailed. As the authors of the manual note, within this structure is a carefully developed scheme for grading the confidence that can be placed on the reliability and accuracy of inputs. The testing results currently available seem to demonstrate that this confidence-grading scheme is adequate for any other country.

The confidence-grading scheme includes a four-grade measure of the reliability of the data (sound records, analysis result, forecast, etc.) and a seven-grade classification for their respective accuracy. The interpretation to be adopted for the questions on data reliability and data accuracy is shown in Table 9.10.

TABLE 9.4 Water Resources Indicators

Indicator	Unit	Suggested level
Inefficiency of use of water resources	%	L1
Resources availability ratio	%	L2

TABLE 9.5 Personnel Indicators

Indicator	Unit	Suggested level
Total personnel		
Employees per connection	No./1000 connections	L1
Personnel per main function		
Management and support personnel	No./1000 connections	L2
Financial, commercial personnel	No./1000 connections	L2
Customer service personnel	No./1000 connections	L2
Technical activities personnel	No./1000 connections	L2
Planning & construction personnel	No./1000 connections	L3
Operations & maintenance personnel	No./1000 connections	L3
Water resources, catchment, and treatment personnel	No./10^6 m^3/year	L3
Transmission, storage, and distribution personnel	No./10^2 km	L3
Laboratory personnel	No./1000 tests	L3
Meter maintenance personnel	No./1000 meters	L3
Other personnel	No./10^6 m^3/year	L3
Personnel qualification		
University degree personnel	%	L3
Personnel with basic education	%	L3
Other personnel	%	L3
Personnel training		
Total training	Days/employee/year	L3
Internal training	Days/employee/year	L3
External training	Days/employee/year	L3
Personnel health and safety		
Working accidents	No./employee/year	L3
Absenteeism	Days/employee/year	L3
Absenteeism due to working accidents or disease	Days/employee/year	L3
Absenteeism due to other reasons	Days/employee/year	L3

TABLE 9.6 Physical Indicators

Indicator	Unit	Suggested level
	Treatment	
Treatment availability	%	L1
	Storage	
Impounding reservoir capacity	Days	L2
Transmission and distribution storage capacity	Days	L2
	Pumping	
Standardized energy consumption	Wh/m^3 at 100 m	L2
Reactive energy consumption	%	L3
Energy recovery	%	L3
	Transmission and distribution network	
Valve density	No./km	L3
Hydrant density	No./km	L3
Meters		
District meter density	No./1000 service connections	L3
Customer meter density	No./service connection	L2
Metered customers	No./customer	L3
Metered residential customers	No./customer	L3

The confidence grades will be an alphanumeric code, which couples the reliability band and the accuracy band, for instance:

A2. Data based on sound records, etc. (highly reliable, band A), which is estimated to be within ±5 percent (accuracy band 2).

C4. Data based on extrapolation from a limited sample (unreliable, band C), which is estimated to be within ±25 percent (accuracy band 4).

The reliability and accuracy bands would form the matrix of confidence grades shown in Table 9.11.

9.5.4 Organization of the IWA PI Manual

The structure of the IWA PI manual is shown in Fig. 9.2. Sections 1 to 3 of the document cover general matters such as concepts and uses of PIs, structure of the

TABLE 9.7 Operational Indicators

Indicator	Unit	Suggested level
Inspection and maintenance		
Pumping inspection	%/year	L2
Storage tank cleaning	%/year	L2
Network inspection	%/year	L2
Leakage control	%/year	L2
Active leakage control repairs	%/year	L2
Hydrant inspection	%/year	L3
Instrumentation calibration		
System flow meters	%/year	L3
Meter replacement	%/year	L2
Pressure meters	%/year	L3
Water-level meters	%/year	L3
On-line water quality monitoring equipment	%/year	L3
Electrical equipment inspection		
Electrical equipment inspection by number	%/year	L3
Electrical equipment inspection by power	%/year	L3
Vehicle availability	Vehicles/km	L3
Mains, service connection, and pumps rehabilitation		
Mains rehabilitation	%/year	L1
Mains relining	%/year	L2
Replaced or renewed mains	%/year	L2
Replaced valves	%/year	L2
Service connection rehabilitation	%/year	L1
Pumps rehabilitation		
Pump refurbishment	%/year	L2
Pump replacement	%/year	L2
Water losses		
Water losses	m^3/connection/year	L1
Apparent losses	m^3/connection/year	L3
Real losses	Liters/connection/day when system is pressurized	L1
Infrastructure leakage index	—	L3

TABLE 9.7 Operational Indicators (*Continued*)

Indicator	Unit	Suggested level
Failures		
Mains failures	No./100 km/year	L1
Service connection failures	No./1000 connections/year	L1
Hydrant failures	No./1000 hydrants/year	L2
Power failures	Hours/pumping station/year	L2
Metering		
Customer reading efficiency*	%	L1
Residential customer reading efficiency[3]	%	L1
Water quality monitoring		
Tests performed (quantity compliance)	%	L1
Aesthetic	%	L2
Microbiological	%	L2
Physical-chemical	%	L2
Radioactivity	%	L3

*These indicators shall be used in alternative.

manual, and, separately, a detailed list and full explanation of the definitions used in the subsequent text. Special attention is paid to the definitions related to water balance, main functions of the undertaking, and financial terminology and conventions.

Section 4 is titled "Data Reporting" and considers the means of assessing the reliability of the contextual and PI information, by use of the confidence-grading scheme presented above.

Section 5 describes the operating context under three headings, profiling the undertaking itself, its operating systems, and the operating region.

Section 6 of the IWA manual is devoted entirely to the tabulation of all 133 suggested indicators together with their respective priority level, base concept, and processing rule in the form of a simple equation. The background data definitions and processing rules are very fully elaborated in an appendix, which accounts for over half of the entire manual. Figures 9.3 and 9.4 show the organization of this chapter, containing the identification, classification, definition, and processing rules of the PIs. Figures 9.5 and 9.6 show the organization of the appendix of the IWA manual, containing the definition of every variable required.

TABLE 9.8 Quality of Service Indicators

Indicator	Unit	Suggested level
Service		
Households and businesses supply coverage*	%	L1
Buildings supply coverage*	%	L1
Population coverage*	%	L1
With service connections	%	L2
Public taps and standpipes	%	L2
Public taps and standpipes		
Distance to households	m	L1
Quantity of water consumed	Liter/person/day	L1
Population per public tap or standpipe	Persons/tap	L2
Pressure of supply adequacy	%	L2
Continuity of supply	%	L1
Water interruptions*	%	L2
Interruptions per connection*	No./1000 connections	L2
Population experiencing restrictions to water service*	%	L2
Days with restrictions to water service*	%	L2
Quality of supplied water (quality compliance)	%	L1
Aesthetic	%	L2
Microbiological	%	L2
Physical-chemical	%	L2
Radioactivity	%	L3
New connection efficiency	%	L2
Connection repair efficiency	%	L2
Customer complaints		
Service complaints	No. complaints/ connection/year	L1
Pressure complaints	%	L2
Continuity complaints	%	L2
Water quality complaints	%	L2
Restrictions or interruptions	%	L2
Billing complaints	No. complaints/ customer/year	L1
Other complaints and queries	No. complaints & queries/customer/year	L2
Response to written complaints	%	L2

*These indicators shall be used in alternative.

TABLE 9.9 Financial Indicators

Indicator	Unit	Suggested level
Annual costs		
Unit total costs	US$/m^3	L2
Unit running costs	US$/m^3	L1
Unit capital costs	US$/m^3	L3
Composition of annual running costs per type of costs		
Internal work force costs ratio	%	L3
External services costs ratio	%	L3
Imported (raw and treated) water costs ratio	%	L3
Energy costs ratio	%	L3
Other costs ratio	%	L3
Composition of annual running costs per main function of the water undertaking		
Management and support costs ratio	%	L3
Financial and commercial costs ratio	%	L3
Customer service costs ratio	%	L3
Technical services costs ratio	%	L3
Composition of annual capital costs		
Depreciation costs ratio	%	L3
Interest expenses costs ratio	%	L3
Annual revenue		
Unit annual revenue	US$/m^3	L2
Sales revenues	%	L2
Other revenues	%	L2
Annual investment		
Average unit investment	US$/m^3	L2
Average annual investments for new assets	%	L3
Average annual investments for assets replacement	%	L3
Average water charges (without public taxes)		
Average water charges for direct consumption	US$/m^3	L1
Average water charges for exported water	US$/m^3	L1
Efficiency indicators		
Total cost coverage ratio	—	L1
Operating cost coverage ratio	—	L1
Delay in accounts receivable	Months equivalent	L2
Investment ratio	%/year	L2

TABLE 9.9 Financial Indicators (*Continued*)

Indicator	Unit	Suggested level
Contribution of internal sources to investment = CTI	%	L1
Average age of tangible assets	%	L2
Average depreciation ratio	%	L3
Late payments ratio	%	L2
Leverage indicators		
Debt service coverage ratio = DSC	%	L2
Debt equity ratio	—	L2
Liquidity indicator		
Current ratio	—	L1
Profitability indicators		
Return on net fixed assets	%	L2
Return on equity	%	L2
Water losses indicators		
Nonrevenue water by volume	%	L1
Nonrevenue water by cost	%	L3

The main text concludes with a section on the relative importance of the PIs to water undertakings, regulators, and users. All 133 indicators are retabulated into the appropriate priority levels for each of the separate user groups.

9.5.5 A Partial In-depth Look into the IWA PI System

Viewpoint. One of the characteristics of the IWA PI system that is frequently pointed out as a comparative advantage to other existing performance assessment systems is the data structure adopted and the level of detail of the indicator and variable definitions. This section aims to illustrate this with a specific example.

As water is the product delivered to the undertaking customers, a proper monitoring of the water balance and the reliable assessment of performance in terms of water losses is fundamental to support managers' decisions. This was the topic chosen to provide an in-depth perspective of the IWA PI system.

Definition of Water Supply System Inputs and Outputs.[3] A well-established water balance is fundamental for water losses assessment. The definitions, terminology, and choices of PIs for water losses used in this document are based

TABLE 9.10 Definition of the Confidence-Grading Scheme

Data reliability	Definition
A—Highly reliable	Data based on sound records, procedures, investigations, or analyses that are properly documented and recognized as the best available assessment methods.
B—Reliable	Generally as in band A, but with minor shortcomings; e.g., some of the documentation is missing, the assessment is old, or some reliance on unconfirmed reports or some extrapolations are made.
C—Unreliable	Data based on extrapolation from a limited sample for which band A or B is available.
D—Highly unreliable	Data based on unconfirmed verbal reports and/or cursory inspections or analysis.
Data accuracy	**Definition**
1—Error (%): [0; 1]	Better than or equal to ±1%
2—Error (%): [1; 5]	Not band 1, but better than or equal to ±5%
3—Error (%):[5; 10]	Not bands 1 or 2, but better than or equal to ±10%
4—Error (%): [10; 25]	Not bands 1, 2, or 3, but better than or equal to ±25%
5—Error (%): [25; 50]	Not bands 1, 2, 3, or 4, but better than or equal to ±50%
6—Error (%): [50; 100]	Not bands 1, 2, 3, 4, or 5 but better than or equal to ±100%
Error (%): >100	Values which fall outside the valid range, such as >100%

substantially on the work of the IWA Task Force on water losses, with some minor adaptations to allow them to be fitted within the overall context of this wider PI report. The reading of the Blue Pages on Losses from Water Supply Systems: Standard Terminology and Performance Measures (published electronically by the IWA; available free of charge for members) is recommended.

TABLE 9.11 Matrix of Confidence Grades

	Reliability bands			
Accuracy bands, %	A	B	C	D
[0; 1]	A1	++	++	++
[1; 5]	A2	B2	C2	++
[5; 10]	A3	B3	C3	D3
[10; 25]	A4	B4	C4	D4
[25; 50]	++	++	C5	D5
[50; 100]	++	++	++	D6

Note: ++ indicates confidence grades that are considered to be incompatible.

IWA Manual of Best Practice—Table of Contents

1. Introduction
2. About this document
3. Definition
4. Data reporting
5. Context information
6. Performance indicators
7. Relative importance of the performance indicators

Appendix. Data definition and processing rules

FIGURE 9.2 Structure of the PI manual.

6.4 Operational indicators

INDICATOR *Level of importance* (*)	(unit)	CONCEPT	Processing rule
INSPECTION AND MAINTENANCE			
Op1 - Pump inspections *L2*	(%/year)	Σ (nominal power of pumps and related ancillaries subjected to *inspection* during the year) / Σ (nominal power of pumps) x 100	Op1 = D5/C5 x 100
Op2 - Storage tank cleaning *L2*	(%/year)	Volume of storage tank cells cleaned during the last five years /total volume of storage tank cells x 100 / 5	Op2= D6/C2 x 100
Op3 - Network inspection *L2*	(%/year)	Length of transmission and distribution mains where at least valves and other fittings were *inspected* during the year / total mains length x 100	Op3 = D7/C6 x 100
Op4 - Leakage control *L2*	(%/year)	Length of mains subject to *active leakage control* / total mains length x 100	Op4 = D8/C6 x 100
Op5 - Active leakage control repairs *L3*	(%/year)	Number of leaks detected and repaired due to *active leakage control* / total mains length x 100	Op5 = D9/C6 x 100
Op6 - Hydrant inspection *L3*	(%/year)	Number of hydrants *inspected* during the year / total number of hydrants x 100	Op6 = D10/C31 x 100
		...	

FIGURE 9.3 General outline of the PI definitions within the manual.

Figure 9.7 illustrates the principal inputs and outputs to a typical water supply system sequentially, from raw water intake to consumption by customers. Some systems will of course be simpler and will not have all the features shown.

The water balance requires estimates of volumes of water to be made at each measurement point applicable to the system under consideration. Where there are actual meters, data from these would normally be used, but in the absence of meters, a "best estimate" based on other related available data and application of sound engineering judgment may be required. The water balance is normally computed over a 12-month period and thus represents the annual average of all components.

Definitions relating to Figs. 9.7 and 9.8, are given below. Because of widely varying interpretations of the term *unaccounted-for water* (UFW) worldwide, the IWA does not recommend use of this term. If the term is used at all, it should be defined and calculated in the same way as *nonrevenue water* in Fig. 9.8.

Water abstracted The annual volume of water obtained for input to water treatment plants (or directly to the transmission and distribution systems) that was abstracted from raw water sources.

Raw water, imported or exported The annual volumes of bulk transfers of raw water across operational boundaries. The transfers can occur anywhere between the abstraction point and the treatment plants.

Treatment input The annual volume of raw water input to treatment works.

Water produced The annual volume of water treated for input to water transmission lines or directly to the distribution system. (The annual volume of water that is distributed to consumers without previous treatment shall also be accounted for in *water produced*.)

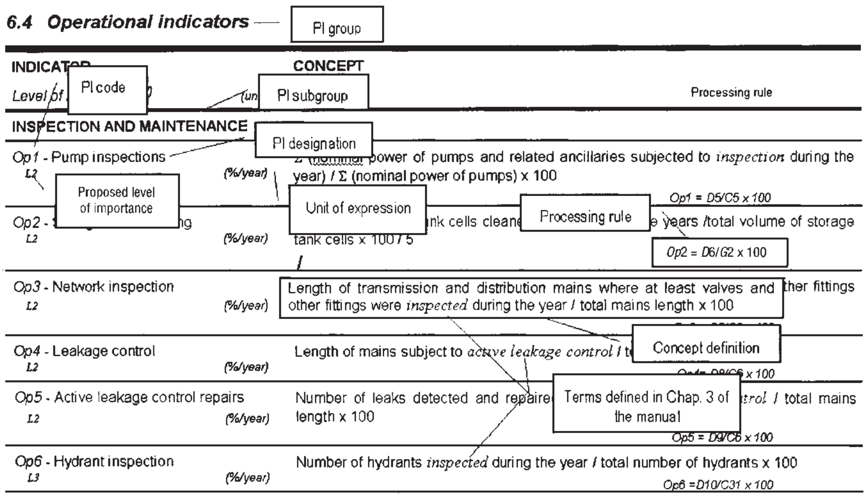


FIGURE 9.4 PI identification and classification, description, terms, and processing rules.

A5 - IMPORTED RAW WATER		
UNIT OF EXPRESSION: m^3/day	PERIOD: *[dd.mm].yy-1 – [dd.mm].yy*	VALID VALUES: ≥ 0
DEFINITION: Total annual volume of raw water transferred from other water supply systems.		
PROCESSING RULE: Input data (Reliability: *[targeted reliability]* ; Accuracy: *[targeted accuracy]*)		
COMMENT:		
RESPONSIBILITY: *[service responsible for this data (e.g. Operational indicators team)]*		
USED FOR VARIABLES: *[A2], A20, A24 (A20), A26*		
USED FOR INDICATORS: *[WR1 (A2)], WR2, Ph2 (A20), Ph3 (A20), Op22 (A20), Op24 (A24), Fi36 (A26), F137 (A24)*		

FIGURE 9.5 General outline of variable definition.

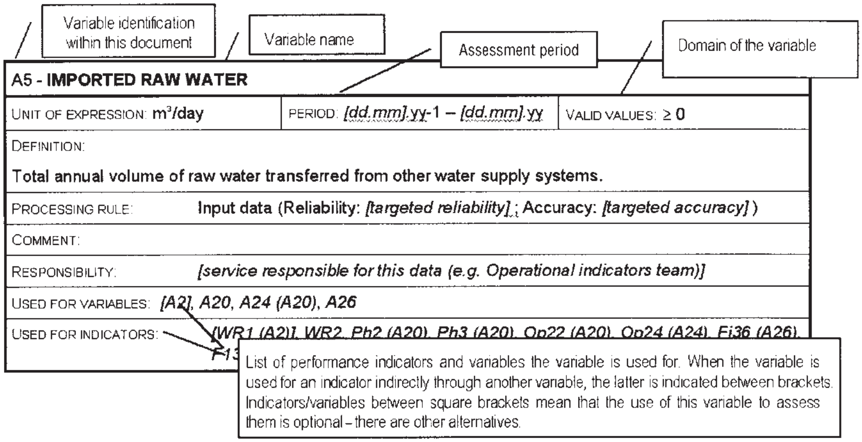


FIGURE 9.6 Variable definition, classification, and description.

Treated water, imported or exported The annual volumes of bulk transfers of treated water across operational boundaries. The transfers can occur anywhere downstream treatment. [The annual volume of water (if any) that is abstracted and delivered to consumers without any treatment shall also be accounted for as *treated water* in the scope of the water balance.]

Transmission input The annual volume of treated water input to a transmission system.

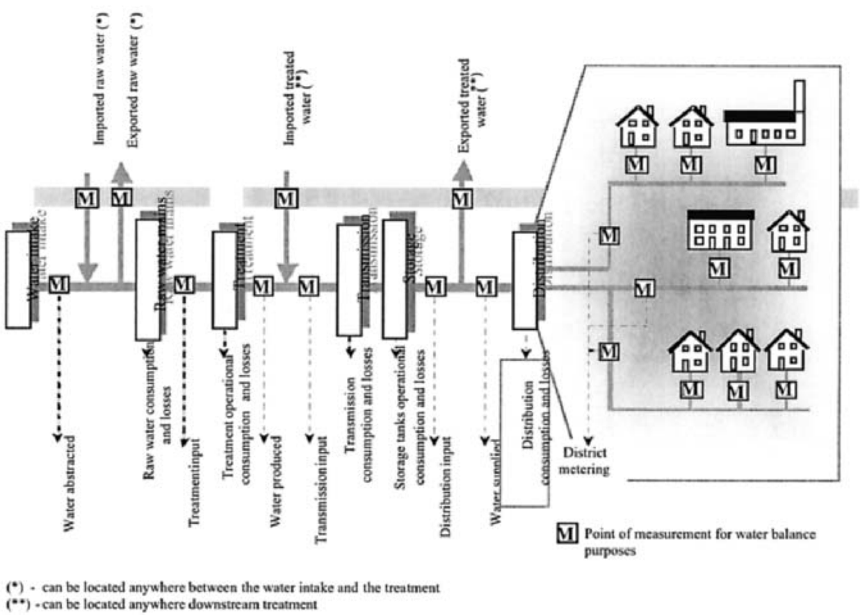


FIGURE 9.7 Definition of water supply system inputs and outputs.

A	B	C	D	E
System input volume [m^3/year]	Authorized consumption [m^3/year]	Billed authorized consumption [m^3/year]	Billed metered consumption (including water exported) [m^3/year]	Revenue water [m^3/year]
			Billed unmetered consumption [m^3/year]	
		Unbilled authorized consumption [m^3/year]	Unbilled metered consumption [m^3/year]	Nonrevenue water [m^3/year]
			Unbilled unmetered consumption [m^3/year]	
	Water losses [m^3/year]	Apparent losses [m^3/year]	Unauthorized consumption [m^3/year]	
			Metering inaccuracies [m^3/year]	
		Real losses [m^3/year]	Real losses on raw water mains and at the treatment works (if applicable) [m^3/year]	
			Leakage on transmission and/or distribution mains [m^3/year]	
			Leakage and overflows at transmission and/or distribution storage tanks [m^3/year]	
			Leakage on service connections up to the measurement point [m^3/year]	

FIGURE 9.8 Components of water balance.

Distribution input The annual volume of treated water input to a distribution system.

Supplied water The *distribution input* minus *treated water exported.* (When it is not possible to separate transmission from distribution, supplied water is the *transmission input* minus *treated water exported*).

System input volume The annual volume input to that part of the water supply system to which the water balance calculation relates.

Authorized consumption The annual volume of metered and/or nonmetered water taken by registered customers, the water supplier, and others who are implicitly or explicitly authorized to do so by the water supplier, for residential, commercial, and industrial purposes. It includes *water exported. Notes:* (1) Authorized consumption may include items such as fire fighting and training, flushing of mains and sewers, street cleaning, watering of municipal gardens, public fountains, frost protection, and building water. These may be billed or unbilled, metered or unmetered, according to local practice. (2) Authorized consumption includes leakage and waste by registered customers that are unmetered.

Water losses The difference between system input volume and authorized consumption. Water losses can be considered as a total volume for the whole system, or for partial systems such as raw water mains, transmission, or distribution. In each case the components of the calculation would be adjusted accordingly. Water losses consist of real losses and apparent losses.

Real losses Physical water losses from the pressurized system, up to the point of measurement of customer use. The annual volume lost through all types of leaks, bursts, and overflows depends on frequencies, flow rates, and average duration of individual leaks. *Note:* Although physical losses after the point of customer flow measurement are excluded from the assessment of *real losses,* they are often significant (particularly where customers are unmetered) and worthy of attention for demand management purposes.

Apparent losses Accounts for all types of inaccuracies associated with production metering and customer metering, plus unauthorized consumption (theft or illegal use). *Note:* Underregistration of production meters, and overregistration of customer meters, leads to underestimation of *real losses.* Overregistration of production meters, and underregistration of customer meters, leads to overestimation of *real losses.*

Nonrevenue water The difference between the annual volumes of *system input* and *billed authorized consumption. Nonrevenue water* includes not only the *real losses* and *apparent losses,* but also the *unbilled authorized consumption.*

Water Balance Components[3] Figure 9.8 shows the recommended standard format and terminology for water balance calculations for one or more sections of a water supply system (e.g., raw water mains, transmission, distribution). If original water balance data is available in any alternative format or terminology, they must be reassembled into the components in Fig. 9.8 in volume per year, before attempting to calculate any PIs. *Note*: Consumption of water by registered customers who pay indirectly through local or national taxation is deemed to be billed authorized consumption for the purposes of the water balance.

The steps for calculating nonrevenue water and water losses are as follows:

Step 1. Define *system input volume* and enter in column A.

Step 2. Define *billed metered consumption* and *billed unmetered consumption* in column D; enter total in *billed authorized consumption* (column C) and *revenue water* (column E).

Step 3. Calculate the volume of *nonrevenue water* (column E) as *system input volume* (column A) minus *revenue water* (column E).

Step 4. Define *unbilled metered consumption* and *unbilled unmetered consumption* in column D; transfer total to *unbilled authorized consumption* (column C).

Step 5. Add volumes of *billed authorized consumption* and *unbilled authorized consumption* in column C; enter sum as *authorized consumption* (column B).

Step 6. Calculate *water losses* (column B) as the difference *between system input volume* (column A) and *authorized consumption* (column B).

Step 7. Assess components of *unauthorized consumption* and *metering inaccuracies* (column D) by best means available; add these and enter sum in *apparent losses* (column C).

Step 8. Calculate *real losses* (column C) as *water losses* (column B) minus *apparent losses* (column C).

Step 9. Assess components of *real losses* (column D) by best means available (night flow analysis, burst frequency/flow rate/duration calculations, modeling, etc.); add these and cross-check with volume of *real losses* in column C.

Particular difficulty is experienced in completing the water balance with reasonable accuracy where a significant proportion of customers are not metered. Authorized consumption in such cases should be derived from sample metering of sufficient numbers of statistically representative individual connections of various categories and subcategories and/or by measurement of total flows into discrete areas of uniform customer profile, also of various categories and subcategories. In the latter method, subtraction of leakage demand from total input is necessary, leakage being determined by analysis of the subcomponents of night demand, adjusting for diurnal pressure variation as appropriate. The confidence grade allocated to the authorized consumption (see Sec. 9.5.3) should reflect the rigor of the investigations.

In this document it is the intention that performance related to management of water losses should be measured from three different viewpoints: financial, technical, and water resources. If the water balance calculations proceed no further than step 3 in Fig. 9.8, as is the case in the simplest traditional water balances, the only performance indicator that can be calculated is the financial one of nonrevenue water by volume (%). The task force on water losses emphasized the importance of completing the calculation up to step 8 (and preferably step 9) in Fig. 9.8—in particular, the importance of attempting, by the best means available, the separation of water losses into apparent and real losses. This allows for the calculation of a required range of water resource, operational, and financial PIs as shown in Fig. 9.9 and Table 9.12.

Definition of the Water Losses Indicators. Table 9.13 reproduces the definition of all the water loss indicators listed above. Each variable identified in the processing rule is defined in detail in forms such as illustrated in Sec. 9.5.4. Table 9.14 summarizes their meanings.

9.5.6 The SIGMA Lite Software

The SIGMA Lite software (see Fig. 9.10) was prepared by the Instituto Tecnológico del Agua (ITA), Spain, as a simplified version of its professional PI evaluation system, SIGMA Pro. Both versions of the software are based on the

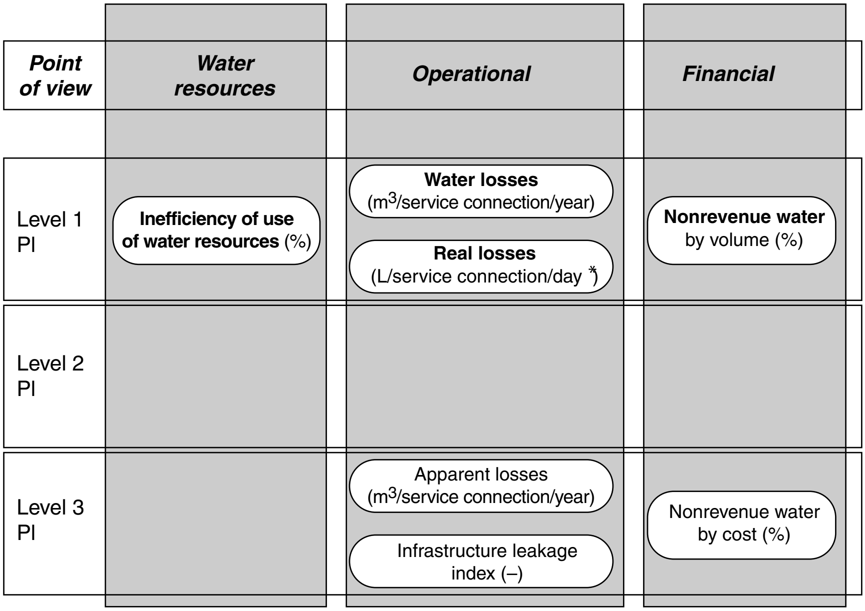


FIGURE 9.9 Points of view and relative importance of the recommended *water losses* and *nonrevenue water* indicators. Asterisk = when system is pressurized.

IWA PI framework. SIGMA Lite is available free of charge through the Internet and can be downloaded at the SIGMA web site: *http://www.sigmalite.com.*

Although dating from 1997 as an independent initiative, the SIGMA project joined the IWA task force on PI in mid-1999, as a follow-up of an IWA invitation to ITA to prepare a freeware version of the program that would be the appropriate complement to the paper document. The Lite version features:

- A stand-alone PI evaluation system, independent of the database systems.
- The complete set of performance indicators from the IWA proposal.
- A user-friendly graphic interface, with intuitive management of the indicators and variables involved.
- Easy operation—calculating PI values is intuitive and fast; chances of calculation errors are greatly reduced.
- Compatibility with MS-Excel for further interpretation and processing of the results.

(See Fig. 9.11.) The Lite version aims to be a good *tryout* tool to discover the potential of the IWA PI system. Additionally it is a powerful educational tool.

The professional version of SIGMA, developed as an independent project of ITA, has not only the basic features of SIGMA Lite, but also a good number of

TABLE 9.12 Performance Indicators for Water Losses and Nonrevenue Water

Indicator	Recommended units	Comments
Inefficiency of use of water resources	Real losses as % of system input volume	Unsuitable for assessing efficiency of management of distribution systems
Water losses	m^3/service connection/year*	Same units as authorized consumption
Apparent losses	m^3/service connection/year*	Same units as authorized consumption
Real losses	Liters/service connection/day* when system is pressurized	Allows for intermittent supply situations
Infrastructure leakage index	Ratio of real losses to technical achievable low-level annual real losses	Technical achievable low-level annual real losses are equal to the best estimate of so-called unavoidable average real losses (UARL). They include system-specific allowance for connection density, customer meter location on service, and current average pressure
Nonrevenue water by volume	Volume of nonrevenue water as % of system input volume	Can be calculated from simple water balance
Nonrevenue water by cost	Value of nonrevenue water as % of annual cost of running system	Allows separate values per m^3 for components of nonrevenue water

*Where service connection density is less than 20 per km of mains, use "/km mains/" instead of "/service connection/" for these performance indicators.

complementary ones. For instance, it allows for the customization of the IWA PI system or even the construction of a completely new set of PIs. It also allows the user to store indicators allowing for time-based comparisons, and report and chart customization features will be available.

9.5.7 International Field Test of the IWA PI System (2000–2003)

The IWA PI system resulted from a careful analysis of the potential uses of each indicator and hopefully corresponds to the most current management needs. However, there is no doubt room for improvement. With the publication of the *Performance Indicators for Water Supply Services* in the Manual of Best Practice Series, the IWA launched a new initiative aiming to promote the use of this

TABLE 9.13 Definition of the Water Losses Indicators

Indicator	Level of importance*	Unit	Concept	Processing rule
WR1: Inefficiency of use of water resources	L1	%	Real losses/water abstracted and imported water × 100 *This indicator must not be used as a measure of efficiency of management of the transmission and/or the distribution system.*	$WR1 = (A24)/(A4 + A5 + A8) \times 100$
Op22: Water losses	L1	m^3/connection/year	Water losses/number of *service connections* *If service connections density <20/km of mains (e.g., transmission networks), then this indicator should be expressed in m^3/km of water mains/year.*	Op22 = $A20$/C32
Op23: Apparent losses	L3	m^3/connection/year	Apparent losses/number of *service connections* *If service connections density <20/km of mains (e.g., transmission networks), then this indicator should be expressed in m^3/km of water mains/year.*	Op23 = $A23$/C32
Op24: Real losses	L1	Liters/connection/day when system is pressurized	Real losses × 1000/(number of service connections × 365 × T/100) (T = % of year system is pressurized†) *If service connections density <20/km of mains (e.g., transmission networks), then this indicator should be expressed in liters/km of water mains/day*	Op24 = $A24 \times 1000/(C32 \times 365 \times D29/365)$
Op25: Infrastructure leakage index	L3	—	Real losses (Op)/ technical achievable low-level annual real losses (when system is pressurized) *The technical achievable low-level annual real losses are equal to the "best estimate" of so-called unavoidable annual real losses (UARL), which can be calculated with the equation derived by the Water Losses Task Force (see AQUA Dec. 1999 and IWA Blue Pages "Losses from water supply systems"):*	$Op25 = Op24/(18 \times C6/C32 + 0.7 + 0.025 \times C33) \times D31$

			UARL (liters/service connection/day) = $(18 \times Lm/Nc + 0.7 + 0.025\ Lp) \times P$ *This equation, based on empirical results of international investigations, recognizes separate influences on real losses from:* *Length of mains Lm in km (C6)* *Number of service connections Nc (C32)* *Average length of service connections Lp in m (C* *Average operating pressure in m (D31)* *Well-managed systems are expected to have low values of this infrastructure leakage index—close to 1.0—while systems with infrastructure management deficiencies will present higher values.*	
Fi36: Nonrevenue water by volume	L1	%	Nonrevenue water/system input volume × 100	Fi36 = A26/(A7 + A8) × 100
Fi37: Nonrevenue water by cost	L3	%	Valuation of nonrevenue water components/annual running costs × 100 This is the sum *of separate valuations for unbilled authorized cons umption, apparent losses and real losses. With respect to real losses, the attributed unit cost (G50) will be the highest of the variable component of bulk supply charge or long-run marginal cost (LRMC) for own sources. It is usually worthwhile to calculate and review the three components of Fi37 separately.*	Fi37 = ((A18 + A23) × G49 + A24 × G50)/G2 × 100

*L1 = level 1 indicator; L2 = level 2 indicator; L3 = level 3 indicator.

†The term "% of year system is pressurized" is only relevant for intermittent supply networks. For permanent supply networks it assumes the value 1.

TABLE 9.14 Synthesis of Variables Used

A4	***Water abstracted*** (m^3/year): The annual volume of water obtained for input to water treatment plants (or directly to the transmission and distribution systems) that was abstracted from raw water sources.
A5	***Imported raw water*** (m^3/year): Total annual volume of raw water transferred from other water supply systems.
A7	***Water produced*** (m^3/year): Total annual volume of water treated for input to water transmission lines or directly to the distribution system.
A8	***Imported treated water*** (m^3/year): Total annual volume of treated water imported from other water undertaking or system.
A18	***Unbilled authorized consumption*** (m^3/year): Total annual amount of unbilled water consumed.
A20	***Water losses*** (m^3/year): The difference between *system input volume* and *authorized consumption.* Water losses can be considered as a total volume for the whole system, or for partial systems such as raw water mains, transmission, or distribution. In each case the components of the calculation would be adjusted accordingly.
A23	***Apparent losses*** (m^3/year): Total annual amount of water unaccounted for due to metering inaccuracies and unauthorized consumption.
A24	***Real losses (m^3/year): Total annual amount of physical water losses from the pressurized system, up to the point of customer metering.***
A26	***Nonrevenue water*** (m^3/year): Difference between the annual volumes of system input and billed authorized consumption (including exported water).
C6	***Mains length*** (km): Total transmission and distribution mains length (service connections not included).
C32	***Service connections*** (number): Total number of service connections.
C33	***Average service connection length*** (m): Frequently water undertakings do not have detailed accurate information to assess the total service connections length. In these cases, a qualitative assessment will be adopted.
D29	***Time system is pressurized*** (hours): Amount of time of the year the system is pressurized.
D31	***Average operating pressure*** (kPa): Average operating pressure at the delivery points when system is pressurized.
G2	***Annual running costs*** (US$/year): Total annual operations and maintenance costs + internal work force costs − capitalized costs of self-constructed assets.
G49	***Average water charges for direct consumption*** (US$/$m^3$): Water sales revenue for direct consumption/billed water.
G50	***Attributed unit cost for real losses*** (US$/$m^3$): Highest of the variable components of bulk supply charge or long-run marginal cost (LRMC).

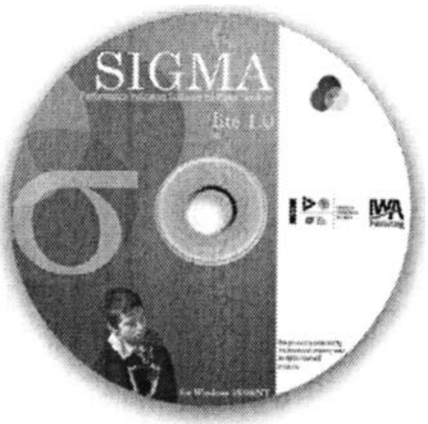


FIGURE 9.10 SIGMA Lite CD-ROM.

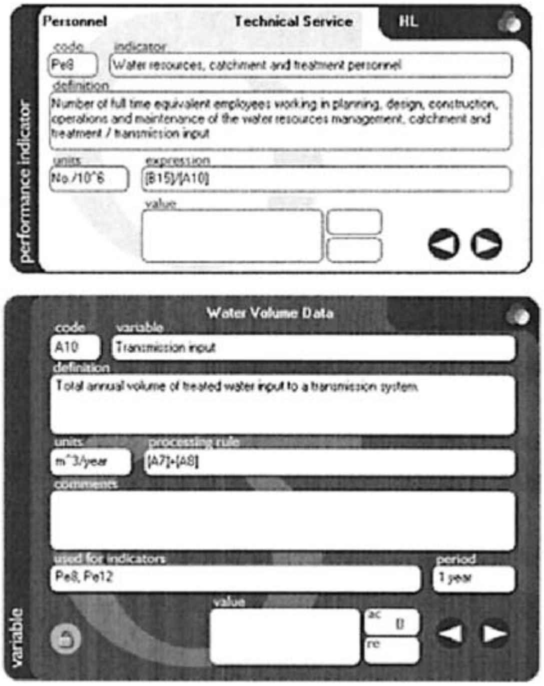


FIGURE 9.11 PI and variable data windows within SIGMA.

manual and field-test it in a broad number of real situations. That work has been ongoing since July 2000 and is planned to continue until July 2003 when the results will be used as the basis for a revised edition of the manual and of the software.

The IWA provided some seed money to make this initiative possible, and a volunteer core team assures the coordination.[4] The participating undertakings are committed to use the system for about 2 years during the course of the project, to reply to three inquiries, one per year, and to share their hands-on experience of using the IWA PI system. Conversely, this is a unique opportunity for them to start using the system counting on technical support from the IWA, while contributing and influencing an IWA-led development process, "by the industry, for the industry," not driven by external financial or commercial interests.

Although the initial target was to have 10 to 15 case studies from different regions and with different characteristics, the number of volunteers exceeded the more optimistic perspectives, and the project currently incorporates 70 undertakings from 19 countries (from Europe, Africa, South America, Middle East, and the Asia-Pacific region). Figure 9.12 presents the geographic locations of the participants. These include organizations of different type (utilities, regulators, holding companies), size, activity (water only or water and other activities; bulk supply or distribution), asset ownership, and type of operations (public or private).

The current English language version of both the text and the software is being translated into Portuguese, Spanish, and German. Translation into French is under consideration, and a simplified Czech version is available.

Given the complexity and variety of situations, any general PI system needs to be sufficiently flexible to accommodate each case, allowing for customization. However, on the other hand, definitions and processing rules must be very precise, to prevent comparisons between "pears and apples." These are underlying principles adopted for the IWA PI system.

The IWA PI system for water supply services is tailored as a tool to support strategic management. It aims to be an evolving tool. Given the rapid evolution in this area of knowledge, such a system will always have room for improvement. If the first edition of the IWA PI manual is already the output of a rather mature work resulting from the experience and know-how from many IWA members, the contributions arising from 2 years of use by such a broad group of undertakings will certainly provide significant improvements to the second edition.

The IWA PI system for wastewater services, currently in an already advanced stage of development, is supported on the same structure of indicators (Matos et al., 2001). The methodology adopted combines the work of a core team, assisted by an ad hoc international advisory group that provides comments and suggestions, with contributions from any interested person. The working documents are available at any time to any interested person who requests it, the process of evolution being completely transparent. The publication of the corresponding manual is scheduled for the beginning of 2003.

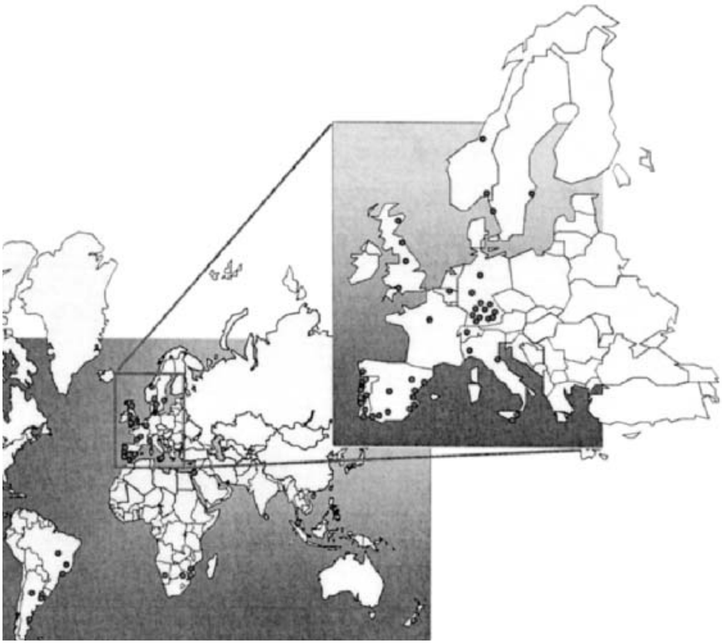


FIGURE 9.12 Participating countries. Number of entities per country: Argentina (4), Belgium (1), Brazil (3), France (1), Germany (13), Israel (1), Italy (4), Malaysia (1), Mozambique (2), Namibia (1), New Zealand (1), Norway (2), Portugal (15), South Africa (1), Spain (12), Sweden (2), Switzerland (1), UK/England (3), UK/Scotland (1).

9.6 IMPLEMENTATION OF A PI SYSTEM

9.6.1 Phases of Implementation

The implementation of a PI system within an organization must be carefully planned. The recommended implementation phases are synthesized in the schematics of Fig. 9.13. The following subsections refer to each one of these phases.

9.6.2 Definition of the Strategic Performance Assessment Policy

Objective Definition. The implementation of any PI system has to be objective-oriented. The number and definition of the PI system to be adopted shall reflect directly the specific objectives of its use. The indicators adequate for a global and synthetic performance assessment for the top managers of a big organization are different from the indicators adequate for a more in-depth analysis, or for a thematic use, such as water losses or rehabilitation.

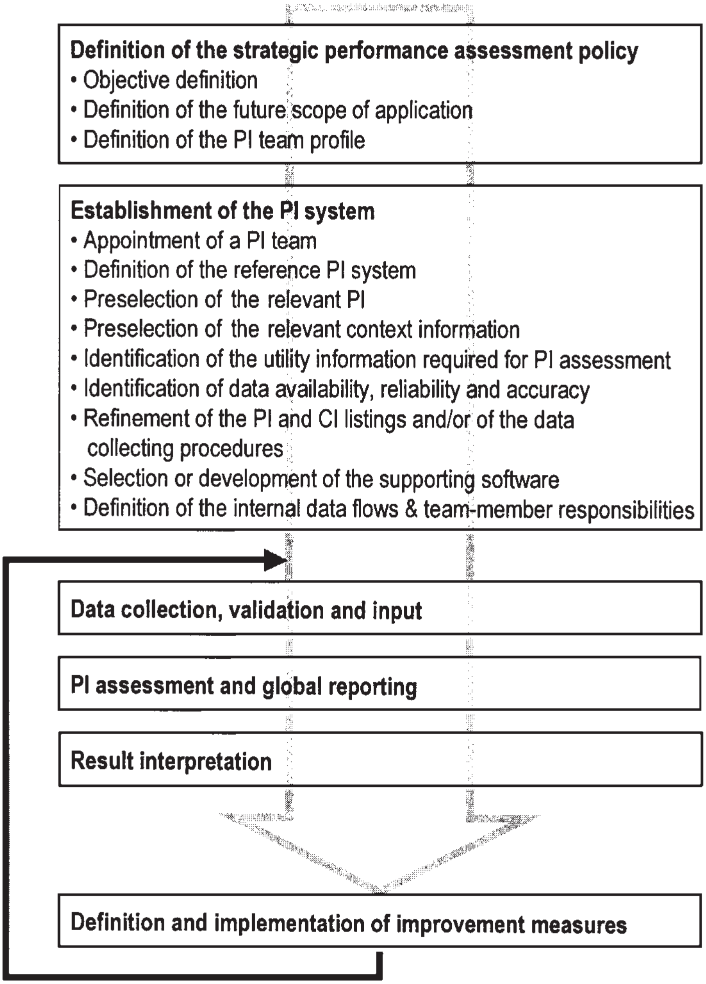


FIGURE 9.13 Phases of implementation of a PI system.

Definition of the Future Scope of Application. The PI definitions to be adopted will depend on the domain within which the indicators will be compared and interpreted; the greater the group of undertakings involved is, the smaller is the degree of flexibility for creating ad hoc indicators; however, it is advisable to use definitions nationally or internationally accepted whenever possible, even though the immediate use is confined to the undertaking or to a small group of undertakings.

Definition of the PI Team Profile. According to the specific objectives and taking into account the organization size resources, the decision makers should define the appropriate profile of the future PI team. A successful implementation of a PI system within an organization requires that the project manager is a senior staff member, if necessary assisted by less experienced collaborators; the involvement and active participation of the decision makers who are going to use the PI values is also indispensable; it is strongly recommended that representatives of the departments who need to provide data are members of the PI team, so that they can participate in all phases of the process; in fact, experience has shown that one of the important benefits of the implementation of PI systems is the sharing of experiences and viewpoints between staff members with different responsibilities and backgrounds; when the PIs are selected, it is recommended that the name of a person or sector is assigned to each data item as being responsible for its provision, validation, and input. Experience shows that the assistance of an experienced external consultant during the implementation period can be beneficial, particularly if it is a coordinated initiative of a group of undertakings (e.g., six Scandinavian cities group project; Portuguese PI project; Waters, German PI projects; WUP PI project).

9.6.3 Establishment of the PI System

Appointment of a PI Team. Based on the profile defined, a team shall be appointed and briefed about the objectives and intended scope of application. The future collaborators within the undertaking shall also be informed and motivated for the process from a very early stage.

Definition of the Reference PI System. Before selecting the specific PI to be adopted, the organization managers need to decide whether they prefer to use an ad hoc system, developed expressly for the objective defined, or an existing external PI system (e.g., the IWA PI system). The former option has the advantage of providing more freedom in terms of the definitions to be adopted; however, the development of a dedicated system has the disadvantages of being a time-consuming and expensive task, preventing future comparisons with the outside world and reflecting the internal experience, leaving aside the contributions of many experienced persons from other organizations. The latter option has the disadvantage of imposing several constraints in terms of data and PI definitions; however, existing systems have the major advantage of being matured and tested and allowing future comparisons with reference levels or with other organizations in a much more reliable way. Added advantages are the cost and time of implementation, since these systems already exist, and software is available for some of them (e.g., the IWA PI system, the World Bank benchmarking toolkit, and the WUP PI system).

Preselection of the Relevant PI. The selection of the final listing of PI to be adopted is sometimes an iterative process. The experience of the decision makers will allow for selecting a preliminary listing that seems to fit the defined objectives. In some cases the choice is trivial, but in other cases it is more complex than it may appear. It is important to analyze a priori the information that a given indicator can provide and check if it actually fits the objectives. Sometimes even different units of expression lead to rather different information. Another factor that cannot be undervalued is the confidence of the data required to assess the indicators. It may be preferable to use a less ambitious indicator that is based on reliable data than another theoretically better indicator which requires data that is not available or reliable within the organization: the cost/benefit balance has to always be taken into account. It is recommended that undertakings use parsimony and start with a small number of indicators. Later on they can embrace new indicators step by step.

Identification of Data Availability, Reliability. and Accuracy. The preliminary listing of PI allows the team to identify all the required input data and work with the data providers in order to assess the data's availability, reliability, and accuracy. Whenever a variable is not available or is inaccurate, decision makers and data providers have to agree on what to do: improve the situation or accept that the indicators using this data cannot be assessed or are inaccurate.

Refinement of the PI Listing and Data Collecting Procedures. The assessment of data availability, reliability, and accuracy allows for a consolidation of the PI listing, as well as for some improvements of the data measurement and collection procedures, when appropriate. Another task at this stage is the definition of the frequency of assessment of the indicators. In general, the indicators for overall assessment of the undertaking are assessed once per year, but there are indicators and objectives for which different frequencies are advisable.

Selection or Development of the Software. Although the processing rules of each indicator tend to be as simple as a short number of algebraic operations, the total number of variables manipulated justifies the use of an adequate software application. It is advisable that the software incorporates analysis of consistency procedures, in order to minimize the errors (e.g., range of valid values and cross-coherence between some data values). The software application can be developed for the specific case or the undertaking or be a software of general use. The spreadsheet application made available by the World Bank (WB benchmarking toolkit) and the software SIGMA Lite (IWA PI system) are available free of charge and can be used if the PIs selected are part of the WB indicators or the IWA indicators, respectively.

Definition of the Internal Data Flows and Team Member Responsibilities. A responsible person has to be assigned to each data item required by the PI system. Each variable measurement (or estimate) must be compatible with the preset PI frequency, be assessed for reliability and accuracy, and be input into the PI system.

If the same person or team does not hold all of these tasks, it is recommended that the procedure and the responsible person for each step of the process be specified.

The synergy with the existing undertaking information systems should be maximized. Data providers need to feel that the PI system is part of the organization's normal activity, and that there are major advantages in using the same data (i.e., defined and expressed in the same way) for the various existing applications requiring similar information. However, the effective integration of the existing data banks is part of a more general strategic objective of most utilities, and not a PI-specific problem.

9.6.4 Data Collection, Validation, and Input

Taking into account the periods of reference for the PI assessment, the PI team requests the data values to the respective responsible. Depending on the internal procedure agreed upon, data input could be performed directly by each responsible person or collectively by the PI team.

9.6.5 PI Assessment and Global Reporting

PI assessment is just a fingertip task, provided that the data are correctly input and an adequate software application is available. Reports should be customized taking into account the intended use of the information. It is important that the reports include not only the PIs but also the relevant context information. General reports can be limited to tables containing the results, but can also include graphical representations showing the PI evolution with time, presenting sets of related PIs in the same graphic, or comparing PIs with preset targets or reference thresholds, for instance.

9.6.6 Results Interpretation

The success of implementing a PI system depends on the effective use of the results as a decision-making tool. The objective is to interpret the results (i.e., to decide whether a given result is satisfactory or not). If the results are not satisfactory, a diagnosis needs to be established and improvement measures identified.

The process of results interpretation requires comparison of the results to a reference value or a reference threshold, taking into account the context information, and analyzing at each time a group of related indicators and not single indicators individually. The sources of reference values are of five main types:

- Equivalent results of previous years
- Targets established in the master plans
- Targets established within legal, contractual, or regulatory frameworks

- Results from other operational units of the same undertaking (e.g., valid for holdings managing more than one undertaking or for undertakings with more than one system or operational unit)
- Results from other undertakings within metric benchmarking initiatives
- Reference values publicly available

The internal comparison is in general the most important and effective. However, it starts to be feasible only after some years of systematic use of the PI system. The comparison with other undertakings or with publicly available reference values is therefore a need in the first years. On the other hand, sharing experiences with others has a significant leverage effect that should not be ignored.

As managers may be interested in keeping some of the indicators confidential, comparisons of these can be held within small carefully selected groups or provided to an independent external entity that collects the data of the various undertakings, processes them, and reports the results of the group in a form in which individual identification is not possible.

9.6.7 Definition and Implementation of Improvement Measures

The definition and implementation of improvement measures, although not being a part of the performance assessment itself, is so important that it must be dis-

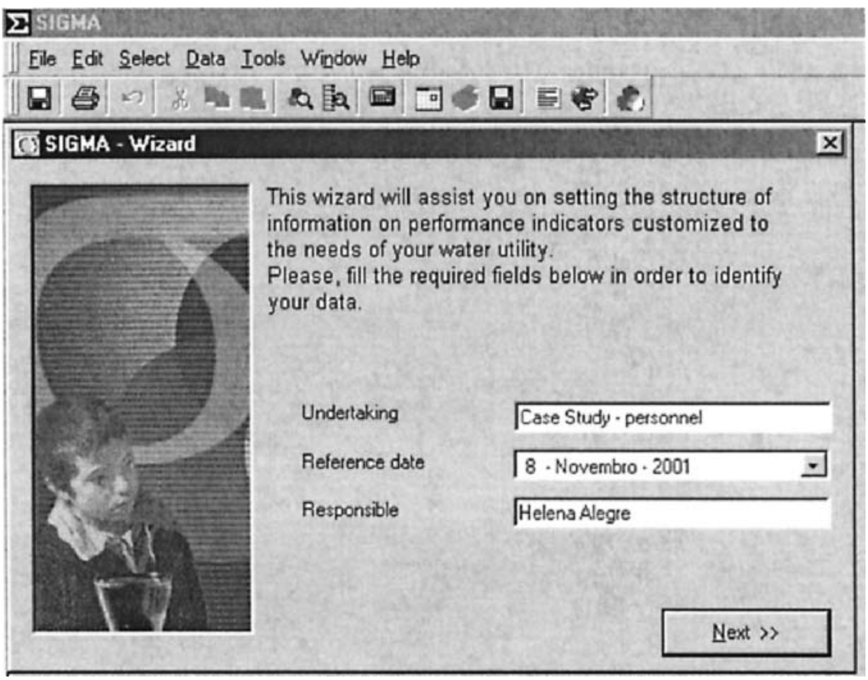


FIGURE 9.14 First SIGMA window: user identification.

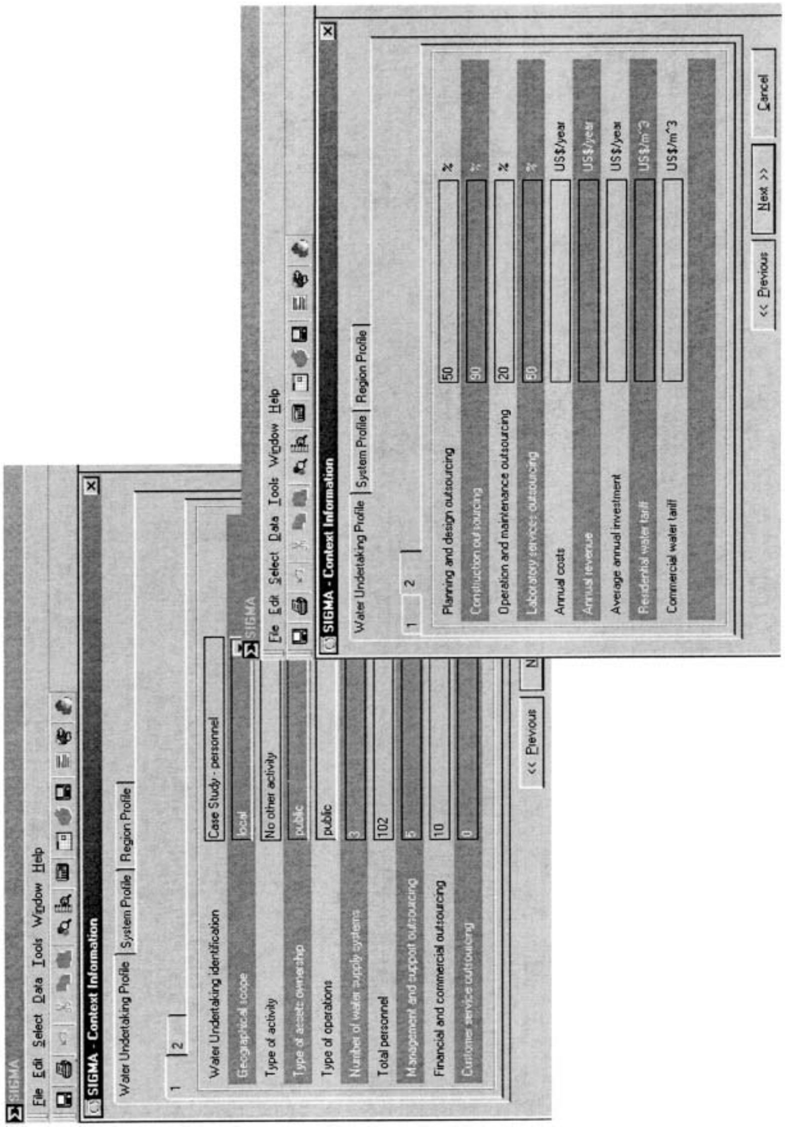


FIGURE 9.15 Context information.

cussed. Based on the results interpretation, weaknesses are identified and a diagnosis should be established. From these, decision makers can establish short-, medium-, or long-term targets for improvement and implement the remedial measures. The systematic use of the PI system will later on allow for monitoring of the results and comparing them with the initial targets. Deviations shall be carefully interpreted in order to introduce corrective measures.

9.7 EXAMPLE OF THE USE OF SIGMA LITE FOR DATA INPUT AND INDICATORS ASSESSMENT

9.7.1 Scope of the Example

This section exemplifies the use of the software SIGMA Lite. It contains a step-by-step procedure:

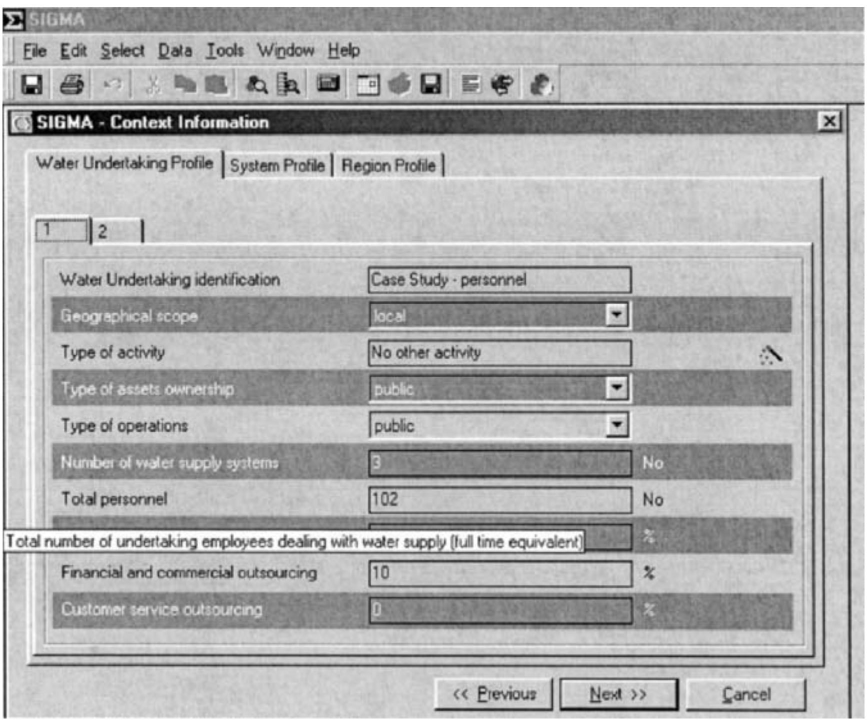


FIGURE 9.16 Example of on-line help.

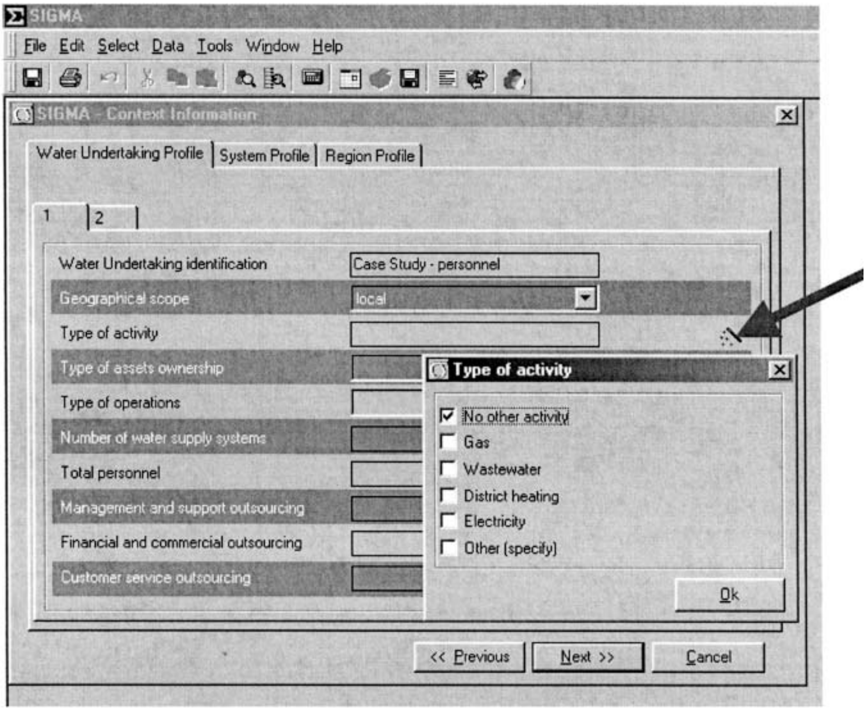


FIGURE 9.17 Example of wizard.

- Input the context information.
- Select the PI to be assessed.
- Input the respective data values.
- Assess the PI.
- Produce PI reports.

A small group of indicators related to personnel is used for this scope:

- Pe1: Employees per connection
- Pe2: Management and support personnel
- Pe3: Financial and commercial personnel
- Pe4: Customer service personnel
- Pe5: Technical services personnel
- Pe6: Planning and construction personnel
- Pe7: Operations and maintenance personnel

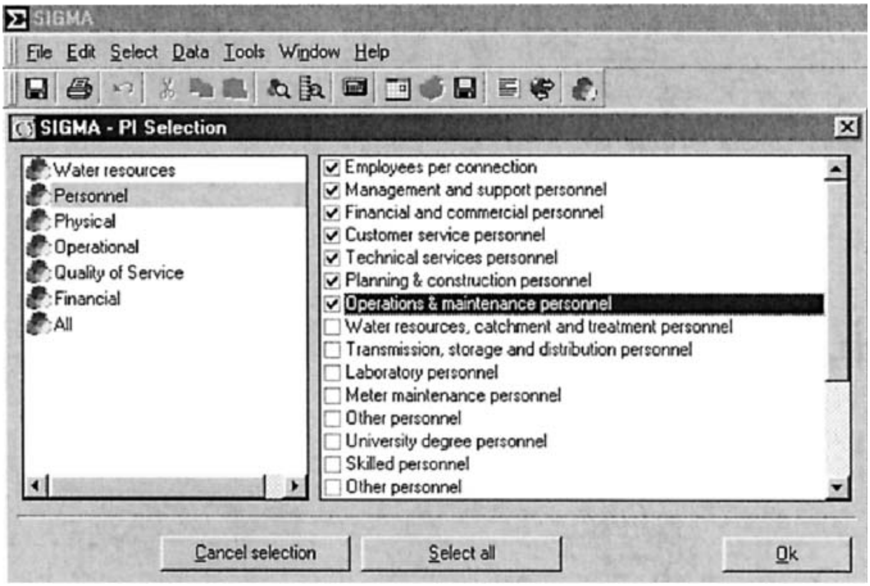


FIGURE 9.18 Selection of the indicators.

9.7.2 Starting to Use SIGMA Lite

The installation of SIGMA is self-explanatory, and the user just needs to reply to the questions asked. A wizard is available the first time the program is run. If the user wishes to reset SIGMA to see the wizard again and input new values, then he or she has to select Tools, Options, Data, or Reset SIGMA from the SIGMA Lite menu. The first step is user identification (Fig. 9.14).

9.7.3 Context Information

The next step is the input of the context information (Fig. 9.15). This information is not used for the assessment of the indicators, but exclusively to help interpret

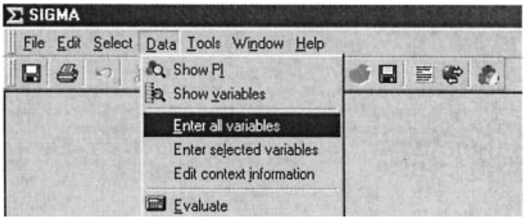


FIGURE 9.19 Selection of data entry.

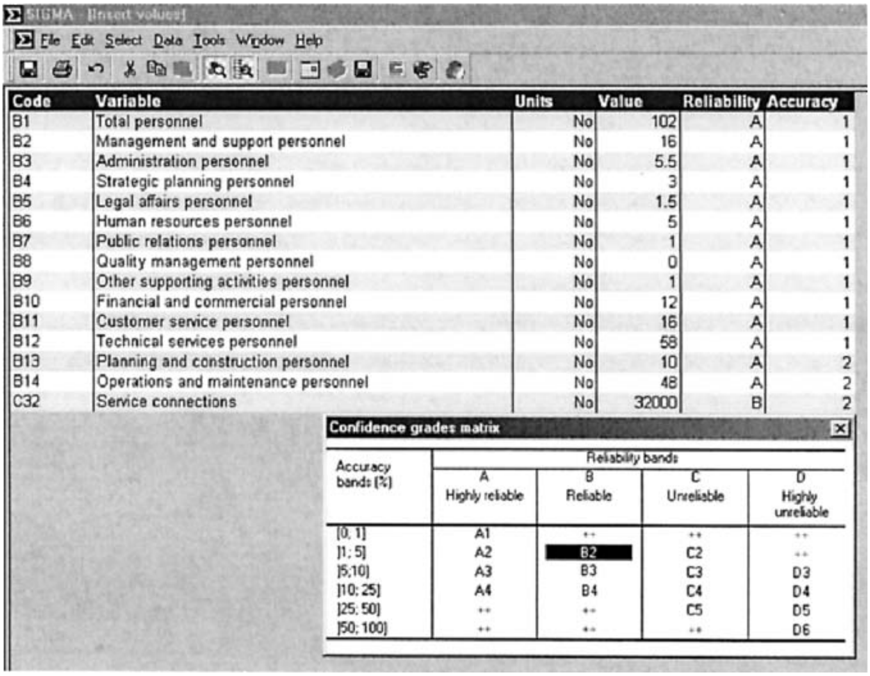


FIGURE 9.20 Data input.

the results. Therefore, the user does not have to fulfill the entire set of tables, but rather select the information considered to be relevant.

In the case of personnel, the most critical context information relates to the type of undertaking and of the service provided, and the percentage of outsourcing for each main function within the organization. In fact, personnel indicators are assessed based on the permanent and temporary personnel of the undertaking, but do not incorporate the work force included as part of contracted services (e.g.,

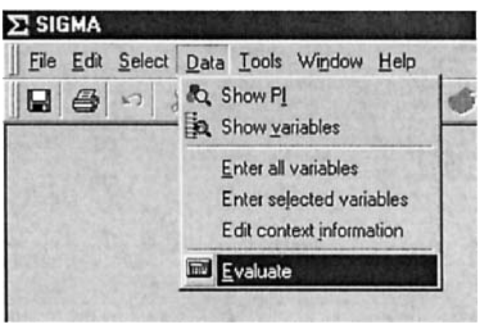


FIGURE 9.21 PI evaluation.

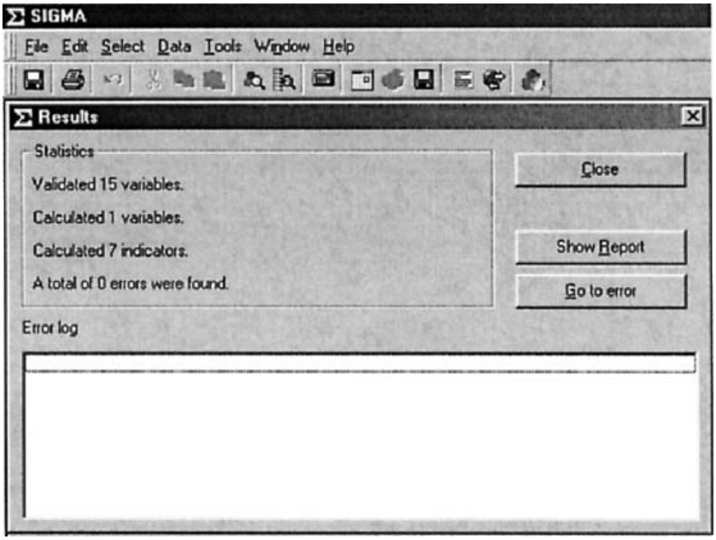
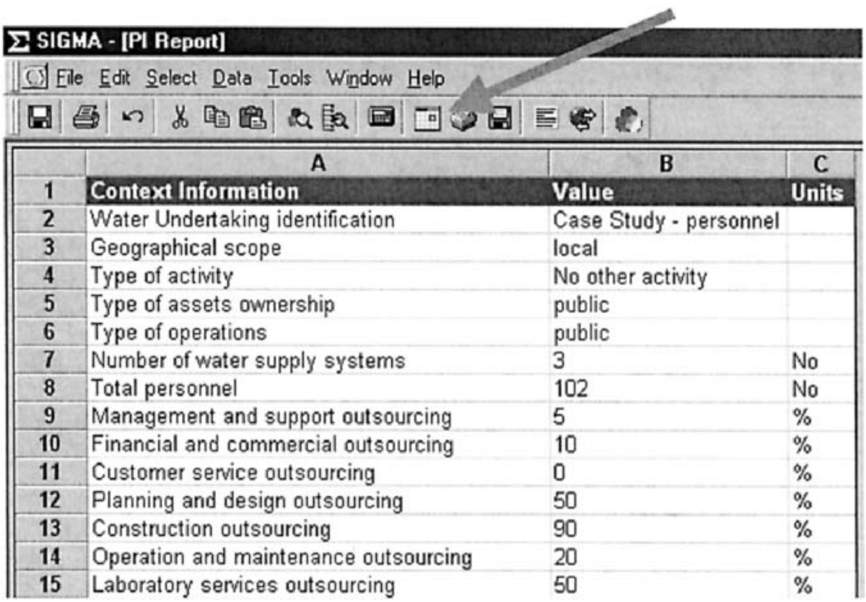


FIGURE 9.22 Statistics and error log file.



	A	B	C
1	Context Information	Value	Units
2	Water Undertaking identification	Case Study - personnel	
3	Geographical scope	local	
4	Type of activity	No other activity	
5	Type of assets ownership	public	
6	Type of operations	public	
7	Number of water supply systems	3	No
8	Total personnel	102	No
9	Management and support outsourcing	5	%
10	Financial and commercial outsourcing	10	%
11	Customer service outsourcing	0	%
12	Planning and design outsourcing	50	%
13	Construction outsourcing	90	%
14	Operation and maintenance outsourcing	20	%
15	Laboratory services outsourcing	50	%

FIGURE 9.23 Report: context information.

SIGMA - [PI Report]

File Edit Select Data Tools Window Help

	A	B	C	D	E
1	Code	Variable	Value	Units	Confidence
2	B1	Total personnel	102	No	A1
3	B2	Management and support personnel	16	No	A1
4	B3	Administration personnel	5,5	No	A1
5	B4	Strategic planning personnel	3	No	A1
6	B5	Legal affairs personnel	1,5	No	A1
7	B6	Human resources personnel	5	No	A1
8	B7	Public relations personnel	1	No	A1
9	B8	Quality management personnel	0	No	A1
10	B9	Other supporting activities personnel	1	No	A1
11	B10	Financial and commercial personnel	12	No	A1
12	B11	Customer service personnel	16	No	A1
13	B12	Technical services personnel	58	No	A1
14	B13	Planning and construction personnel	10	No	A2
15	B14	Operations and maintenance personnel	48	No	A2
16	C32	Service connections	32000	No	B2

FIGURE 9.24 Report: variables.

construction). The analysis of the personnel indicators cannot be held without the background knowledge of how much of the service is outsourced.

On-line help and a system of wizards guide the user, as exemplified in Figs. 9.16 and 9.17. The context information can be edited at any time, at a subsequent stage.

9.7.4 PI Selection and Data Input

The next step is the specification of the indicators that the user selected. The SIGMA window shows which indicators have been selected. If only a subset is to

SIGMA - [PI Report]

File Edit Select Data Tools Window Help

	A	B	C	D	E
1	Code	Indicator	Value	Units	Confidence
2	Pe1	Employees per connection	3,19	No./1000 connections	B2
3	Pe2	Management and support personnel	0,5	No./1000 connections	B2
4	Pe3	Financial and commercial personnel	0,38	No./1000 connections	B2
5	Pe4	Customer service personnel	0,5	No./1000 connections	B2
6	Pe5	Technical services personnel	1,81	No./1000 connections	B2
7	Pe6	Planning & construction personnel	0,31	No./1000 connections	B2
8	Pe7	Operations & maintenance personnel	1,5	No./1000 connections	B2

FIGURE 9.25 Report: performance indicators.

be used, it is necessary to choose Cancel Selection (see Fig. 9.18). Then the user can tick all the relevant indicators, either from all the indicators or group by group.

After the indicators have been selected, the user introduces the values of the corresponding variables, automatically selected by the program, as shown in Figs. 9.19 and 9.20. To fill in the reliability and accuracy values, pick a cell, press the right button of the mouse, and select from the table shown at the bottom of Fig. 9.20.

9.7.5 PI Assessment and Reporting

The next step is the assessment of the indicators. The user can select Data, Evaluate from the menu (Fig. 9.21), or press the Evaluate icon from the tool bar. A window with a log file appears with the statistics (Fig. 9.22). If the program detects any errors, such as missing variables to assess the selected indicators, these are listed, and the user is guided to the location where the error occurred in order to correct it.

To see the report, press the Spreadsheet icon. It is separated into three parts: context information, variables, and performance indicators (Figs. 9.23 to 9.25). The results are presented in a spreadsheet-like format that can be exported to other applications.

9.8 EXAMPLE OF THE USE OF THE IWA PI SYSTEM FOR WATER LOSSES CONTROL

9.8.1 Scope of the Example

This section presents an example of assessment of the PI selected by a given undertaking to support the process of water losses control. It aims to highlight that PIs can be used not only for the overall assessment of the company but also for specific uses.

9.8.2 Objective Definition and Form of Use of the Indicators

In this example the objective is to identify whether or not a given network is problematic in terms of water losses. Managers are interested in comparing their situation with other networks of the same region, in order to be able to prioritize interventions. Not only will the operational point of view be taken into account, but also the water resources and financial ones. Improvement of customer metering is part of the undertaking policy, in order to minimize the unmetered consumption. The indicators will be used exclusively internally to support managers' decision making. However, comparison with external reference values will be held whenever feasible.

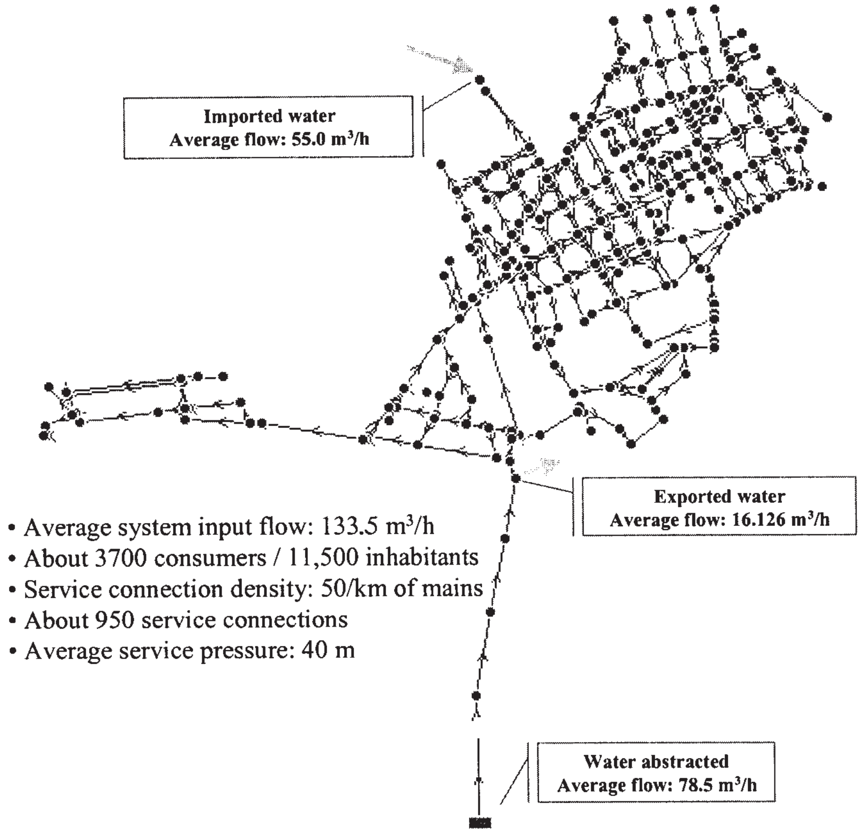


FIGURE 9.26 Case study network.

9.8.3 Appointment of the PI Team and Definition of the Reference PI System

The team has been established in order to create an effective group and to involve decision makers and data providers. A part-time project manager with 10 years of professional experience and a background on operations of water supply networks was appointed. A younger full-time assistant was contracted. One top manager assured his participation in periodic meetings. The contribution of field staff, customer service, and financial and commercial personnel was assured by their participation in periodic meetings. In a later stage of the process, specific responsibilities will be committed to these staff members.

A	B	C	D	E
System input volume *1,169.460 m^3*	Authorized consumption *855,754 m^3*	Billed authorized consumption *771,804 m^3*	Billed metered consumption (including water exported) *695,334 m^3*	Revenue water *771,804 m^3*
			Billed unmetered consumption *76,470 m^3*	
		Unbilled authorized consumption *530 m^3*	Unbilled metered consumption *12,950 m^3*	Nonrevenue water *397,656 m^3*
			Unbilled unmetered consumption *71,000 m^3*	
	Water losses *313,706 m^3*	Apparent losses *115,416 m^3*	Unauthorized consumption *15,000 m^3*	
			Metering inaccuracies *100,416 m^3*	
		Real losses *198,290 m^3*	Real losses on raw water mains and at the treatment works (if applicable) [m^3/year]	
			Leakage on transmission and/or distribution mains [m^3/year]	
			Leakage and overflows at transmission and/or distribution storage tanks [m^3/year]	
			Leakage on service connections up to the measurement point [m^3/year]	

FIGURE 9.27 Components of water balance for the selected case study.

9.8.4 Selection of the Indicators

As a starting point, the undertaking managers analyzed the water losses indicators recommended by the IWA PI system, with emphasis to the level 1 indicators (Sec. 9.5.5), because the managers had agreed to start with a small number of indicators. None of the IWA PI level 2 or level 3 indicators were considered necessary to select at this stage.

The managers recognized that the indicator "inefficiency of use of water resources" was appropriate to represent the water resources perspective. They also agreed that the indicator "nonrevenue water by volume" should be selected to represent the financial viewpoint. However, they were surprised that the network efficiency was not included as an operational indicator. In fact, the network efficiency, expressing the water losses as percentage of the input volume, is the operational indicator of water losses most commonly used worldwide. Its adoption is attractive because it is apparently simple to assess and interpret. However, it has the major advantage of being strongly dependent on the consumption, a scaling factor that leads to misinterpretations, as discussed further on. Conversely, this indicator has the shortcoming of not taking into account any of the relevant factors that influence real losses, such as the total length of the mains and of the service connections, the service pressure, the density of service connections, and the location of domestic meters. These are the reasons why the IWA supports the recommendation of the National Committees of Germany, South Africa, and the United Kingdom, as well as of the British regulator OFWAT, which states that the network efficiency is not an adequate PI for technical use.

To understand how misleading the network efficiency can be, two simple comparisons are suggested:

1. Let us take one network and analyze the water losses in 2 subsequent years. Because of the occurrence of a very wet summer during the first year monitored, the total consumption was significantly above average. The volume of water losses was approximately the same in both cases. However, the observation of the network efficiency, that consequently increased, might lead to the false conclusion that the situation had improved in terms of losses.

2. Let us take two networks exactly equal in terms of material, age, failure rates, diameters, average service pressure, and density of service connections—therefore with the same amount of water losses. The former supplies single-floor buildings, while the later supplies an area of multistore apartment blocks, with a higher consumption per connection and per pipe length. From the operational point of view it is important to find an indicator that provides equivalent values for both situations. The network efficiency does not, and therefore is inappropriate.

The indicator "total water losses" and particularly the indicator "real losses per service connection," as defined in Tables 9.13 and 9.14, are more robust operational indicators. Being aware of these reasons, managers decided to select these two indicators.

Revisiting their objectives, managers verified that only one had been left uncovered: the measure of the performance regarding the minimization of the unmetered consumption. Several options were considered:

- Number of unmetered authorized consumers/total number of authorized consumers
- Unmetered consumption/total input volume
- Unmetered billed consumption/total metered billed consumption
- Unmetered authorized consumption/total authorized consumption

The first indicator was abandoned because metering prioritization should take into account the consumption, and not only the number of consumers. The second indicator was eliminated as well because the other two seemed comparatively better.

The decision between these two latter indicators requires a further clarification of the objectives. If the objective is to minimize the percentage of bills based on estimates, the focus should be put on the billed consumption. If the objective is more general and aims to improve the confidence of the water balance figures, in order to have a proper monitoring of the water losses, the focus should be on the authorized consumption. Managers chose the fourth indicator: "unmetered authorized consumption/total authorized consumption."

9.8.5 Assessment of the Water Balance

Step 0. Define the system (or part of the system) to audit and define the reference dates (1 year) (Fig. 9.26)

Step 1. Define *system input volume* and enter in column A (see Fig. 9.27).

- ***a.*** Water abstracted: $78.5 \times 24 \times 365 = 687{,}660\ m^3$
- ***b.*** Imported treated water: $55.0 \times 24 \times 365 = 481{,}800\ m^3$
- ***c.*** Imported treated water: $0\ m^3$
- ***d.*** *System input volume:* $687{,}660\ m^3 + 481{,}800\ m^3 + 0 = 1{,}169{,}460\ m^3$

Step 2.

- ***a.*** Define *billed metered consumption* (the type of consumers can be different from this example).
  - (**1**) Direct supply:
    - (***a***) Domestic consumption: 428.145 m^3
    - (***b***) Industrial consumption: 75.555 m^3
    - (***c***) Public consumption: 50.370 m^3
    - (***d***) Total direct consumption: 554.070 m^3
  - (**2**) Exported water: $16{,}126 \times 24 \times 365 = 141.264\ m^3$
  - (**3**) *Billed metered consumption:* $554.070\ m^3 + 141.264\ m^3 = 695.334\ m^3$
- ***b.*** Define *billed unmetered consumption* in column D.
  - (**1**) Unmetered consumers (12%, 150 L per capita per day): 75.555 m^3
  - (**2**) Landscape irrigation: 540 m^3
    - (***a***) *Example:* By estimating the average daily irrigation time, the average flow consumption per irrigation sprinklers, the approximate number of irrigation sprinklers, and the number of months per year of irrigation.
  - (**3**) Street cleaning: 375 m^3
    - (***a***) *Example:* By estimating the average number of vehicles used per day, the number of times each one is fulfilled per day, their water storage capacity, and the average number of days per year they are used.
  - (**4**) *Billed unmetered consumption:* $75.555\ m^3 + 540\ m^3 + 375\ m^3 = 76.470\ m^3$

c. Enter total in *billed authorized consumption* (column C) and *revenue water* (column E).

(**1**) *Billed authorized consumption* (= *revenue water*): 771.804 m^3 + 76.470 m^3 = 771.804 m^3

Step 3. Calculate the volume of (column E) as *system input volume* (column A) minus *revenue water* (column E).

a. *Nonrevenue water:* 1.169.460 m^3 − 771.804 m^3 = 397.656 m^3

Step 4.

a. Define *unbilled metered consumption* in column D.

(**1**) Undertaking's own consumption: 450 m^3

(**2**) Public consumption: 12.500 m^3

(**3**) Other: 0 m^3

(**4**) *Unbilled metered consumption:* 450 m^3 + 12.500 m^3 + 0 m^3 = 12.950 m^3

b. Define *unbilled unmetered consumption* in column D.

(**1**) Landscape irrigation: 20.000 m^3

(**2**) Street cleaning: 43.000 m^3

(**3**) Fire fighting: 5.000 m^3

(**4**) Pipe flushing and storage tank cleaning: 3.000 m^3

(**5**) Other: 0 m^3

(**6**) *Unbilled unmetered consumption:* 20.000 m^3 + 43.000 m^3 + 5.000 m^3 + 3.000 m^3 = 71.000 m^3

c. Assess *unbilled authorized consumption* and transfer total to column C.

(**1**) *Unbilled authorized consumption:* 12.950 m^3 + 71.000 m^3 = 83.950 m^3

Step 5. Add volumes of *billed authorized consumption* and *unbilled authorized consumption* in column C; enter sum as *authorized consumption* (column B).

a. *Authorized consumption:* 771.804 m^3 + 83.950 m^3 = 855.754 m^3

Step 6. Calculate *water losses* (column B) as the difference *between system input volume* (column A) and *authorized consumption* (column B).

a. *Water losses.* 1.169.460 m^3 − 855.754 m^3 = 313.706 m^3

Step 7.

a. Assess components of *unauthorized consumption* (column D) by best means available.

(**1**) Unauthorized temporary use of fire and irrigation hydrants

(**2**) Unauthorized connections

(**3**) *Unauthorized consumption:* unknown (10.000 to 20.000 m^3)

b. Assess components of *metering inaccuracies* (column D) by best means available.

(**1**) Meter reading and registration errors (10%): 10% × (695.334 + 12.950) = 70.828 m^3

(**2**) Inaccuracies of unmetered authorized consumption (20%): 20% $\times$ (76.940 + 71.000) = 29.588 m^3

(**3**) *Metering inaccuracies:* 70.828 m^3 + 29.588 m^3 = 100.416 m^3 (these values can be positive or negative)

c. Add *unauthorized consumption* and *metering inaccuracies* and enter sum in *apparent losses* (column C).

(**1**) *Apparent losses:* 15.000 + 100.416 = 115.416 m^3

Step 8. **Calculate *real losses*** (column C) as *water losses* (column B) minus *apparent losses* (column C).

a. *Real losses:* 313.706 m^3 − 115.416 m^3 = 198.290 m^3

Step 9. Assess components of *real losses* (column D) by best means available (night flow analysis, burst frequency/flow rate/duration calculations, modeling, etc.), add these and cross-check with volume of *real losses* in column C.

9.8.6 Assessment of the Indicators

Water Resources Indicators

$$\text{Inefficiency of use of water resources (\%)} = \frac{\text{volume of real losses}}{\text{system input volume} \times 100}$$

$$\frac{198{,}290}{1{,}169{,}460} \times 100 = 17.0\ \%$$

Financial Indicator

$$\text{Nonrevenue indicators (by volume) (\%)} = \frac{\text{volume of nonrevenue water}}{\text{system input volume}} \times 100$$

$$\left(\frac{397{,}656}{1{,}169{,}460}\right) \times 100 = 34.0\%$$

Operational Indicators

$$\text{Water losses (m}^3\text{ per connection per year)} = \frac{\text{volume of water losses}}{\text{number of service connections}}$$

$$\frac{313{,}706}{950} = 330\ \text{m}^3 \text{ per connection per year}$$

Real losses (L per connection per day)

$$= \frac{\text{real losses} \times 1000}{\text{number of service connections} \times 365 \times T/100}$$

where T is the percent of the year when the system is pressurized.

$$\frac{198{,}290 \times 1000}{(950 \times 365 \times 100/100)} = 572 \text{ L per connection per day}$$

Unmetered consumption (%) =

$$\frac{\text{Billed unmetered consumption} + \text{uqnbilled unmetered consumption)}}{\text{authorized consumption (\%)}}$$

$$\frac{76{,}470 + 71{,}000}{855{,}754} = 17.2\%$$

9.8.7 Results Interpretation

The results show that the situation of this network in terms of water losses is problematic. In financial terms, the efficiency is low, with 34 percent being nonrevenue water. Real losses of 572 L per connection per day are high, even taking into account that the average service pressure (40 m) is above average. The percentage of unmetered consumption, 17 percent, shows that the consumption assessed from estimates is still significant, although it demonstrates that the undertaking already manages to measure most of the authorized consumption. In a real situation, the interpretation would need to be completed with the analysis of the explanatory factors and the identification of the main aspects where there is room for improvement.

9.9 FINAL REMARKS

The systematic use of performance indicators, when preceded by an adequate planning and implementation and followed by an in-depth comparative analysis of the results and by the improvements derived from it, has a major leverage effect on the increase of efficiency and productivity of an organization.

The implementation of this type of management tool requires the active participation and engagement of the top managers in every stage of the process. The establishment of a team led by a senior staff member and representatives of the main departments of the organization that have clear responsibilities in the process is also a key factor for its success.

The selection of the PI system and of the specific PIs to adopt is another step of major importance. The use of standardized PI systems, such as the IWA's, is recommended as a generally better alternative to ad hoc systems developed expressly to a specific objective. A subset of relevant PIs shall be selected from the listing, and whenever necessary customized (e.g., changing metric units into imperial units) or expanding in some cases (e.g., splitting the PI into more detailed ones).

Any PI value is meaningless if the confidence of the input data is not known. Therefore, it is recommended that the reliability and accuracy of each variable required by the PIs be carefully assessed according to preset criteria.

All the team members shall in one way or another participate in the process of interpretation of the PI values, assisting in the identification and implementation of the procedures where there is room for improvement. This phase is critical to keep the staff motivated.

The continuous use of PIs allows the comparison of the actual results with management objectives and with the results of previous years. This type of analysis is in general more important than the comparison with other undertakings.

One could hardly choose better concluding remarks than the words of Finn Johansen, Director of Oslo Water and Sewage Works (Johansen and Helland, 2000):

> There is today a strong call for efficiency, productivity, quality and good service. As a director of municipally-owned water and wastewater undertaking, I experience these demands from the politicians and administrators in the City Hall and from the users of our services. And in addition there are requirements and demands from the State regulators and policy-making Departments. Last, but not least, we have the constant interest from the media and from environmental organizations and other pressure groups.…
>
> Let me make it clear that I don't experience the present call for efficiency and effectiveness as a threat, but rather as a stimulant for our organization to become even better.
>
> Benchmarking routines can be implemented internally in our organizations—comparing our results this year with last year's—or externally—comparing our performance with other undertakings with similar size and conditions, and preferably with undertakings that are better than us in certain areas or activities.…
>
> The most important benchmarking is done within each undertaking, when we compare the current year's PI with previous years' indicators. This is a relevant motivational factor for employees in our different divisions, who will use the all of PI used by the whole group plus several more specific division indicators.…
>
> The big question for us has been to find good performance indicators for benchmarking purposes.

Oslo is one of the 70 undertakings participating in the IWA PI field-testing project.

9.10 ACKNOWLEDGMENTS

The author expresses her acknowledgments to Wolfram Hirner, Jaime Melo Baptista, and Renato Parena, her coauthors of *Performance Indicators for Water Supply Services* in IWA Manual of Best Practices Series; and to Francisco Cubillo, Enrique Cabrera Jr., and Jaime Melo Baptista, members of the coordinating team of the ongoing field-testing phase.

A word of appreciation is also due to the 70 organizations who are participating in the field-testing of the IWA PI system and to everyone who contributed by hosting meetings, providing comments and suggestions, or assisting with the background work.

On behalf of the International Water Association, the author acknowledges her employer, the National Laboratory of Civil Engineering, for all the support provided to this project.

9.11 REFERENCES

Adamsson, J., *The Swedish Experience and Viewpoints,* IWSA Workshop on Performance Indicators for Transmission and Distribution Systems, Lisbon, Portugal, 1997.

Adamsson, J., et al., *Performance Benchmarking among 6 Cities in Scandinavia,* working report, 2000.

Alegre, H., J. M. Baptista, and A. L. Faria, *A General Framework of Performance Indicators in the Scope of Water Supply,* IWSA Workshop on Performance Indicators for Transmission and Distribution Systems, Lisbon, Portugal, 1997.

Alegre, H., W. Hirner, J. M. Baptista, and R. Parena, *Performance Indicators for Water Supply Services,* Manual of Best Practice Series, IWA Publishing, London, 2000.

Ancarani, A., "The Evaluation of the Quality of the Public Services Provided by Public/Private Firms: The Water Service Firms Case Study," Ph.D. Thesis in Economical and Managerial Engineering, Rome, December 1999.

AWA, "Australian Non Major Urban Water Utilities," *Performance Monitoring Report 1998–1999.* Australian Water Association, 2000.

AWWARF, "Distribution System Performance Evaluation," *Report of the American Water Works Research Foundation,* 1995.

AWWARF, "Performance Benchmarking for Water Utilities," *Report of the American Water Works Research Foundation,* 1996.

Cabrera, Jr., E., "Design of System for the Management Assessment of Urban Water Supply Systems," Ph.D. thesis, Universidade Politécnica de Valência, Valência, Spain, 2001 (in Spanish).

Cardoso, A., R. Matos, H. Alegre, and S. T. Coelho, "Performance Assessment of Water Supply and Wastewater Systems," *Seminar on Performance Indicators for Urban Drainage,* University of Hertfordshire, Reino Unido, December 6, 2000.

Cardoso, M. A., J. S. Matos, R. S. Matos, and S. T. Coelho, "A New Approach for Diagnosis and Rehabilitation of Sewerage Systems through the Development of Performance Indicators," *Proceedings of the 8th International Conference on Urban Storm Drainage,* Sydney, Australia, September 1999.

Coelho, T., *Performance Indicators in Water Supply—A Systems Approach,* Water Engineering and Management Systems Series, Research Studies Press, John Wiley and Sons, UK, 1997.

Deb, A., and L. Cesario, "Water Distribution Systems Performance Assessment," *IWSA Workshop on Performance Indicators for Transmission and Distribution Systems,* Lisbon, Portugal, 1997. Project funded by the AWWA Research Foundation.

Faria, A. L., and H. Alegre, "Paving the Way to Excellence in Water Supply Systems: a National Framework for Levels-of-Service Assessment Based on Consumer Satisfaction," The Maarten Schalekamp Award—1995, *AQUA,* Vol. 45, No. 1, pp. 1–12, 1996.

Guérin-Schneider, L., "Regulation of the Water Services in France and the Performance Indicators. Instrumentation and Organization," Ph.D. thesis, ENGREF—École National du Génie Rural, des Eaux et Forêts, Paris, 2001 (in French).

Johansen, F., and B. Helland, "Benchmarking Experience in Nordic Water and Wastewater Utilities," in *H2Objecttivo 2000—XI Edition,* Turin, Italy, May 2000, pp. 3–5.

MWA, *Performance Indicators for Water Supply,* A proposal from the Malaysian Water Association for the consideration of member countries of ASPAC, 1996.

MWA, *Malaysia Water Industry Guide 2001,* Water Supply Branch of the Malaysian Water Association, 2001.

Matos, R., M. A. Cardoso, and P. Duarte, "Preliminary Listing of Definitions, Context Information and Performance Indicators for Wastewater Services," Working document, *Workshop on Performance Indicators for Wastewater Services,* International Water Association, World Congress 2001, Berlin, Germany, 2001.

McIntosh, A. C., and C. Yñoguez, C., *Second Water Utilities Data Book—Asian and Pacific Region,* Asian Development Bank's Regional Technical Assistance No. 5694, Manila, Philippines, 1997.

Merkel, W., "Performance Assessment in the Water Industry," *International Report,* International Water Association, World Congress 2001, Berlin, Germany, 2001.

OFWAT, "Levels-of-Service for the Water Industry in England & Wales," *1999–2000 Report,* Office of Water Services, United Kingdom, 2000.

Paralez, L., "Status of Performance Assessment in Water and Wastewater Utility Systems in the United States," *National Report on Performance Assessment in the Water Industry,* International Water Association, World Congress 2001, Berlin, Germany, 2001, 11 pp.

van der Willigen, F., "Dutch Experience and Viewpoints on Performance Indicators," *IWSA Workshop on Performance Indicators for Transmission and Distribution Systems,* Lisbon, Portugal, 1997.

Westerhoff, G., D. Gale, P. Reiter, S. Haskins, and J. Gilbert, in J. B. Mannion (ed.), *The Changing Water Utility: Creative Approaches to Effectiveness and Efficiency,* American Water Works Association, 1998.

Yeppes, G., and A. Dianderas, *Water & Wastewater Utilities Indicators,* Water and Sanitation Division, The World Bank, 1996.

9.12 ENDNOTES

1. In England and Wales most domestic water consumption is not metered.
2. In 1999 the former International Water Services Association (IWSA) and the former International Association of Water Quality (IAWQ) merged, giving place to the IWA.
3. Text extracted from *Performance Indicators for Water Supply Services* in the IWA Manual of Best Practices Series.
4. Helena Alegre (LNEC, Portugal); Francisco Cubillo (Canal de Isabel II, Spain); Enrique Cabrera Jr. (ITA, Spain); Jaime Melo Baptista (LNEC, Portugal); Patrícia Duarte (LNEC, Portugal).

CHAPTER 10
RELIABILITY AND AVAILABILITY ANALYSIS OF WATER DISTRIBUTION SYSTEMS

Kevin Lansey
Department of Civil Engineering and Engineering Mechanics
University of Arizona
Tucson, Arizona

Larry W. Mays
Department of Civil and Environmental Engineering
Arizona State University
Tempe, Arizona

Y. K. Tung
Department of Civil Engineering
Hong Kong University of Science and Technology
Clear Water Bay, Kowloon, Hong Kong

10.1 FAILURE MODES FOR WATER DISTRIBUTION SYSTEMS

To ensure delivery of finished water to the user, the water distribution system must be designed to accommodate a range of expected emergency loading conditions. These emergency conditions may generally be classified into three groups: broken pipes, fire demands, and pump and power outages. Each of these conditions must be examined with an emphasis on describing its impact on the system, developing relevant measures of system performance, and designing into the system the capacity required to handle emergency conditions with an acceptable measure of reliability. *Reliability* is usually defined as the probability that a system performs its mission within specified limits for a given period of time in a specified environment.

Reviews of the literature (see Mays, 1989, 1992, 2000) reveal that there is currently no universally accepted definition or measure of the reliability of water distribution systems. For layer systems with many interactive subsystems, it is extremely difficult to analytically compute the mathematical reliability. Accurate calculation of the mathematical reliability of water distribution systems requires knowledge of the precise reliability of the basic subsystems or components of the water distribution system and the impact on mission accomplishment caused by the set of all possible subsystem (component) failures.

Many researchers, municipal engineers, urban planners, institutes, government agencies, and others have discussed the need to develop explicit measures and methodologies to evaluate water distribution system reliability and performance under emergency loading conditions. Some researchers have proposed candidate approaches using concepts of reliability factors, economic loss functions, forced redundancy in the designs, and so on. All of these approaches have limitations in problem formulation and/or solution technique. Some investigators discuss the need to explicitly incorporate measures of reliability into optimization models to predict system operation under emergency loading conditions. At present, however, no "optimization-reliability" evaluation or design technique with general application has been developed.

The reliable delivery of water to the user requires that the water distribution system be designed to handle a range of expected emergency loading conditions. These emergency conditions can be classified:

1. Fire demands
2. Broken links
3. Pump failure
4. Power outages
5. Control valve failure
6. Insufficient storage capability

A general application methodology that considers the minimum cost and reliability aspects must consider each of these emergency conditions. These conditions should be examined within the methodologies to

1. Describe their importance to the system
2. Develop relevant measures of system performance
3. Design into the system the capacity for the emergency loading conditions with an acceptable measure of reliability

10.1.1 Distribution Repair Definitions

The American Water Works Association Research Foundation (AWWARF) in conjunction with the U.S. Environmental Protection Agency developed a guidance manual entitled *Water Main Evaluation for Rehabilitation/Replacement* (O'Day et al., 1986). Portions of that document will be cited throughout this section. One aspect of the AWWARF manual was the clarification of terminology regarding leak and break reporting.

The lack of consistency between utilities defining leaks and breaks presents difficulties when trying to compare distribution system conditions. The general term *leak* can include both structural failures as well as water loss, which occurs from improperly sealed joints, defective service taps, and corrosion holes. *Breaks,* on the other hand, imply a structural failure of the water main, which results in either a crack or a complete severance of a water main. Breaks cause leakage and water loss; therefore, it can be argued that breaks are a special form of leaks.

Distribution system repairs can be classified by the component which was repaired (main, service, hydrant), the nature of the leak (joint leak, break), and the type of repair. Distribution repair records should identify the specific component repaired. Distribution system components include (O'Day et al., 1986):

- Main: barrel, joint, plug
- Hydrant: branch line, valve, cap barrel
- Line valve: bolts, bonnet, stem
- Service: tap, pipe, curb stop

A set of definitions was developed to help standardize the reporting and logging of distribution repairs (O'Day et al., 1986):

leak repair All actions taken to repair leaks in water mains, line valves, hydrant branches, and service lines are leak repairs.

main leak A water main leak includes all problems that lead to leakage of water from the main. These include joint leaks, holes, circumferential breaks, longitudinal breaks, defective tapes, and split bells. Hydrant service line and valve-related leaks are not considered main leaks.

joint leak Joint leaks represent a loss of water from the joint between adjacent main sections. This is not a structural problem, rather it is a separation of the main sections caused by expansion and contraction, settlement, movement, or movement of joint materials because of pressure or pipe deflection.

main breaks Main breaks represent the structural failure of the barrel or bell of the pipe due to excessive loads, undermining of bedding, contact with other structures, corrosion, or a combination of these conditions. Main break types include (1) circumferential, (2) longitudinal, and (3) split bells, including bell failures from sulphur compound joint materials.

hydrant leak repairs All actions taken to repair hydrant branch lines, hydrant valves, and hydrant barrels are hydrant repairs.

service leak repairs All actions taken to repair leaking taps, corporation stops, service pipes, and curb stops are service leak repairs.

valve leak repairs All actions taken to repair leaks in valve flanges, bonnet, or body are valve leak repairs.

10.1.2 Failure Modes

Although water distribution networks have many features in common with other types of networks (Templeman and Yates, 1984), there are a number of significant differences between them when they are examined from a failure-based point of view. The differences arise primarily as a result of the ways in which water distribution networks can "fail." In its most basic sense, *failure* of water distribution networks can be defined as either the pressure and/or flow falling below specified values at one or more nodes within the network.

Under such a definition, there are two major modes of failure of water distribution networks: (1) *performance failure,* i.e., demand on the system being greater than the design value; and (2) *component (mechanical) failure,* which includes the failure of individual network components, e.g., pipes, pumps, and valves. Both these modes of failure have probabilistic bases which must be incorporated into any reliability analysis of networks.

Performance Failure. In the first case, the network fails in a traditional sense of the load exceeding the design capacity. However, it is not necessarily catastrophic, as might be the case of failure of a structural network, e.g., a truss. Rather, either the pressure at one or more nodes may drop below the required minimum and/or the amount of flow available at the node(s) may fall below the required level. The implications of such a failure are difficult to define as they may constitute only a decreased level of services to domestic households over a relatively brief period of time, e.g., 6 h. On the other hand, the failure of the water supply when there is a major fire could result in increased monetary damages and possible loss of life.

The probabilistic aspect of failure in this first case arises from the probability distribution of the loading condition under which the distribution network is expected to operate. In most cases, the probability distributions of the demands, which the network must fulfill, are reasonably well known. The choice of a design value for the demands on the network, for example, a 10-year, 6-h maximum flow, is not a deterministic decision and implies acceptance of a probability of the network not being able to perform to specifications. What has been missing in network analysis to date is an explicit recognition of the probabilistic processes.

The implications of failure as a result of hydraulic loads being greater than design loads are similar to those resulting from component failure; flows and/or supply pressures can fall below the desired levels. The means of improving the reliability of water distribution as measured by the probability of either failure occurring are also similar. Obviously, larger component sizes will provide greater overall flow capacity in the network and therefore reduce the probability that actual demands exceed the design values. In addition, previous work (e.g., Kettler and Goulter, 1985; Mays, 1989) has shown that pipe breakage rates in existing distribution networks are strongly correlated to diameter. Larger-diameter pipes break less frequently than smaller-diameter pipes. Thus, selection of larger-diameter pipes can reduce the probability of component failure while simultaneously improving the flow-exceedance-based measure of reliability.

It is important to note, however, that good performance according to a reliability measure based on one mode of failure does not necessarily mean good performance according to the other. Consider, for example, a large pipe near the source of supply for a particular network. Since this pipe is near the source, a larger percentage of the flow passing to more distant parts of the network will flow through it and, therefore, a large-diameter, large-capacity pipe will be required. This large-diameter pipe will have a low frequency of breakage and would therefore indicate good reliability according to component failure probabilities. The probability that the flow that same pipe is required to carry will be greater than its design capacity may, however, be quite large. If this is the case, its *flow-exceedance* reliability is high. Improvements to both measures of reliability may still be achieved through the selection of larger-diameter pipes.

The improvements to network reliability that are achieved through specification of large pipes can also be achieved through selection of larger capacities for the other important components of a network. By the provision of more pumps, pump station failure is reduced. The additional pumps also provide additional capacity for the pump station thereby reducing the probability that demands on the station will exceed the design capacity. A similar argument can be made for facilities on the network.

Component (Mechanical) Failure. The probability of component failure can be derived from historical failure records and modeled using an appropriate probability distribution (see Sec. 10.2 for time-to-failure analysis). Goulter and Coals (1986) proposed a Poisson probability distribution with parameter λ breaks per kilometer per year (breaks/km/yr) for pipe breakage. The probability of demands being greater than design values can also be derived from historical data and modeled appropriately. The most difficult problem, however, is that even knowing the appropriate probability of the two modes of failure it is difficult to define reliability explicitly. Reduction in either the probability of component failure or the probability of demands being greater than design capacity obviously improves system reliability. The question is, by how much? If a pipe (or any other component) fails, what is the effect on the remainder of the network? Obviously, the more critical the pipe in question to the network, the more widespread and serious is the effect. The actual quantification of the effect is very difficult to assess without extensive simulation of network operation. Similarly, the networkwide implication of a particular demand being greater than the design value for that area is difficult to assess without considerable effort. Furthermore, if network reliability, as defined from measures based on either mode of failure is unsatisfactory, What is the best way to improve it?

There is currently no universally accepted definition or measure of the reliability of water distribution systems. This discussion is a general definition of terms for the reviewer not familiar with the reliability and availability aspects. *Reliability* is usually defined as the probability that a system performs its mission within specified limits for a given period of time in a specified environment. For a large system, with many interactive subsystems (such as a water distribution system), it is extremely difficult to analytically compute a mathematical reliability. Accurate calculation of a mathematical reliability requires knowledge of the precise reliability of the basic subsystems or components and the impact on mission accomplishment caused by the set of all possible subsystem (component) failures. See Cullinane et al. (1992) and Mays (1989, 2000) for complete discussions of the aspects of reliability and availability analysis.

The traditional engineering definition of reliability is the ability of a system to meet demand. Some authors (see Cullinane et al., 1992) prefer to label this concept "availability," while reserving the term reliability for the length of time that a system can be expected to perform without failure. The term reliability in its general sense is defined as any measure of the system's ability to satisfy the requirements placed upon it.

10.1.3 Reliability Indices

A number of different measures of *reliability indices* have been proposed (see Cullinane et al., 1992), but no single one is universally accepted. Each can be useful in reliability analyses, depending on the purpose of the analysis.

Reliability is often defined as the ability of a system to meet demand under a defined set of contingencies. A system may be judged reliable if it can, for example, satisfy demand during a period of drought or despite the failure of a key piece of equipment. Early design criteria were based on such rules of thumb instead of quantitative indices, such as those we will soon discuss. Contingency criteria suffer from arbitrariness in choice of contingency and inconsistency, because different designs able to withstand the same contingency can have different probabilities of failure.

Availability is defined as the probability at a given moment that the system will be found in a state such that demand does not exceed available supply or capacity or that operating conditions are not otherwise unsatisfactory (e.g., low water pressure).

Average availability is the mean probability over a period of time of being found in such a state. In this and Sec. 10.2, the term *availability* will be used in the sense of average availability. The probability of system failure is sometimes defined as 1 minus availability.

Severity indices describe the size of failures. In water supply studies, *reliability* has been defined as the ratio of available annual supply to demand.

Frequency and duration indices indicate how often failures of a given severity occur and how long they last. Such indices have found increasing use in electric utility planning studies.

Economic indices are indices of the economic consequences of shortages, also referred to as *vulnerability.* Letting S_u represent water supply and Q the quantity of water demanded in units of m^3/h, these indices can be phrased as follows:

- *Contingency.* A system is reliable if $P(S_u < Q + \text{contingency}) = 0$.
- *Availability.* The availability of the system is $P(S_u > 0)$.
- *Probability and severity of failure.* The probability of a failure can be defined as $P(S_u < Q - s)$ at a particular moment or over a particular period, where s represents some predefined level of failure severity.
- *Frequency and duration.* The *expected frequency* $E(F)$ (in units of $1/T$) of a failure event of at least severity s in m^3/h is related to the expected duration $E(D)$ by the equation $E(F)E(D) = P(S_u < Q - s)$.

Economic consequences (measured in dollars) are related nonlinearly to the frequency, duration, and severity of events.

Many models have been reported in the literature for the reliability analysis of water distribution systems. Unfortunately the water distribution industry (engineering community) has not incorporated these methodologies into practice. There are many reasons for this, two of which are that engineers are not familiar with the methodologies and the proper software has not been developed to make this an easier task. The list of references includes several of the reliability methodologies that have been reported in the literature: Awumah et al. (1991), Awumah and Goulter (1992), Bao and Mays (1994a,b), Biem and Hobbs (1988), Boccelli et al. (1998), Bouchart and Goulter (1989, 1991, 1992), Cullinane et al. (1989, 1992), Duan and Mays (1990), Duan et al. (1990), Fujiwara and de Silva (1990), Fujiwara and Gannesharajah (1993), Fujiwara and Li (1998), Fujiwara and Tung (1991), Germanopoulis et al. (1986), Goulter (1988, 1992a,b, 1995a,b), Goulter and Bouchart (1990), Goulter and Coals (1986), Goulter and Jacobs (1989), Gupta and Bhave (1994), Hobbs and Biem (1988), Jacobs (1992), Jacobs and Goulter (1988, 1989, 1991), Jowitt and Xu (1993), Kessler et al. (1990), Kettler and Goulter (1983, 1985), Lansey et al. (1989, 1992), Ormsbee and Kessler (1990), Park and Liebman (1993), Quimpo and Shamsi (1991), Shamsi (1990), Shinstine (1999), Su et al. (1987), Tanyimboh et al. (2001), Tanyimboh and Templeman (1993, 1995), Wagner et al. (1988a,b), Wu et al. (1993), Xu and Goulter (1998), Yang et al. (1996a,b).

10.2 COMPONENT (MECHANICAL) RELIABILITY ANALYSIS

A system or its components can be treated as a black box or a lumped-parameter system, and their performances are observed over time. This reduces the reliability analysis to a one-dimensional problem involving time as the only variable. In such cases, the *time to failure* (TTF) of a system or a component of the system is the random variable. It should be pointed out that the term *time* could be used in a more general sense. In some situations other physical scale measures, such as distance or length, may be appropriate for system performance evaluation.

The time-to-failure analysis is particularly suitable for assessing the reliability of systems and/or repairable components. For a system that is repairable after its failure, the time period it would take to have it repaired to the operational states is uncertain; therefore, the *time to repair* (TTR) is also a random variable.

This section will focus on characteristics of failure, repair, and availability of repairable systems by time-to-failure analysis.

10.2.1 Failure Density, Failure Rate, and Mean Time to Failure

Any system will fail eventually; it is just a matter of time. Because of the presence of many uncertainties that affect the operation of a physical system, the time that the system fails to satisfactorily perform as intended is a random variable.

The *failure density function* is the probability distribution that governs the time occurrence of failure. The failure density function serves as the common thread in the reliability assessments in time-to-failure analysis. The reliability of a system or a component within a specified time interval $[0, t]$ can be expressed as

$$p_s(t) = P(\mathrm{TTF} > t) = \int_t^\infty f(\tau)\, d\tau \qquad (10.1)$$

where TTF is the random time to failure having $f(t)$ as the failure density function. The reliability $p_s(t)$ represents the probability that the system experiences no failure within $[0, t]$. The *failure probability* or *unreliability* can be expressed as $p_f(t) = 1 - p_s(t)$. Note that unreliability $p_f(t)$ is the probability that a component or a system would experience its first failure within the time interval $[0, t]$. Conversely, the failure density function can be obtained from reliability or unreliability as

$$f(t) = -\frac{d[p_s(t)]}{dt} = \frac{d[p_f(t)]}{dt} \qquad (10.2)$$

The time to failure is a continuous, nonnegative random variable by nature. Many distribution functions described in Sec. 10.2.1 are appropriate for modeling the stochastic nature of time to failure. Among them, the exponential distribution is perhaps the most widely used. The exponential distribution has been found, both phenomenologically and empirically, to adequately describe the time-to-failure distribution for components, equipment, and systems involving components with mixtures of life distributions.

The *failure rate* is defined as the number of failures occurring per unit time in a time interval $(t + \Delta t)$ per unit of the remaining population at time t. The instantaneous failure rate $m(t)$ can be obtained as

$$m(t) = \lim_{\Delta t \to 0} \left[\frac{N_F(\Delta t)/\Delta t}{N(t)} \right] = \frac{f(t)}{p_s(t)} \tag{10.3}$$

where $N_F(\Delta t)$ is the number of failures in a time interval $(t, t\ \Delta t)$ and $N(t)$ is the number of failures from the beginning up to time t. This *instantaneous failure rate* is also called the *hazard function* or *force of mortality function*.

The failure rate for many systems or components has a bathtub shape, as shown in Fig. 10.1, in that three distinct life periods can be identified. They are *early life* (or *infant mortality*) *period*, *useful life period*, and *wear−out life period*. In the early life period, quality failures and stress-related failures dominate, with little contribution from wear-out failures. During the useful life period, all three types of failures contribute to the potential failure of the system or component, and the overall failure rate increases with age, because wear-out failures and stress-related failures are the main contributors, and wear out becomes a more dominating factor for the failure of the system with age.

The reliability can be directly computed from the failure rate as

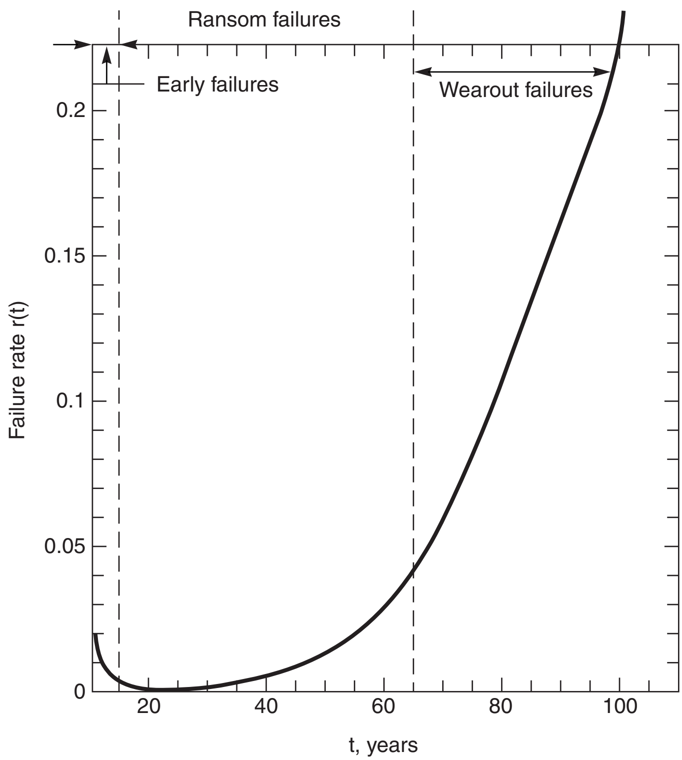


FIGURE 10.1 Failure rate $r(t)$ versus t. (Henley and Kumamoto, 1981)

$$p_s(t) = \exp\left(\int_0^t m(\tau)\, d\tau\right) \tag{10.4}$$

Substituting Eq. (10.4) into Eq. (10.3), the *failure density function* $f(t)$ can be expressed, in terms of failure rate, as

$$f(t) = m(t)\left[\exp\left(\int_0^{-t} m(\tau)\, d\tau\right)\right] \tag{10.5}$$

In general, the reliability of a system or a component is strongly dependent on its age. This can be mathematically expressed by the condition probability as

$$p_s(\xi\,|t) = \frac{p_s(t+\xi)}{p_s(t)} \tag{10.6}$$

in which t is the age of the system and up to that point the system has not failed and in which $p_s(\xi\,|t)$ is the reliability over a new mission period ξ, having successfully operated over a period of t.

A commonly used reliability measure of system performance is the *mean time to failure* (MTTF) which is the expected time to failure. The MTTF can be mathematically defined as

$$\text{MTTF} = E(\text{TTF}) = \int_0^\infty \tau f(\tau)\, d\tau \tag{10.7}$$

The mean time to failure for some failure density functions is given in Table 10.1.

Repairable Systems. For repairable water resources systems, such as pipe networks, pump stations, and storm runoff drainage structures, failed components within the system can be repaired or replaced so that the system can be put back into service. The time required to have the failed system repaired is uncertain, and consequently, the total time required to restore the system from its failure to operational states is a random variable.

Like the time to failure, the random time to repair (TTR) has the repair density function $g(t)$ describing the random characteristics of the time required to repair a failed system when failures occur at time 0. The *repair probability* $G(t)$ is the probability that the failed system can be restored within a given time period $[0, t]$,

$$G(t) = P[\text{TTF} \le t] = \int_0^t g(\tau)\, d\tau \tag{10.8}$$

The repair probability $G(t)$ is one measure of maintainability.

The *repair rate* $r(t)$, similar to the failure rate, is the conditional probability that the system is repaired per unit time, given that the system failed at time 0 and is

TABLE 10.1 Mean Time to Failure for Some Failure Density Functions

Distribution	PDF, $f(t)$	Reliability, $p_s(t)$	Failure rate, $m(t)$	MTTF
Normal	$\frac{1}{\sqrt{2\pi}\sigma_T} \exp\left[-\frac{1}{2}\left(\frac{t-\mu_r}{\sigma_T}\right)^2\right]$	$\phi\left(\frac{t-\mu_T}{\sigma_T}\right)$	$\phi\left(\frac{t-\mu_T}{\sigma_T}\right)$	μ_T
Lognormal	$\frac{1}{\sqrt{2\pi}\sigma_{\ln T}t} \exp\left[-\frac{1}{2}\left(\frac{\ln(t)-\mu_{\ln T}}{\sigma_{\ln T}}\right)^2\right]$	$\phi\left(\frac{\ln(t)-\mu_{\ln T}}{\sigma_{\ln T}}\right)$	$\sigma\left(\frac{\ln(t)-\mu_{\ln T}}{\sigma_{\ln T}}\right)$	$\exp\left(\mu_{\ln T}+\frac{\sigma_{\ln T}^2}{2}\right)$
Exponential	$\beta e^{-\beta t}$	$e^{-\beta t}$	β	$\frac{1}{\beta}$
Rayleigh	$\frac{t}{\beta^2} \exp\left[-\frac{1}{2}\left(\frac{t}{\beta}\right)^2\right], \beta > 0$	$\exp\left[-\frac{1}{2}\left(\frac{t}{\beta}\right)^2\right]$	$\frac{t}{\beta^2}$	$1.253\,\beta$
Gamma	$\frac{\beta}{\Gamma(\alpha)} (\beta t)^{\alpha-1} e^{-\beta t}$	$\int_t^\infty f(\tau)\,d\tau$	$\frac{f(t)}{p_s(t)}$	$\frac{\alpha}{\beta}$
Gumbel	$e^{\pm y} - e^{\pm y};\ y = \frac{t-t_0}{\beta}$	$1 - e^{e^{\pm y}}$	$\frac{f(t)}{p_s(t)}$	$x_0 \pm 0.577\,\beta$
Weibull	$\frac{\alpha}{\beta}\left(\frac{t-t_0}{\beta}\right)^{\alpha-1} e - \left(\frac{t-t_0}{\beta}\right)^a$	$\exp\left(\frac{t-t_0}{\beta}\right)^a$	$\frac{\alpha(t-t_0)^{\alpha-1}}{\beta^\alpha}$ $f(t)$	$t_0 + \beta\Gamma\left(1+\frac{1}{\alpha}\right)$
Uniform	$\frac{1}{b-a}$	$\frac{b-t}{b-a}$	$f(t)$ $\frac{1}{b-t}$	$\frac{a+b}{2}$

still not repaired at time *t*. The quantity $r(t)\,dt$ is the probability that the system is repaired during the time interval $(t, t + dt)$ given that the system fails at time *t*. The relation between the repair density function, repair rate, and repair probability is

$$r(t) = \frac{g(t)}{1 - G(t)} \tag{10.9}$$

Given a repair rate $r(t)$, the density function and the maintainability can be determined, respectively, as

$$g(t) = r(t) \exp\left(\int_0^{-t} r(\tau)\,d\tau\right) \tag{10.10}$$

$$G(t) = 1 - \exp\left(\int_0^{-t} r(\tau)\,d\tau\right) \tag{10.11}$$

The *mean time to repair* (MTTR) is the expected value of the time to repair of a failed system, that is,

$$\text{MTTR} = \int_0^\infty tg(t)\,dt \tag{10.12}$$

The MTTR measures the elapsed time required to perform the maintenance operation and is used to estimate the downtime of a system. It is also a measure for maintainability of a system. Various types of maintainability measures are derivable from the repair density function.

The MTTR is a proper measure of the mean life span of a nonrepairable system. For a repairable system, a more representative indicator for the fail-repair cycle is the *mean time between failures* (MTBF) which is the sum of MTTF and MTTR, that is,

$$\text{MTBF} = \text{MTTF} + \text{MTTR} \tag{10.13}$$

The *mean time between repairs* (MTBR) is the expected value of the time between two consecutive repairs, and it is equal to MTBF.

10.2.2 Availability and Unavailability

A *repairable system* experiences a repetition of repair-to-failure and failure and failure-to-repair processes during its service life. Hence, the probability that a system is in operating condition at any given time *t* for a repairable system is different than that of a nonrepairable system. The term availability $A(t)$ is generally used for repairable systems, to indicate the probability that the system is in operating condition at any given time *t*. On the other hand, reliability $p_s(t)$ is appropriate for nonrepairable systems, indicating the probability that the system has been continuously in its operating state starting from time 0 up to time *t*.

Availability can also be interpreted as the percentage of time that the system is in operating condition within a specified time period. In general, the availability and reliability of a system satisfies the following inequality relationship

$$0 \leq p_s(t) \leq A(t) \leq 1 \tag{10.14}$$

The equality holds for nonrepairable systems. The reliability of a system decreases monotonically to 0 as the age of the system increases, whereas the availability decreases but converges to a positive probability, as shown in Fig. 10.2.

The complement to availability is the *unavailability* $U(t)$, which is the probability that a system is in the failed condition at time *t,* given that it is in operating condition at time 0. In other words, unavailability is the percentage of time that the system is not available for the intended service in time period [0, *t*], given it is operational at time 0. Availability, unavailability, and unreliability satisfy the following relationships:

$$A(t) + U(t) = 1.0 \tag{10.15}$$

$$0 \leq U(t) \leq p_f(t) < 1 \tag{10.16}$$

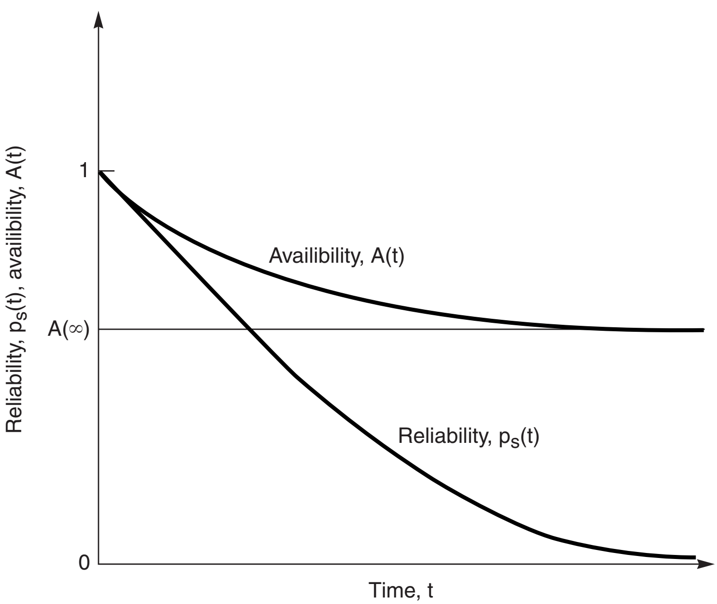


FIGURE 10.2 Variation of reliability and availability with time. (Tung, 1996)

For a nonrepairable system, the unavailability is equal to the unreliability, that is, $U(t) = p_f(\mathrm{t})$.

Determination of availability or unavailability of a system requires full accounting of the failure and repair processes. The basic elements that describe processes are failure and repair processes. The basic elements that describe such processes are the failure density function $f(t)$ and the repair density function $g(t)$. Consider a system with a *constant failure rate* λ and a *constant repair rate* η. The availability and unavailability, respectively, are

$$A(t) = \frac{\eta}{\lambda + \eta} + \frac{\lambda}{\lambda + \eta}\, e^{-(\lambda+\eta)t} \tag{10.17}$$

$$U(t) = \frac{\lambda}{\lambda + \eta}\left[1 - e^{-(\lambda+\nu)t}\right] \tag{10.18}$$

As the time approaches infinity ($t \rightarrow \infty$), the system reaches its stationary condition. Then, the *stationary availability* $A(\infty)$ and *stationary unavailability* $U(\infty)$ are

$$A(\infty) = \frac{\eta}{\lambda + \eta} = \frac{1/\lambda}{1/\lambda + 1/\eta} = \frac{\mathrm{MTTF}}{\mathrm{MTTF}} + \mathrm{MTTR} \tag{10.19}$$

$$U(\infty) = \frac{\lambda}{\lambda + \eta} = \frac{1/\eta}{1/\lambda + 1/\eta} = \frac{\mathrm{MTTR}}{\mathrm{MTTF}} + \mathrm{MTTR} \tag{10.20}$$

10.3 METHODOLOGY FOR RELIABILITY AND AVAILABILITY ANALYSIS FOR WATER DISTRIBUTION NETWORKS

10.3.1 Reliability of a System

The methodology presented here summarizes the work of Shinstine et al. (2002). *Reliability* of a water distribution system can be defined using the approach of Goulter (1995):

> The ability of a water distribution system to meet the demands that are placed on it where such demands are specified in terms of:
>
> **1.** The flows to be supplied (total volume and flow rate); and
> **2.** The range of pressures at which those flows must be provided

A useful extension or interpretation of this definition as given by Cullinane et al. (1992) is "the ability of the system to provide service with an acceptable level of interruption in spite of abnormal conditions."

A key feature of this latter definition is that it implies, or introduces, the concept of both the period of time over which the system is unable to meet demands and the particular point in time (i.e, circumstances which caused the system to be unable to supply demands) as being important determinants in defining and calculating reliability. This interpretation of reliability also recognizes that system failure may occur either through the flows associated with the demands not being supplied, and/or the flows associated with the demands being supplied but at pressures lower than the minimum acceptable or minimum specified for the particular circumstances. Of equal importance is that this definition also recognizes that failures should only occur in association with abnormal conditions, e.g., component failure and/or abnormally high demands.

The range of combinations of ways in which failure can occur in a water distribution system constitutes one, and arguably the major, source of the many theoretical and practical difficulties which have been encountered in establishing suitable, i.e., comprehensive and computationally tractable, measures of reliability which can be used in the practical design and operation of water distribution systems. It does have to be recognized, however, that reliability has been an explicit factor in the design and operation of water distribution networks for a considerable time as demonstrated by the existence of looped networks. Looped networks are not the most cost-efficient solution for water supply networks. The presence of loops does, however, add reliability and redundancy to the system by providing excess capacity and alternative paths for the supply of water in the face of failure of one of the system components. The question facing practitioners in the water distribution industry is how to make quantitatively explicit the reliability of supply. This quantitative specification of reliability can then be used by municipal authorities and the like to define compliance levels for the reliability to be provided by the water utility operators both private and public. Such a specification can similarly be used as a point against which the operators can fine-tune and optimize the performance of the system.

Reliability models to compute system reliability for a given system and layout have been developed by Cullinane (1985), Tung (1985), Wagner et al. (1988a,b), Duan and Mays (1990), and Bao and Mays (1990). These approaches allow a modeler to determine the reliability of a system and account for such factors as the probability and duration of pipe and pump failure, the uncertainty in demands, and the variability in the decay of pipes. Notably, Tung (1985) defines system reliability as "probability that flow can reach all the demand points in the network." Among the six reliability methods identified, the cut-set method, defined later in this chapter, appeared to be the most efficient and is easily programmed on a digital computer.

Hydraulic-failure reliability measures have been included in optimization models by Su et al. (1987), Goulter and Bouchart (1990), Lansey et al. (1989), and Duan et al. (1990). Using the minimum cut-set method as the reliability model for the system and the nodes, Su et al. (1987) added a reliability constraint to an optimal network design model.

The second category of reliability measures overcomes the computational problems. Examples include Goulter and Coals (1986) and Shamir and Howard (1981). Goulter and Coals's (1986) model incorporates reliability aspects into a linear programming model. One disadvantage of their model is their assumption that all pipes connecting a node have similar diameters and hence have similar values of failure probability, which is not applicable in real pipe networks. Su et al.'s (1987) model is one example where this problem is solved.

The results from several of the above heuristic measures generally appear consistent with expectations of systems as reliability requirements are increased. However, since the system pressures are not evaluated in these techniques, over- or under-designed components may result (Cullinane et al., 1992). Therefore, Cullinane et al. (1992) described an approach that is a balance between the two types of measures reported by Mays (1989). Hydraulic failure is verified for several conditions, but not for all as required by other failure-based techniques. Thus, computational burden is reduced, but heuristics and judgment are necessary to define critical conditions.

10.3.2 Methodology

Hydraulic Availability. "*Hydraulic availability*, as it relates to water distribution system design, can be defined as the ability of the system to provide service with an acceptable level of interruption in spite of abnormal conditions" (Cullinane et al., 1992). The evaluation of hydraulic availability relates directly to the basic function of the water distribution system, i.e., delivery of the specified quantity of water to a specific location at the required time under the desired pressure.

Availability is evaluated in terms of developing a required minimum pressure. Generally, for fire conditions the desirable minimum pressure is 137.9 kilonewtons per square meter (kN/m^2) (20 psi) with an absolute minimum pressure of 68.9 kN/m^2 (10 psi). These pressures are required to overcome the head losses between hydrants and fire-engine pumps and are essentially estimates of the net positive suction head (NPSH) required for pump operation. Pressures between 137.9 kN/m^2 (20 psi) and 551.6 kN/m^2 (80 psi) are considered to be desirable pressures under normal daily demands.

Previous works (Su et al., 1987; Goulter and Coals, 1986) proposed the use of a discrete relationship (Fig. 10.3) between availability and pressure. Here, availability during a time period t can be expressed by the following mathematical relationship:

$$\mathrm{HA}_j = \begin{cases} 1 & \text{for } P_j \geq \mathrm{PR} \qquad (10.21) \\ 0 & \text{for } P_j < \mathrm{PR} \qquad (10.22) \end{cases}$$

where HA_j = hydraulic availability of node j
P_j = pressure at node j
PR = required minimum pressure

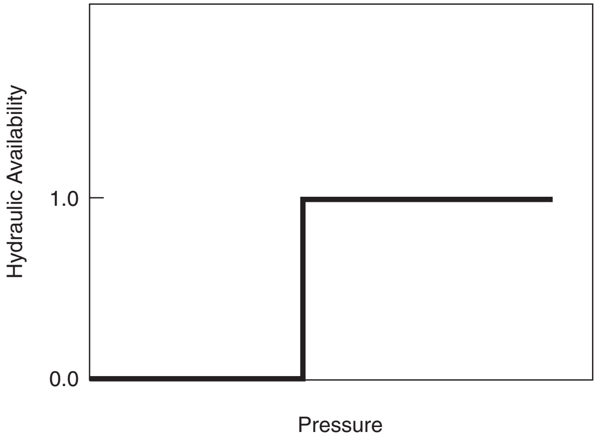


FIGURE 10.3 Hydraulic availability step function. (Shinstine et al., 2002)

The *network hydraulic availability* is the product of the nodal hydraulic availabilities. With this approach, a zero availability index is assigned for all pressure values below the required minimum pressure. For example, if the required minimum pressure is set at 137.9 kN/m^2 (20 psi), a residual of 137.2 kN/m^2 (19.9 psi) results in the same availability index as a residual pressure of 6.9 kN/m^2 (1 psi). Thus, the use of this discontinuous relationship does not adequately represent the engineering reality of the problem.

Cullinane et al. (1992) adopted a more appropriate representation that describes the availability index as a continuous "fuzzy" function. Using a continuous function of this shape, a significant index value may be assigned to pressure values slightly less than the arbitrarily assigned required minimum pressure value PR. The shape of the curve shown in Fig. 10.4 is similar to the cumulative normal distribution, which is mathematically stated as follows:

$$\mathrm{HA}_j = P(\mathrm{PR} \le P_j) = \frac{1}{\sqrt{2\pi}} \int^{(H-\mu_H)/\sigma_H} e^{-t^{5/2}}\, dt = P\left(\frac{H-\mu_H}{\sigma_H}\right) \tag{10.23}$$

where P_j = value of nodal pressure
μ_H = mean nodal pressure
σ_H = standard deviation of pressure

Values of μ_H and σ_H = can be selected to adjust the position and shape of the function, respectively.

Pipe Failure Probabilities. Goulter and Coals (1986) and Su et al. (1987) used similar methods to determine the probability of failure of individual pipes. The probability of failure of pipe *i*, P_i, is determined using the *Poisson probability distribution*:

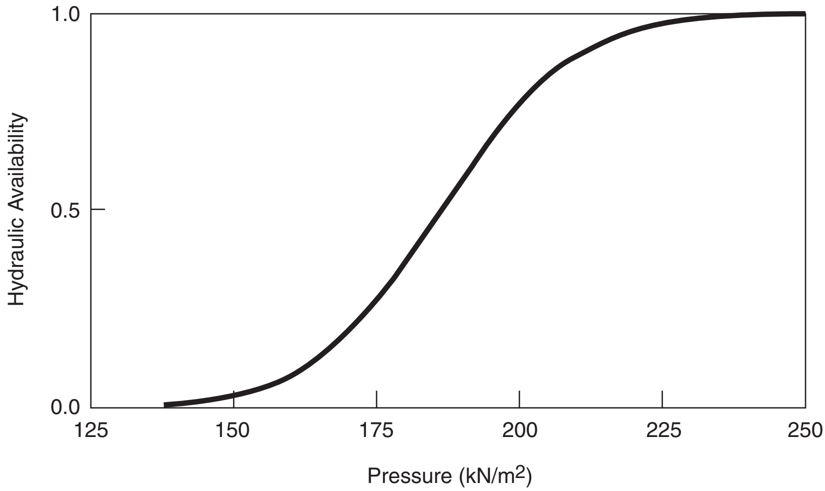


FIGURE 10.4 Continuous hydraulic availability function. (Shinstine et al., 2002)

$$P_i = 1 - e^{-\beta_i} \tag{10.24}$$

and

$$\beta_i = r_i L_i \tag{10.25}$$

where β_i = expected number of failures per year for pipe i
r_i = expected number of failures per year per unit length of pipe i
L_i = the length of pipe i

Nodal and System Reliability. The *nodal* and *system reliabilities*, R_{node} and R_S, respectively, are obtained from minimum cut-sets. Su et al. (1987) defines the *minimum cut-set* as "a set of system components (e.g., pipes) which, when failed, causes failure of the system." However, when any one component of the set has not failed, it does not cause system failure (Billinton and Allan, 1983).

Assuming that a pipe break can be isolated from the rest of the system, minimum cut-sets are determined by closing a pipe or combination of pipes in the water distribution system and using a hydraulic simulation model (e.g., KYPIPE) to obtain the values of pressure head at each demand node of the system. By comparing these pressure heads with minimum pressure-head requirements, the reliability model can determine whether or not this pipe or combination of pipes is a minimum cut-set of the system or an individual demand node. *A minimum cut-set for a node* is one that causes reduced hydraulic availability at that node, while a *minimum cut-set for the system* is a cut-set that reduces the hydraulic availability for any node in the system. This procedure is repeated until all the combinations of pipes have been considered and hence all minimum cut-sets of the demand nodes and the total system have been determined. The values of system and nodal reliability are then computed. A flowchart of the procedure is shown in Fig. 10.5. Note that prior to hydraulic simulation a test is performed to examine if the failure(s) cause a part of the system to be isolated from all sources and thus have no available supply. This disconnection results in failure of all isolated nodes.

For n components (pipes) in the ith minimum cut-set of a water distribution system, the failure probability of the *ith minimum cut-set* (MC_i) is

$$P(MC_i) = \prod_{i=1}^{n} P_i = P_1 \cdot P_2 \cdot P_3 \cdot \cdots \cdot P_n \tag{10.26}$$

Using the step function for hydraulic availability and assuming that the occurrence of the failure of the components within a minimum cut-set are statistically independent for a water distribution network with four minimum cut-sets (MC_i), for the system reliability R_s, the failure probability of the system P_s is then defined as (Billinton and Allan, 1983):

$$P_s = P(MC_1 \cup MC_2 \cup MC_3 \cup MC_4) \tag{10.27}$$

By applying the principle of inclusion and exclusion (Ross, 1985), Eq. (10.27) can be reduced to

$$P_s = P(MC_1) + P(MC_2) + P(MC_3) + P(MC_4) \tag{10.28}$$

$$= \sum_{i=1}^{4} P(MC_i) \tag{10.29}$$

In general form,

$$P_s = \sum_{i=1}^{M} P(MC_i) \tag{10.30}$$

The *system reliability* R_s is expressed as

$$R_s = 1 - P_s = 1 - \sum_{i=1}^{M} P(MC_i) \tag{10.31}$$

where M is the number of minimum cut-sets in the system. To further provide physical significance to the reliability, it is possible to weight the nodal terms as a function of the nodal demand. Nodal reliabilities can be computed with the same relationship including only failures that affect the individual node.

Using the continuous hydraulic availability concept, a true minimum cut-set does not exist. The probability of a cut-set occurring is consistent; however, the *reliability* is defined as the product of pipe reliability and hydraulic unavailability (1 − HA). The system reliability is then,

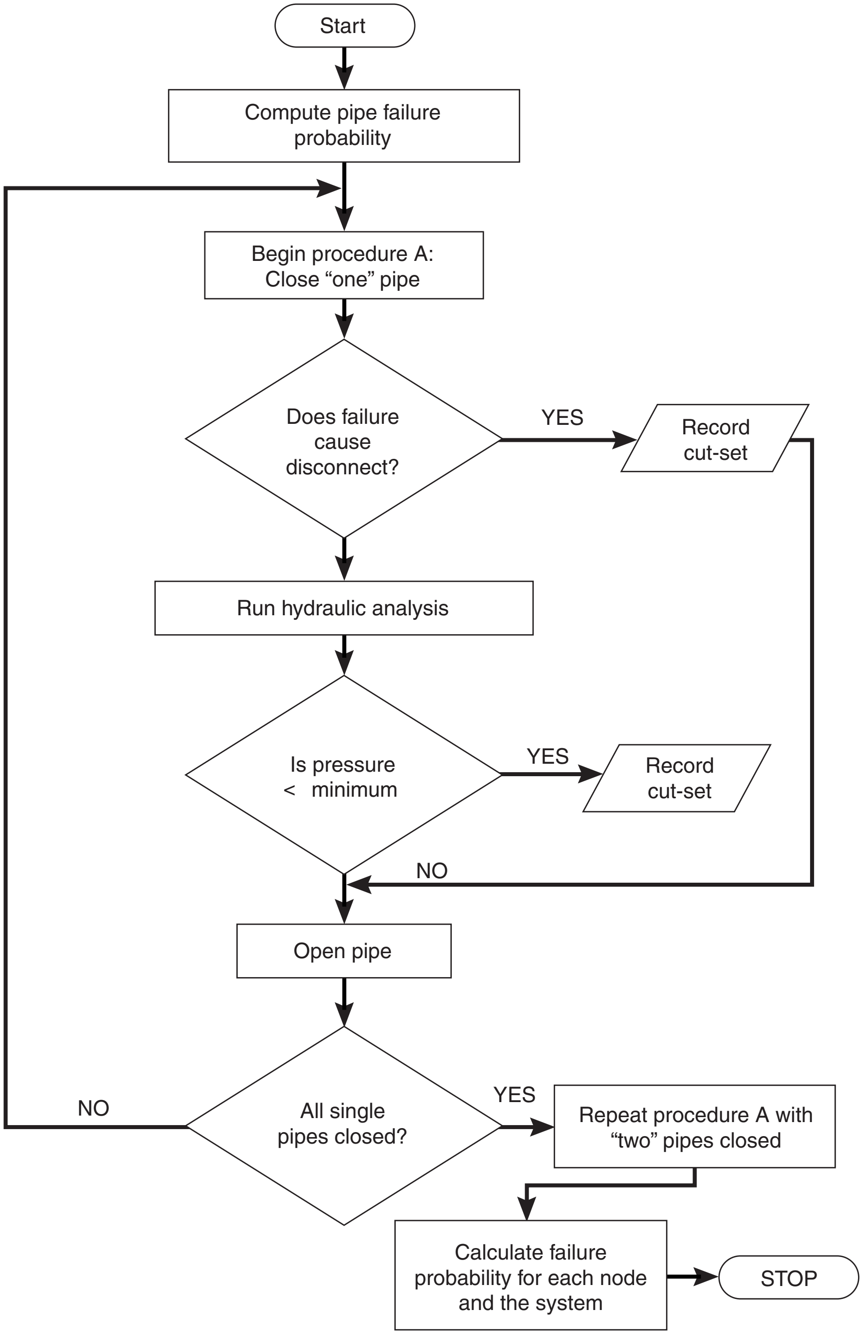


FIGURE 10.5 Minimum cut-set reliability flowchart with step hydraulic availability. (Shinstine et al., 2002)

$$R_s = 1 - P_s = 1 - \sum_{i=1}^{M} (1 - \mathrm{HA}_{\mathrm{net}}{}^{i})\, P(\mathrm{MC}_i) \tag{10.32}$$

where $\mathrm{HA}_{\mathrm{net}}{}^{i}$ is the network hydraulic availability (Fujiwara and de Silva, 1990):

$$\mathrm{HA}_{\mathrm{net}} = \prod_{j=1}^{j} = \mathrm{HA}_j \tag{10.33}$$

where HA_j is the hydraulic availability of node *j*. If HA equals 1, the failure probability for the cut-set is not included in Eq. (10.32); thus it is identical to Eq. (10.31) for the step-function hydraulic availability case. To compute the system reliability with continuous hydraulic availability, all cut-sets are included.

Mechanical Availability of Piping Components. The *probability of pipe failure* does not consider that pipes are repairable components. In this case, mechanical availability is a more appropriate measure. The *mechanical availability*, MA, of piping can be evaluated in terms of the availability of the individual pipes. For the *i*th component, when the mean time between failure (MTBF_i) and mean time to repair (MTTR_i) are known, the mechanical availability (MA_i) is (Ross, 1985; Cullinane, 1986)

$$\mathrm{MA}_i = \frac{\mathrm{MTBF}_i}{\mathrm{MTBF}_i + \mathrm{MTTR}_i} \tag{10.34}$$

where $\mathrm{MTBF}_i = 1/[\text{probability of failure}, P_i]$.

Mechanical unavailability (MU_i), the probability that a component is not operational, is then given by

$$\mathrm{MU}_i = 1 - \mathrm{MA}_i = 1 - \frac{\mathrm{MTBF}_i}{\mathrm{MTBF}_i + \mathrm{MTTR}_i} \tag{10.35}$$

For the entire water distribution network, the probability $\mathrm{MA}_{\mathrm{tot}}$ (total availability) that the system is fully operational in all its components (e.g., pipes) is given by the complement to the overall probability that at least one component will be removed from service. In brief, $\mathrm{MA}_{\mathrm{tot}}$ can be evaluated as the probability that all the components are in operating order:

$$\mathrm{MA}_{\mathrm{tot}} = \prod_{i=1}^{\mathrm{mc}} = \mathrm{MA}_i \tag{10.36}$$

where mc is the total number of components.

The probability of a failure of the *i*th component and all other components remaining operational is given by (Fujiwara and de Silva, 1990; Fujiwara and Tung, 1991)

$$u_i = \mathrm{MA}_{\mathrm{tot}} \frac{\mathrm{MU}_i}{\mathrm{MA}_i} \tag{10.37}$$

Similarly, the *probability of the event of simultaneous failure* of two (the *i*th and *k*th components) and all others remaining operational is

$$u_{ik} = \mathrm{MA}_{\mathrm{tot}} \frac{\mathrm{MU}_i}{\mathrm{MA}_i} \frac{\mathrm{MU}_k}{\mathrm{MA}_k} \tag{10.38}$$

Nodal and System Availability. Nodal and system availabilities are computed using a product form similar to equation (10.32). System availability is obtained from (Fujiwara and de Silva, 1990)

$$A_s = \mathrm{HA}_{\mathrm{net}}{}^0 \cdot \mathrm{MA}_{\mathrm{tot}} + \sum_{j=1}^{\mathrm{mc}} \mathrm{HA}_{\mathrm{net}}{}^i \cdot u_i + \sum_{i=1}^{\mathrm{mc}-1} \sum_{k=j-1}^{\mathrm{mc}} \mathrm{HA}_{\mathrm{net}}{}^{ik} \cdot u_{ik} + \cdots \tag{10.39}$$

where mc = total number of mechanical components taken into account
$\mathrm{HA}_{\mathrm{net}}{}^0$ = network hydraulic availability with a fully operational system
$\mathrm{HA}_{\mathrm{net}}{}^i$ = network hydraulic availability with only *i*th component removed
$\mathrm{HA}_{\mathrm{net}}{}^{ik}$ = network hydraulic availability with only *i*th and *k*th components removed

Each term in Eq. (10.39) is the probability of the network being in a certain condition times the hydraulic availability for that condition. Either definition for hydraulic availability can be used in Eq. (10.39). Nodal availabilities are computed with Eq. (10.39) using the HA for the node of concern. Note that compared to Eq. (10.32) the terms are summed so that the total approaches 1 while in Eq. (10.32) the failure conditions are subtracted from 1.

Only first- and second-order cut-sets are assumed to be important given the low failure probabilities, so Eq. (10.39) is truncated at this term. This reduces the number of computations without a significant loss of accuracy as shown in the application. Figure 10.6 illustrates the computational procedure with continuous hydraulic availability (HA) and mechanical availability (MA).

In summary, four measures for analyzing a system have been identified: system reliability [Eq. (10.32)] with step and continuous hydraulic availability and system availability [Eq. (10.39)] using step and continuous hydraulic availability. Section 10.3.3 discusses the values of these measures for two midsize municipal water distribution networks.

10.3.3 Application of Methodology

The Tucson, Arizona, water distribution system is divided into many water service areas (WSAs), with some areas isolated from the main system and some set up to serve specific zones. Water service area WSA-C6 (consisting of 109 pipes and 89 nodes), shown in Fig. 10.7, was chosen to illustrate the methodology for computing system reliability. The pipe failure rates listed in Table 10.2 were determined

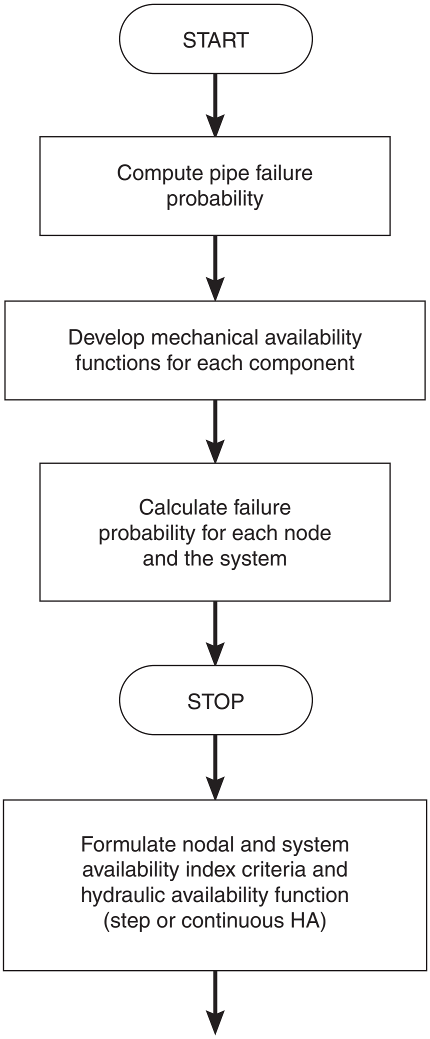


FIGURE 10.6 Steps in computing mechanical availability or hydraulic reliability with pipe failure given step or continuous hydraulic availability. (Shinstine et al., 2002)

from the available maintenance and repair data from Tucson Water. The likelihood of failure rate for a given pipe length and type was estimated from the available information on the total length of pipe by size and material. The two sources of data, maintenance database and pipe inventories, were used to provide failure rate per mile per year for each pipe size and material. Based on the maintenance records, the mean time to repair (MTTR) for all pipes was assumed to be 1 day.

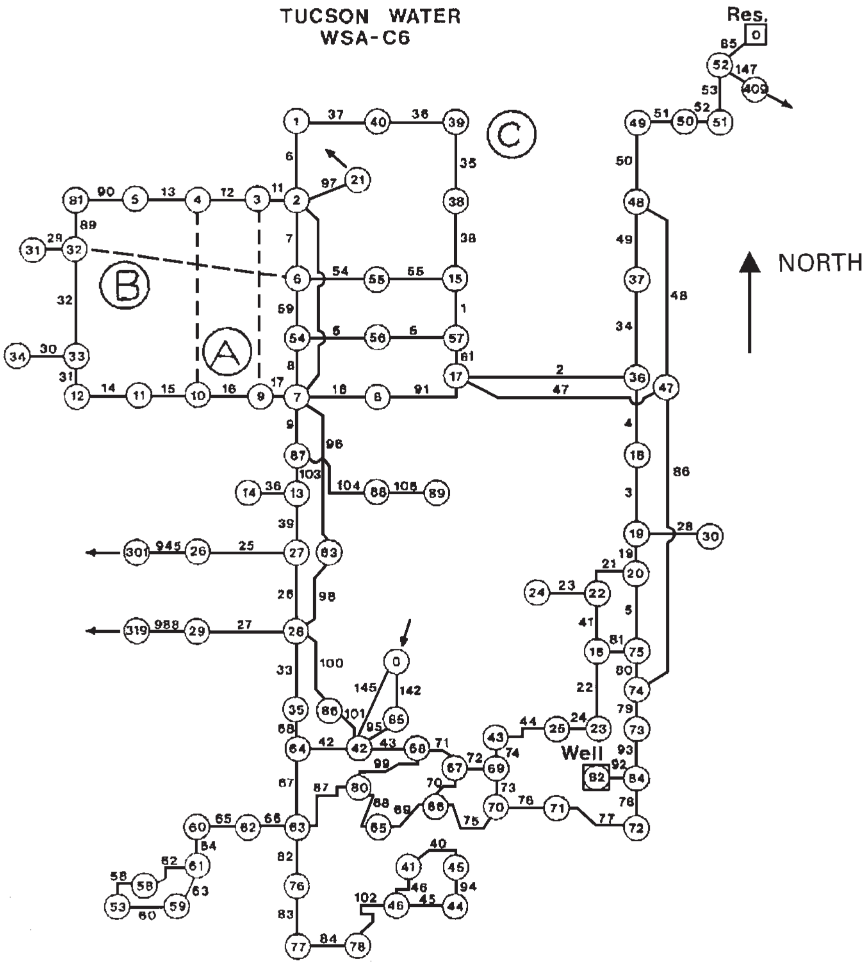


FIGURE 10.7 WSA-C6 water service area in Tucson, Arizona.

The *minimum operational pressure* for Tucson Water is 241.3 kN/m^2 (35 psi) for normal conditions and 137.9 kN/m^2 (20 psi) during fire flows. For an initial reliability analysis, a pressure of 6.9 kN/m^2 (1 psi) was used to identify pipe sets resulting in a geometric disconnection or mechanical failure of the system. Identifying mechanical failure provided a baseline reliability and simplified the process of checking the cut-sets resulting from hydraulic failures.

When using continuous hydraulic availability, the modeler must define the parameters for the function. Here, two different cases were analyzed. In both cases, the standard deviation was 17.25 kN/m^2 and $H_{\min}$ and $H_{\max}$ were two standard deviations from the mean pressure requirement. For case 1, mean nodal pres-

TABLE 10.2 Failure Rates and Cumulative Lengths of Pipes (Tucson and St. Louis)

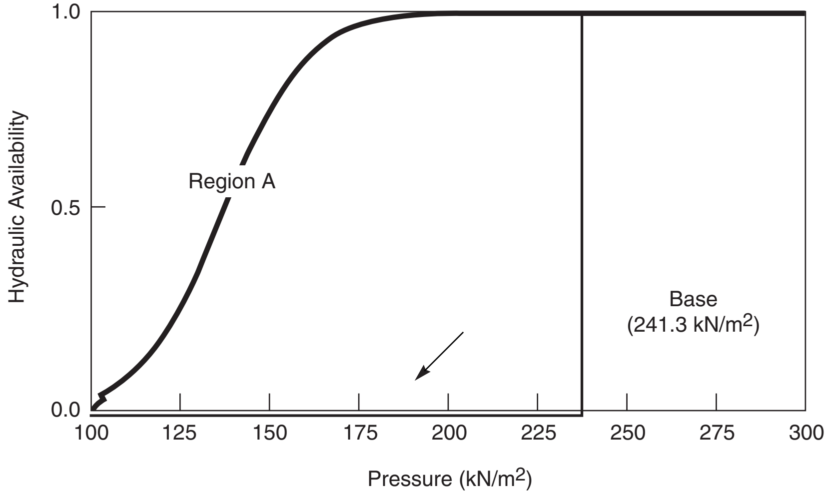
Pipe diameter, m	Failure rate, km/yr		Cum. pipe length, m (WSA-C6)
	Tucson	St. Louis	
0.051	0.022	0.387	402.34
0.102	0.022	0.387	882.70
0.152	0.009	0.225	12,173.10
0.203	0.007	0.171	9,048.30
0.254	0.011	0.130	
0.305	0.005	0.099	4,375.40
0.406	0.007	0.058	3,414.06
0.610	0.002	0.019	2,344.83
0.914	0.005	0.004	226.16

sure, minimum nodal pressure (H_{min}), and maximum nodal pressure (H_{max}) were defined as 137.9 kN/m^2, 103.4 kN/m^2, and 172.4 kN/m^2, respectively. The corresponding pressure values for case 2 were 241.3 kN/m^2, 206.8 kN/m^2, and 275.8 kN/m^2, respectively. Figure 10.8 shows the continuous hydraulic availability curves for the two pressure ranges, superimposed on the step function with minimum required pressure (PR) of 241.3 kN/m^2 (35 psi).

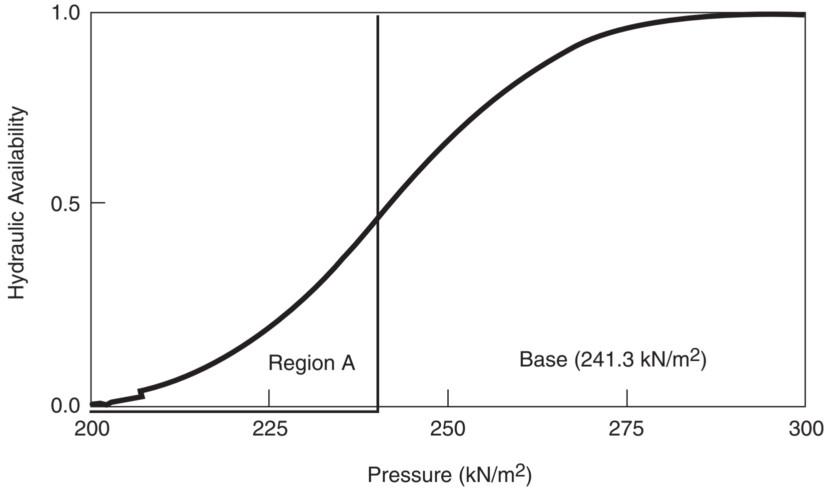
As described above, although the pressure heads were less than the minimum required pressure, i.e., the base condition pressure in region A (Fig. 10.8), nodes with pressures in that range were assigned to nonzero hydraulic availability. From the selected parameters and resulting shape of the curves, it is seen that the estimated reliability/availability (R/A) for case 1 will be larger than for case 2, since case 1 has a larger range of nonzero hydraulic availabilities.

Results. Tables 10.2 to 10.4 are used to show the results of applying the methodology to the WSA-C6 water service area in Tucson, Arizona. Four possible measures were analyzed for hydraulic reliability and availability in terms of the shape of hydraulic availability (HA) function (step or continuous):

1. Step hydraulic availability with mechanical reliability [Eq. (10.32) with Fig. 10.3]
2. Continuous hydraulic availability with mechanical reliability [Eq. (10.32) with Fig. 10.4]
3. Step hydraulic availability with mechanical availability [Eq. (10.39) with Fig. 10.3]
4. Continuous hydraulic availability with mechanical availability [Eq. (10.39) with Fig. 10.4]



(a)



(b)

FIGURE 10.8 (*a*) Continuous hydraulic availability function (case 1). (*b*) Continuous hydraulic availability function (case 2). (Shinstine et al., 2002)

Measures 1 and 2 are described as reliability measures, and measures 3 and 4 are availability terms. Results from the four measures are compared (Tables 10.3 and 10.4), and design proposals (Table 10.5) to improve R/A are discussed for a few cases.

To simulate pipe aging, the roughness coefficient was decreased by 10 percent. The reliability remained the same as the base condition. Since hydraulic failure

TABLE 10.3 WSA-C6 System Reliability Results with Step HA Function

Mechanical reliability (measure 1)				Mechanical availability (measure 3)			
Condition	Cut-sets	System reliability, %	ΔR_s	Condition	Cut-sets	System availability, %	ΔA_s
Pressure, kN/m^2				Pressure, kN/m^2			
6.9	243	96.39231	0.380	6.9	243	99.98991	0.001
137.9	243	96.39231	0.380	137.9	243	99.98991	0.001
241.3	293	96.01107	0	241.3	311	99.98922	0
310.3	944	75.15366	−20.860	310.3	944	99.93055	−0.059
Demand				Demand			
50%	301	96.38246	0.370	50%	391	99.99281	0.004
120%	411	96.00548	−0.010	120%	411	99.93190	−0.057
150%	412	96.00449	−0.010	150%	500	99.93055	−0.059
200%	576	96.00325	−0.010	200%	659	99.93055	−0.059
Roughness				Roughness			
−10%	309	96.00979	−0.001	−10%	309	99.93055	−0.059

Note: ΔR_s and ΔA_s = difference in reliability and availability, respectively, relative to base condition: 241.3 kN/m^2 (35 psi).

TABLE 10.4 System Reliability and Availability Using Step Hydraulic Availability Function

	First-order pipe sets only	First- and second-order pipe sets
Results using Tucson failure rates		
System reliability* (%)	96.15006	96.01107
System availability† (%)	99.97846	99.98922
Results using St. Louis failure rates on Tucson water systems		
System reliability* (%)	18.44431	0
System availability† (%)	99.54907	99.77650

Note: Base condition: 241.3 kN/m^2 (35 psi).
*Using mechanical reliability (measure 1).
†Using mechanical availability (measure 3).

TABLE 10.5 WSA-C6 Design Alternatives

Design alternative	Cut-sets	Reliability, % (measure 1)*	ΔR_{s1}	Reliability, % (measure 2)†	ΔR_{s2}	Availability, % (measure 3)*	ΔA_{s3}
Base	293	96.01107	0	96.39186	0	99.98922	0
A (add pipes between nodes 9-3 & 10-4)	185	96.04030	0.03	96.42105	0.029	99.98922	0
B (add pipe between nodes 6-32)	193	96.03641	0.02	96.41730	0.025	99.9922	0
C (add booster pump adjacent to node 39)	267	96.37476	0.36	96.48752	0.096	99.99027	0.001
A & C (add pipes in the northwest zone & booster pump in north zone)	167	96.39935	0.39	96.49693	0.105	99.99027	0.001

Note: FGN = fixed grade node. ΔR_s and ΔA_s = difference in reliability and availability relative to base condition.

*Base condition: 241.3 kN/m^2

†Base condition: mean = 137.9 kN/m^2, H_{min} = 103.4 kN/m^2, H_{max} = 172.4 kN/m^2.

did not occur until a large demand change was imposed, the lower C values did not cause higher head losses that could alter the system R/A.

All results were compared to the base conditions of 241.3 kN/m^2 (35 psi) and average (normal) demand, with a system reliability for the base conditions of 96.01 percent (Table 10.5). The reliability for the system at the operational pressure of 241.3 kN/m^2 (35 psi) included conditions resulting in hydraulic and mechanical failure. Noteworthy, at node 37 in WSA-C6, due to its elevation, the pressure runs slightly lower (227.9 kN/m^2 or 33.1 psi) than the operational minimum of 241.3 kN/m^2 (35 psi). Therefore, the minimum pressure at this node was set at 206.8 kN/m^2 (30 psi). At 137.9 kN/m^2 (20 psi), only mechanical failure was observed. In other words, only cut-sets that geometrically disconnect the system influence the reliability at lower pressures. However, only a 0.4 percent decrease in reliability occurred when considering hydraulic failure at the 241.3-kN/m^2 (35-psi) requirement.

Pressure requirement, PR, changes had a more significant impact on R/A. The reliability dropped dramatically when the pressure requirement was increased to 310.3 kN/m^2 (45 psi) (Table 10.3). Many nodes with pressures between 241 and 310 kN/m^2 (35 and 45 psi) are considered as failing when the PR is increased.

Perhaps the most interesting result from this work is the relatively high values of reliability and, even more so, availability. In all cases, availability was greater than 99.9 percent of the time (364.8 of 365 days). When failure does occur, it may be low pressure or limited periods without service for only a portion of the network. Recall that the mean time to repair after identification of a failure was 1 day. Examination of nodal reliabilities is useful to determine specific system weaknesses.

As expected, the system availability is higher in all cases compared to the system reliability. The probability failure is higher than the probability of being in a failed condition since the pipes are likely repaired very quickly. In addition, the continuous hydraulic availability relationship results in higher reliability and availabilities compared to the step-function relationship. This result is also expected as described above regarding the hydraulic availability functions.

Equations (10.32) and (10.39) have been truncated at the second-order cut-set term. To see the effect of considering different orders of cut-sets, the system R/A for first- and second-order pipe sets for the step hydraulic availability function are listed in Table 10.4. The reliabilities are essentially the same with and without the second-order pipe sets. This clearly shows that the probability of two pipes failing simultaneously is very small and justifies the assumption that considering at most higher-order failures is unnecessary. However, it is notable and expected that the system reliability decreases with the inclusion of second-order pipe sets. The probability of failure conditions is subtracted from 1 in the reliability relationship [Eq. (10.32)] and more failures occur when second-order cut-sets are included. Availability, on the other hand, sums the conditions that provide acceptable operations, so they increase very slightly in WSA-C6 when second-order sets are included. Because of significant differences in results between first- and second-order pipe sets, consideration of second-order sets is important for reliability measures with high failure rates.

Because of mechanical failure and disconnections on branch portions of the networks with the higher failure rates, the reliability was significantly impacted (to less than 10 percent for both systems). Since it properly considers pipe repair, the availability fell only slightly. Under those conditions, the general network layout would probably not be modified significantly if the network was located in St. Louis rather than Tucson.

System Reliability and Design Changes. The availability of this network is already quite high. Parallel pipes have apparently been added for redundancy and provide additional reliability (e.g., pipes 48 and 86). However, a number of smaller areas are supplied through a single pipe (e.g., southwest downstream of node 63). Examining nodal R/A values, the north zone (nodes 39 and 40) is most affected by hydraulic failure and the northwestern zone is affected by a combination of mechanical and hydraulic failures. Thus, unlike WSA-W1, the weaknesses are dispersed throughout the system. Several alternatives were examined to improve matters on the north side. Referring to Table 10.5, no alternative provided significant improvement with the booster pump ensuring pressure in the north region giving the maximum return. Alternatives A, C, and A & C suggest that the changes are independent of their influence on the system.

10.4 OTHER APPROACHES TO ASSESSMENT OF RELIABILITY

Consider in the instance, a reliability measure based on deficits in volumes of water supplied. In recognition of (1) the need to incorporate both demand failure and component failure (2) the fact that demands are not actually located at nodes as is normally assumed in network analysis and design, and (3) the probability nature of the events that give rise to failure, Bouchart and Goulter (1991, 1992) formulated the following measures for *volume deficits*.

1. Volume deficit arising from "demand variation failure"

$$\mathrm{ED}_d^{\eta} = K \int_{Q_{de}^{\eta}}^{\infty} (Q^{\eta} - Q_{de}^{\eta})\, p(Q^{\eta})\, dQ^{\eta} \tag{10.40}$$

where ED_d^{η} = expected volume of deficit arising from demand at node η being larger than design value Q_{de}^{η} at that node

Q^{η} = demand per unit time at node η

K = factor to convert expected flow rates to expected volumes

Q_{de}^{η} = design demand at node η

$p(Q^{\eta})$ = probability of demand Q_{η} at node η

2. Volume deficit arising from "component (pipe) failure"

$$\mathrm{ED}_f^{s} = q_d t_f E(\mathrm{NF}) \tag{10.41}$$

where ED_f^{s} = expected volume of deficit arising from a section of pipe in links failing and having to be isolated for repair

q_d = average demand per unit time in section of pipe isolated

t_f = average duration of repair

$E(\mathrm{NF})$ = average number of failures per unit time of pipe section (equal to length of pipe section between valves multiplied by number of breaks per unit length of pipe)

The expected deficit for the volume deficits due to component (pipe) failure and demand failure now have the same units. Hence, they can be added directly to give the total expected volume of deficit over the network, i.e.,

$$\text{TND} = \sum_{\text{all links}_s} \text{ED}_{f^s} + \sum_{\text{all links}\eta} \text{ED}_d^{\eta} \tag{10.42}$$

where TND is the total expected volume of deficit due to both component failure and demand variation failure.

It should be recognized that the statement of reliability in Eq. (10.42) as used by Bouchart and Goulter (1991, 1992) does not address the impacts on the rest of the network caused by the reduction in hydraulic of the network associated with removal of one of the components. For the same reason, it examines the impacts of pump or valve failure (valve failure requires the isolation of the section of the network containing the valve so that the valve can be repaired and has some of the characteristics as pipe failure). The important aspect of the approach is that it recognizes both demand variation failure and component failure in a probabilistic sense. However, the technique is only concerned with purely volumetric deficit and does not take into consideration the ability of a network to meet flow demands at reduced pressures.

An approach for measurement of reliability which does take into account the pressure of delivery formulated and applied by Cullinane (1986) and Cullinane et al. (1992) is based on the concept of availability, which is defined as the percentage of time the pressures in the network are below some preset level. In the case of a node, availability represents the percentage of time the pressure at the node is less then some predetermined level, i.e.,

$$A_j = \sum_{i=1}^{M} \frac{A_i t_i}{T} \tag{10.43}$$

where A_j = nodal availability at node j
A_i = availability during time period i
t_i = length of time period i
M = number of time period i
T = time of simulation = time horizon for analysis of reliability

The network can be defined as

$$A = \sum_{j=1}^{N} \frac{A_j}{N} \tag{10.44}$$

where A is the network availability and N is the total number of nodes. This measure is capable of considering both component (mechanical) and demand variation failure. The technique can also be extended to include the probabilistic aspects of mechanical and demand variation failure though the following expression:

$$\text{AE}_{jl} = \text{R}_{jl} \cdot \text{A}_{jl} + Q_{jl} \cdot A_j \tag{10.45}$$

where AE_{jl} = expected value of availability at node j, considering component l
R_{jl} = probability that component l is operational
Q_{jl} = probability that component l is not operational
A_{jl} = availability of node j with component l operational

A_j = availability of node j with component l nonoperational

An interesting feature of the discussion by Cullinane et al. (1992) of this measure was their observation that "the judgement and experience of the modeler-designer is critical in system evaluation." The importance and relevance of this observation is discussed later in relation to simulation models for assessment of network reliability.

It should be noted that demand variation failure is incorporated in Eq. (10.45) through the R_{jl} term in which failure to meet the predetermined pressure can only occur through a condition arising from demand failure. On the other hand, both component failure and demand failure are recognized in the $Q_j \cdot A_j$ term; component failure specifically through the Q_{jl} term; with the demand failure condition imposed through the A_j term. However, Eq. (10.45) does not explicitly consider the probabilistic aspects of demand failure. Consideration of the probabilistic aspects may be obtained by modifying:

$$\mathrm{AE}_{jl}^{T} = \int_0^\infty (R_{jl} A_{jl} + Q_{jl} A_j)\, p(Q^\eta)\, dQ^\eta \qquad (10.46)$$

where AE_{jl}^{T} is the expected availability at node j considering demand variation and mechanical failure and all other terms are as described previously.

Although Eq. (10.46) recognizes deficit in pressure as a factor to be addressed in defining network failure, the measure is not able to distinguish between failure conditions involving a large number of failures, each of which results in relatively small decreases in pressure below the minimum acceptable, and failure conditions involving a small number of failures, each of which causes a significant decrease in pressure below the minimum acceptable. For example, a pipe failure that results in a reduction in delivery pressure heads at a node from 30 to 10 m is considerably more serious than a pipe failure of the same duration that results in reduction in delivery pressure head at the same node from 30 to 29 m. Equation (10.46) cannot distinguish between these failures.

One approach to the problem is to use a measure of the following general form:

$$\mathrm{ND} = \sum_{i=1}^{N} q^j H_{\min} - \sum_{i=1}^{N} q_a^{\,j} H_{aj} \qquad (10.47)$$

$$= Q_T * H_{\min} - \sum_{j=1}^{N} q_a^{\,j} H_{aj} \qquad (10.48)$$

where q^i = demand at node j

h = total number of nodes

Q_T^* = total network demand

$q_a^{\,j}$ = actual flows supplied at node j

H_{aj} = actual delivery pressure at node j if $H_{aj} > H_{\min}$ set $H_{aj} = H_{\min}$

ND = network deficit

Note that this measure does not consider the stochastic nature of the demands and applies only for known deterministic nodal demands. The primary strength of this measure lies, however, in its ability to recognize explicitly the relationship between flows and pressures. If all flow demands are met, then any deficit is solely due to deficits in supply pressure. If all demand pressures are satisfactory, then any deficit is due to the supply not meeting the flow demand. In both cases, the size of the deficit increases as the amount by which the flow or pressure requirements are not met increases. However, if both flow and pressure requirements are not met simultaneously, i.e., flows less than flow demands are supplied and these flows are only able to be supplied at pressures below minimum acceptable, then the products of the q and H terms in Eq. (10.48) will be even less than if only one aspect of demand was not met. In this sense the expression is very comprehensive in its interpretation of reliability.

However, the units of the measure are volume-pressure, i.e., (m^3/h)(m) and as such do not directly represent reliability as engineers tend to know or assess it. For this reason the expression should be considered as a heuristic measure of reliability wherein it is known that if the deficit determined by the measure decreases, the reliability of the system is improving. Additionally, this expression does not consider the duration or probabilities of the events that give rise to pressure or supplied flow rates being less than the requirements. These failures can be incorporated in the same fashion as they were for the expected availability measure of Eq. (10.46) by replacing the Q_j and A_{jl} terms in that equation by elements from the right side of Eq. (10.48) disaggregated by node, i.e.,

$$\mathrm{ED}_{jl} = R_{jl}[q^j H_{\min}] + Q_{jl}[q_a{}^i H_{aj}] \tag{10.49}$$

with the *total network deficit* given by

$$\mathrm{TND}' = \sum_{j=1}^{n} \mathrm{ED}_{jl} \tag{10.50}$$

where ED_{jl} is the expected deficit at node j measured by the product of flow and pressure deficits and TMD is the total expected network deficit as measured by the product of flow and pressure deficits and without recognition of the stochastic nature of the demands.

Incorporation of demand variation can be achieved for each node by the following expression:

$$\mathrm{ED}_{jl}^{T} = \int_0^{\infty} \mathrm{ED}_{jl} p(Q^n)\, dQ^n \tag{10.51}$$

Even when these probabilistic aspects have been incorporated in this manner, the measure only gives the average condition; it does not provide any indication of the extent to which the deficit varies around this mean; i.e., again it does not indicate whether the mean is caused by a small number of very large and possibly unacceptable deficits or a large number of acceptable deficits.

This issue aside, the above discussion might indicate that the problem of assessing and measuring reliability in water distribution networks has been essentially solved. However, this is clearly not the case. It is generally accepted in the literature, e.g., Cullinane et al. (1992), Goulter (1992, 1995a) that there is currently no method or measure for assessment of reliability that is both comprehensive in its interpretation of reliability and computationally practical. Goulter et al. (2000) discuss the simulation and analytical models presently available for assessment of reliability in water distribution networks and highlight the strengths and weaknesses of the approaches.

10.5 RELIABILITY-BASED DESIGN OPTIMIZATION MODELS

10.5.1 Framework for Reliability-based Design Model

Su et al. (1987) developed a basic framework for a model that can be used to determine the optimal (least-cost) design of a water distribution system subject to the various design constraints and also subject to reliability constraints. The optimization model for determining pipe diameters can be stated as follows:

Minimize cost subject to

- Conservation of flow and conservation of energy constraints
- Pressure head bounds to satisfy pressures throughout the system
- Reliability constraints [Eq. (10.21)] for both system and demand nodes

The solution procedure used a nonlinear optimization model interfaced with the KYPIPE simulation model so that the hydraulic constraints (conservation of flow and conservation of energy) were actually solved by the simulator, at each iteration of the optimization model, and pressure requirements were guaranteed through a penalty function method for the optimization procedure.

10.5.2 Reliability-based Optimization Model for Operation Considering Uncertainties of Water Quality

Consider a water distribution system with M pipes; N total nodes, of which there are K junction nodes and S storage nodes (tank or reservoir); and P pumps which are operated for time period T. The basic optimization model for water distribution system operation can be stated in general form as

$$\text{Minimize } z = \text{minimize } k \sum_{p,t} \frac{Q_{pt}\text{TDH}_{pt}\text{UC}_t\text{D}_{pt}}{\text{EFF}_{pt}} \qquad (10.52)$$

subject to

$$\sum_i (q_{ik})_t - \sum_j (q_{kj})_t - Q_{kt} = 0 \qquad \forall i, j, \in M; k = 1, \ldots, K; t = 1, \ldots, T \tag{10.53}$$

$$h_{it} - h_{jt} = f(q_{ij})_t \qquad \forall i,j \in M; t = 1, \ldots, T \tag{10.54}$$

$$y_{st} = y_{s(t-1)} + \frac{q_{s(t-1)}}{A_s} \Delta t \qquad s = 1, \ldots, S; t = 1, \ldots, T \tag{10.55}$$

$$\frac{\delta(C_{ij})_t}{\delta t} = -\frac{(q_{ij})_t}{A_{ij}} \frac{\delta(C_{ij})_t}{\delta x_{ij}} + \Theta(C_{ij})_t \qquad \forall i, j \in M; t = 1, \ldots, T \tag{10.56}$$

$$\underline{H}_{kt} \le H_{kt} \le \overline{H}_{kt} \qquad k = 1, \ldots, K; t = 1, \ldots, T \tag{10.57}$$

$$\underline{y}_{st} \le y_{st} \le \overline{y}_{st} \qquad s = 1, \ldots, S; t = 1, \ldots, T \tag{10.58}$$

$$\underline{C}_{nt} \le C_{nt} \le \overline{C}_{nt} \qquad n = 1, \ldots, N; t = 1, \ldots, T \tag{10.59}$$

The objective function [Eq. (10.52)] minimizes the total energy cost as a function of the power rate per kilowatt hour of pumping during time period $t(\mathrm{UC}_t)$, flow from pump p during time period $t(Q_{pt})$, operating point total dynamic head for pump p during time period $t(\mathrm{TDH}_{pt})$, efficiency of pump p in time period t (EFF_{pt}), and the state (on or off) of pump p during time period $t(D_{pt})$, wherein the pump is either on ($D_{pt} = \Delta t$) or off ($D_{pt} = 0$) during the period. Constraint equations (10.53) and (10.54) are the hydraulic constraints: conservation of mass and conservation of energy, respectively. In these equations, $(q_{ij})_t$ is the flow in pipe m connecting nodes i and j at time t, Q_{kt} is the flow consumed (+) or supplied (−) at node k at time t, h_{it} is the hydraulic grade line elevation at node i at time t, $f(q_{ij})_t$ is the functional relation between head loss and flow in a pipe connecting nodes i and j at time t.

Constraint equation (10.55) states that the height of water stored at a storage node for the current time period, y_{st} is a function of the height of water stored from the previous time period and the flow entering or leaving the tank during the period Δt. In this equation A_s is the cross-sectional area of a tank and q_s is the flow entering or leaving the tank during period $t - 1$. The water quality constraint equation (10.56) is the conservation of mass of the chlorine in each pipe m connecting nodes i and j in the set of all pipes M. In this equation $(C_{ij})_t$ is the concentration of chlorine in pipe m connecting nodes i and j as a function of distance and time, x_{ij} is the distance along the pipe connecting nodes i and j, A_{ij} is the cross-sectional area of the pipe connecting nodes i and j, $\Theta(C_{ij})_t$ is the rate of reaction of the chemical within the pipe connecting nodes i and j at time t. Equations (10.57) to (10.59) define the lower and upper bounds on the pressure heads at node k at time t, the lower and upper bounds on the depth of water stored at storage node s at time t, and the lower and upper bounds on chlorine concentration at node n at time t, respectively.

Since the minimum and maximum residual chlorine concentrations are uncertain, they are chosen as independent random variables from the viewpoint of operation. Equation (10.59) is replaced by a probabilistic statement known as a *chance constraint.* The chance-constrained formulation of the reduced problem is to optimize the objective function, Eq. (10.52), subject to Eqs. (10.57) and (10.58) and the following chance constraint:

$$P[\underline{C}_{nt} \leq C_{nt} \leq \overline{C}_{nt}] \geq \alpha_{Ct} \qquad n = 1, \ldots, N; t = 1, \ldots, T \tag{10.60}$$

The storage bound constraint [Eq. (10.58)] has to be considered if the desired minimum and maximum limits of the tank water level storage heights are different than the allowable minimum and maximum tank water elevations which are due to physical limitations of the storage tanks. Otherwise, this constraint is automatically satisfied within the simulation procedure by keeping the storage heights within the given allowable values of the minimum and maximum tank water elevations. Constraint equation (10.60) is expressed as the probability $P(\ )$ that the residual chlorine concentration at each demand node during the time t is between lower and upper bounds with a probability level of α_{Ct}. In general the value of the constraint performance reliability α_{Ct} can be specified and manipulated to consider the effect of uncertainty. The above model [Eqs. (10.52), (10.56), (10.58), and (10.60)] can be transformed into a deterministic form using the concept of the cumulative probability distribution (Mays and Tung, 1992).

The chlorine residual data for water distribution systems are minimal or nonexistent, and it is difficult to assign the appropriate probability distribution for the residual chlorine concentration. The minimum and maximum chlorine residual concentration variables were assumed to follow either the normal or lognormal probability distribution. The deterministic form of the chance constraint expressed by Eq. (10.60) is

$$\mu_{\underline{C}_{nt}} + \sigma_{\underline{C}_{nt}} \cdot z_{(1+\alpha_{Ct})/2} \leq C_{nt} \leq \mu_{\overline{C}_{nt}} + \sigma_{\overline{C}_{nt}} z_{[1-(1+\alpha_{Ct})/2]} \tag{10.61}$$

for the normal distribution or

$$\exp\{\mu_{\ln \underline{C}_{nt}} + \sigma_{\ln \underline{C}_{nt}} z_{(1+\alpha_{Ct})/2}\} \leq C_{nt} \leq \exp\{\mu_{\ln \overline{C}_{nt}} + \sigma_{\ln \overline{C}_{nt}} z_{[1-(1+\alpha_{Ct})/2]}\} \tag{10.62}$$

for the lognormal distribution.

The reduced deterministic formulation of the chance-constrained model is expressed by the objective function and constraints, Eqs. (10.52), (10.57), (10.58), and (10.61) or (10.62). The model is a programming problem in which Eq. (10.61) or (10.62) is treated as a simple bound constraint because the right-hand side and the left-hand side are known; in fact, $\mu_{\overline{C}_{nt}}$, $\sigma_{\overline{C}_{nt}}$, $\mu_{\underline{C}_{nt}}$, $\sigma_{\underline{C}_{nt}}$, and α_{Ct} are all specified. The simulated annealing code, by Goldman (1998) and described in Chap. 12, can be used to solve the deterministic form of the chance-constrained model [Eqs. (10.52), (10.57), (10.58), and (10.61) or (10.62)]. This model links the EPANET simulation model with the method of simulated annealing to optimize the operation of water distribution systems for hydraulic and water quality behavior.

10.6 REFERENCES

Awumah, K., and I. Goulter, "Maximizing Entropy-defined Reliability of Water Distribution Networks," *Engineering Optimization,* 20(11):57–80, 1992.

Awumah, K., I. Goulter, and S. Bhatt, "Entropy-based Redundancy Measures in Water Distribution Network Design," *Journal of Hydraulic Engineering,* ASCE, 117(3):595–614, 1991.

Bao, Y. X., and L. W. Mays, "Model for Water Distribution System Reliability," *Journal of Hydraulic Engineering,* ASCE, 116(9):1119–1137, 1990.

Bao, Y. X., and L. W. Mays, "New Methodology for Optimization of Freshwater Inflows to Estuaries," *Journal of Water Resources Planning and Management,* ASCE, 120(2):199–217, 1994a.

Bao, Y. X., and L. W. Mays, "Optimization of Freshwater Inflows to the Lavaca-Tres Palacios, Texas, Estuary," *Journal of Water Resources Planning and Management,* ASCE, 120(2):218–236, 1994b.

Biem, G., and G. Hobbs, "Analytical Simulation of Water System Capacity Reliability. 2: A Markov-Chain Approach and Verification of Models," *Water Resources Research,* 24(9):1445–1458, 1988.

Billinton, R., and R. N. Allan, *Reliability Evaluation of Engineering Systems: Concepts and Techniques.* Pitman Books Limited, London, 1983.

Boccelli, D. L., M. E. Tryby, J. G. Uber, L. A. Rossman, M. L. Zierolf, and M. Polycarpou, "Optimal Scheduling of Booster Disinfection in Water Distribution Systems," *Journal of Water Resources Planning and Management,* ASCE, 124(2):99–11, 1998.

Bouchart, F., and I. C. Goulter, "Implications of Pipe Failure on the Hydraulic Performance of Looped Water Distribution Networks," *Proceedings of the First Caribbean Conference on Fluid Dynamics,* St. Augustine, Trinidad, 1989, pp. 291–297.

Bouchart, F., and I. Goulter, "Reliability Improvements in Design of Water Distribution Networks Recognizing Valve Location," *Water Resources Research,* 27(12):3029–3040, 1991.

Bouchart, F., and I. Goulter, "Selecting Valve Location to Optimize Water Distribution Network Reliability," *Proceedings of the 6th IAHR International Symposium on Stochastic Hydraulics,* J.-T. Kuo and G.-F. Lin (eds.), May 18–20, 1992, Taipei, Taiwan, 1992, pp. 155–162.

Cullinane, M. J., "Reliability Evaluation of Water Distribution System Components," *Proceedings of the Specialty Conference on Hydraulics and Hydrology in the Small Computer Age,* ASCE, New York, 1985, pp. 353–358.

Cullinane, M. J., "Hydraulic Reliability of Urban Water Distribution Systems," in M. Karamoyz et al. (eds.), *Water Forum 1986: World Water Issues in Evolution,* ASCE, 1986, pp. 1264–1271.

Cullinane, M. J., K. E. Lansey, and C. Basnet, "Water Distribution System Design Considering Component Failure During Static Conditions," *Proceedings of the National Conference on Hydraulic Engineering,* ASCE, New York, 1989, pp. 762–767.

Cullinane, M., K. Lansey, and L. Mays, "Optimization-Availability Based Design of Water-Distribution Networks." *Journal of Hydraulic Engineering,* ASCE 118(3):420–441, 1992.

Duan, N., and L. Mays, "Reliability Analysis of Pumping Systems." *Journal of Hydraulic Engineering,* ASCE, 116(2):230–248, 1990.

Duan, N., L. Mays, and K. Lansey, "Optimal-based Design of Pumping and Distribution Systems," *Journal of Hydraulic Engineering,* ASCE, 116(2):249–268, 1990.

Fletcher, R., "An Ideal Penalty Function for Constrained Optimization," *Journal of Institute of Mathematics and Its Applications,* 15:319–342, 1975.

Fujiwara, O., and A. de Silva, "Algorithm for Reliability Based Optimal Design of Water Networks," *Journal of Environmental Engineering,* ASCE, 116(3):575–587, 1990.

Fujiwara, O., and T. Ganesharajah, "Reliability Assessment of Water Supply Systems with Storage and Distribution Networks, *Water Resources Research,* 29(8):2917–2924, 1993.

Fujiwara, O., and J. Li, "Reliability Analysis of Water Distribution Networks in Consideration of Equity, Redistribution, and Pressure Dependent Demand," *Water Resources Research,* 34(7):1843–1850, 1998.

Fujiwara, O., and H. Tung, "Reliability Improvement for Water Distribution Networks through Increasing Pipe Size," *Water Resources Research,* 27(7):1395–1402, 1991.

Germanopoulis, G., P. Jowitt, and J. Lumbers, "Assessing the Reliability of Supply and Level of Service for Water Distribution Systems," *Proceedings of the ICE,* Part 1, 80 (Apr.):413–428, 1986.

Goldman, F. E., "The Application of Simulated Annealing for Optimal Operation of Water Distribution Systems," Ph.D. dissertation, Department of Civil and Environmental Engineering, Arizona State University, Tempe, Ariz., 1998.

Goulter, I. C., "Measures of Internal Redundancy in Water Distribution Network Layouts," *Journal of Information and Optimization Science,* 9(3):363–390, 1988.

Goulter, I. C., "Systems Analysis in Water Distribution Network Design: from Theory to Practice." *Journal of Water Resources Planning and Management,* ASCE, 118(3):238–248, 1992a.

Goulter. I. C., "Modern Concepts of a Water Distribution System: Policies for Improvement of Networks with Shortcomings," in E. Cabrera, and F. Martinez (eds.), *Water Supply Systems: State of the Art and Future Trends,* Computational Mechanics Publications, Southampton, U.K., 1992b, pp. 119–138.

Goulter, I. C., "Analytical and Simulation Models for Reliability Analysis in Water Distribution Systems," in E. Cabrera and A. Vela (eds.), *Improving Efficiency and Reliability in Water Distribution Systems,* Kluwer Academic Publishers, Dordrecht, Netherlands, 1995a, pp. 235–266.

Goulter, I. C., "Measures of Internal Redundancy in Water Distribution Network Layouts," *Journal of Information and Optimization Science,* 9(3):363–390, 1995b.

Goulter, I., and F. Bouchart, "Reliability-Constrained Pipe Network Model," *Journal of Hydraulic Engineering,* ASCE, 116(2):211–229, 1990.

Goulter, I., and A. Coals, "Quantitative Approaches to Reliability Assessment in Pipe Networks," *Journal of Transportation Engineering,* ASCE, 112(3):287–301, 1986.

Goulter, I., and P. Jacobs, "Discussion of 'Water Distribution' Reliability: Analytical Method by Wagner, Shamir and Marks," *Journal of Water Resource Planning and Management,* ASCE, 115(5):709–711, 1989.

Goulter, I., T. Walski, L. W. Mays, A. B. Sakarya, F. Bouchart, and Y. K. Tung, "Reliability Analysis for Design," Chap. 18 in L. W. Mays (ed.), *Water Distribution Systems Handbook,* McGraw-Hill, New York, 2000.

Gupta, R., and P. Bhave, "Reliability Analysis of Water Distribution Systems," *Journal of Hydraulic Engineering,* ASCE, 120(2):447–460, 1994.

Henley, E. J., and H. Kumamoto, *Reliability Engineering and Risk Assessment,* Prentice-Hall, Englewood Cliffs, N.J., 1981.

Hobbs, B., and G. Biem, "Analytical Simulation of Water System Capacity Reliability: Modified Frequency-Duration Analysis," *Water Resources Research,* 24(9):1431–1444, 1988.

Jacobs, P., "A Removal Set Based Approach to Urban Water Supply Distribution Network Reliability," Ph.D. thesis, University of Manitoba, Winnipeg, 1992.

Jacobs, P., and I. Goulter, "Evaluation of Methods for Decomposition of Water Distribution Networks for Reliability Analysis," *Civil Engineering Systems,* 5(2):58–64, 1988.

Jacobs, P., and I. Goulter, "Optimisation of Redundancy in Water Distribution Networks Using Graph Theoretic Principles," *Engineering Optimization* 15(1):71–82, 1989.

Jacobs, P., and I. Goulter, "Estimation of Maximum Cut-set Size for Water Network Failure," *Journal of Water Resources Planning and Management,* ASCE, 117(5):588–605, 1991.

Jowitt, P., and C. Xu, "Predicting Pipe Failure Effects in Water Distribution Networks." *Journal of Water Resources Planning and Management,* ASCE, 119(1):18–31, 1993.

Kessler, A., L. Ormsbee, and U. Shamir, "A Methodology for Least-cost Design of Invulnerable Water Distribution Networks," *Civil Engineering Systems,* 1(1):20–28, 1990.

Kettler, A. J., and I. C. Goulter, "Reliability Consideration in the Least-Cost Design of Looped Water Distribution Networks," *Proceedings of the 12th International Conference on Urban Hydrology, Hydraulics, and Sediment Control,* University of Kentucky, Lexington, 1983, pp. 305–312.

Kettler, A. J., and I. C. Goulter, "Analysis of Pipe Breakage in Urban Water Distribution Networks," *Canadian Journal of Civil Engineering,* 12(2):286–293, 1985.

Lansey, K., K. Awumah, Q. Zhong, and I. Goulter, "A Supply Based Reliability Measure for Water Distribution System," *Proceedings of the 6th IAHR International Symposium on Stochastic Hydraulic,* J-T. Kuo and G-F Lin (eds.), May 18–20, 1992, Taipei, Taiwan, 1992, pp. 171–178.

Lansey, K., N. Duan, L. Mays, and Y-K Tung, "Water Distribution Design Under Uncertainties," *Journal of Water Resources Planning and Management,* ASCE, 115(5):630–645, 1989.

Mays, L. W. (ed.), *Reliability Analysis of Water Distribution Systems,* American Society of Civil Engineers, New York, 1989.

Mays, L. W. (ed.), *Water Distributions Systems Handbook,* McGraw-Hill, New York, 2000.

Mays, L. W., and Y. K. Tung, *Hydrosystems Engineering and Management,* McGraw-Hill, New York, 1992.

O'Day, K. D., R. Weiss, S. Chiaveri, and D. Blair, *Water Main Evaluation for Rehabilitation/Replacement,* prepared for American Water Works Association Research Foundation, Denver, and USEPA, Cincinnati, 1986.

Ormsbee, L., and A. Kessler, "Optimal Upgrading of Hydraulic-Network Reliability," *Journal of Water Resources Planning and Management,* ASCE, 116(6):784–802, 1990.

Park, H., and J. Liebman, "Redundancy-Constrained Minimum Design of Water-Distribution Nets," *Journal of Water Resources Planning and Management,* ASCE, 119(1):83–98, 1993.

Quimpo, R. J., and U. Shamsi, "Reliability-based Distribution System Maintenance," *Journal of Water Resources Planning and Management,* ASCE, 117(3):321–339, 1991.

Shamir, U., and C. D. Howard, "Water Supply Reliability Theory," *Journal of the American Water Works Association,* 73(7):379–384, 1981.

Shamsi, U., "Computerised Evaluation of Water Supply Reliability," *IEEE Transactions on Reliability,* 39(1):35–41, 1990.

Shinstine, D. S., "Reliability Analysis Using the Minimum Cut-Set Method for Three Water Distribution Systems," M.S. thesis (Civil Engineering), University of Arizona, Tucson, 1999.

Shinstine, D. S., I. Ahmed, and K. E. Lansey, Reliability Availability Analysis of Municipal Water Distribution Networks: Case Studies, *Journal of Water Resources Planning and Management*, ASCE, 128(2):140–151, 2002.

Su, Y., L. Mays, N. Duan, and K. Lansey, "Reliability Based Optimization for Water Distribution Systems," *Journal of Hydraulic Engineering,* ASCE, 113(12):589–596, 1987.

Tanyimboh. T. T., M. Tabesh, and R. Burrows, "Appraisal of Source Head Methods for Calculating Reliability of Water Distribution Networks," *Journal of Water Resources Planning and Management,* ASCE, 127(4):206–213, 2001.

Tanyimboh, T. T., and A. B. Templeman, "Optimum Design of Flexible Water Distribution Networks," *Civil Engineering Systems,* 10(3):243–258, 1993.

Tanyimboh, T. T., and A. B. Templeman, "A New Method for Calculating the Reliability of Single-Source Networks," *Developments in Computational Techniques for Civil Engineering,* B. H. V. Topping (ed.), Civil-Comp Press, Edinburgh, U.K., 1995, pp. 1–9.

Templeman, A. B., and D. F. Yates, "Mathematical Similarities in Engineering Network Analysis," *Civil Engineering Systems,* 104(1), 1984.

Tung, Y. K., "Evaluation of Water Distribution Network Reliability," *Proceedings of the ASCE Hydraulics Specialty Conference,* New York, 1985, pp. 359–364.

Tung, Y. K., "Uncertainty and Reliability Analysis," in L. W. Mays (ed.), *Water Resources Handbook,* McGraw-Hill, New York, 1996.

Wagner, J., U. Shamir, and D. Marks, "Water Distribution Reliability: Analytical Methods." *Journal of Water Resources Planning and Management,* ASCE, 114(3):253–275, 1988a.

Wagner, J., U. Shamir, and D. Marks, "Water Distribution Reliability: Simulation Methods," *Journal of Water Resources Planning and Management,* ASCE, 114(3):276–293, 1988b.

Wu, S-J., Y-H. Yoon, and R. Quimpo. "Capacity-Weighted Water Distribution System Reliability." *Reliability Engineering and System Safety,* 42:39–46, 1993.

Xu, C., and I. Goulter, "Probabilistic Model for Water Distribution Reliability," *Journal of Water Resources Planning and Management,* ASCE, 124(4):218–228, 1998.

Yang, S.-L., N. S. Hsu, P. W. F. Louie, and W. W.-G. Yeh, "Water Distribution Network Reliability: Connectivity Analysis," *Journal of Infrastructure Systems,* ASCE, 2(2):54–64, 1996a.

Yang, S.-L., N.-S. Hsu, P. W. F. Louie, and W. W.-G. Yeh, "Water Distribution Network Reliability: Stochastic Simulation," *Journal of Infrastructure Systems,* ASCE, 2(2):65–72, 1996b.

CHAPTER 11

GEOGRAPHIC INFORMATION SYSTEMS APPLIED TO URBAN WATER SUPPLY SYSTEMS

Gonzalo López Patiño
Francisco Javier Martínez Solano
Fluid Mechanics Group
Universidad Politécnica de Valencia
Valencia, Spain

11.1 INTRODUCTION

When studying a *water supply system* (WSS), one needs to consider a huge amount of information in order to understand the WSSs, hydraulic performance or to determine how to manage the existing resources of the WSS efficiently. There are many types of information, but the information can be classified into three main groups, depending on its nature and its later usage. These groups are

- Physical features of the network elements
- Economic information of the WSS
- Spatial information about the location of both economic and physical data

Traditionally, this information has been saved in different formats. The information about network elements (diameters, lengths, starting date, suppliers, etc.) was saved in work plots or small inventory databases. The economic information was the most carefully kept database in the system. This database stored all the consumer data, including customers' demands, addresses, and registering dates, and other relevant data for correct economic management of the system. Finally, spatial information was usually scattered in various topographic maps where the izolines of the supplied geographic area, the location of the mains, and the distribution pipes layout appeared. However, most of the time these were not updated sketches. In any case, a connection rarely existed among the three information systems.

However, nowadays there is a greater need to link spatial, economic, and physical information together. This is now possible thanks to the implementation of a proper *geographic information system* (GIS). This system allows us not only to link geographic or spatial data with other alphanumeric data, but also to update in a simple way the included data, through an appropriate graphical interface.

This chapter presents an introduction to the GIS concept, reviewing briefly the elements it consists of and the different solutions that can be found for the information storage. Then various applications of the GIS within the limits of the WSS, such as planning, projecting, operation, and technical management of the network, will be presented. Finally, some proposals of future trends of the GIS technology applied to the water supply management will be given.

11.2 OVERVIEW OF GEOGRAPHIC INFORMATION SYSTEMS

11.2.1 Definition of a Geographic Information System

Multiple definitions of a GIS exist. Some of them are more complete than other ones, but all highlight two basic features of the GIS:

- A GIS is, basically, a database that must have the same tools as ordinary databases.
- A GIS stores and links spatial data (position or location) with thematic data (alphanumeric attributes).

With these two features in mind, one of the simplest definitions of a GIS is "a system for the geographical information capture, storage, extraction, analysis and visualization" (Hernández Rodríguez, 1995). This, without any doubt, shows the origin of the GIS as the integration of the computer-aided design (CAD) software with digital cartography management utilities and links with a database manager software.

The differences among the various solutions commercially presented as GIS are based, on one hand, on the type of spatial data they manage and, on the other hand, on the way these data are stored and related with attributes.

It is also possible to understand a GIS as a management philosophy. In this case, it would be a way of making decisions within an organization based on information that is managed in a centralized way but ordered depending on the organization's geographic location.

A more global definition could identify the three interconnected components of the GIS, as shown schematically in Fig. 11.1 (Parsons, 1997).

- A set of data ordered depending on their geographic location
- Some equipment and software to manage and link these data
- A specific problem or target that is intended to be solved from the spatially distributed data with the aid of the available tools

In conclusion, it can be said that a GIS is a georeferenced database, that is, a database in which the data are located by means of geographic coordinates with respect to some reference system. This is the main difference between GIS and CAD software (Rossiter, 1994).

11.2.2 Spatial Data Representations

The different spatial data representations are called *spatial data models.* The spatial data models used by the GIS can be basically classified into two elemental models: the *raster model* (also called the *grid model*) and the *vector model.*

The Raster Model. The raster model of spatial data representation emulates reality through the creation of a regular grid. Not only the contour of each object is stored but also its interior. A grid or a mesh in which every cell has the same

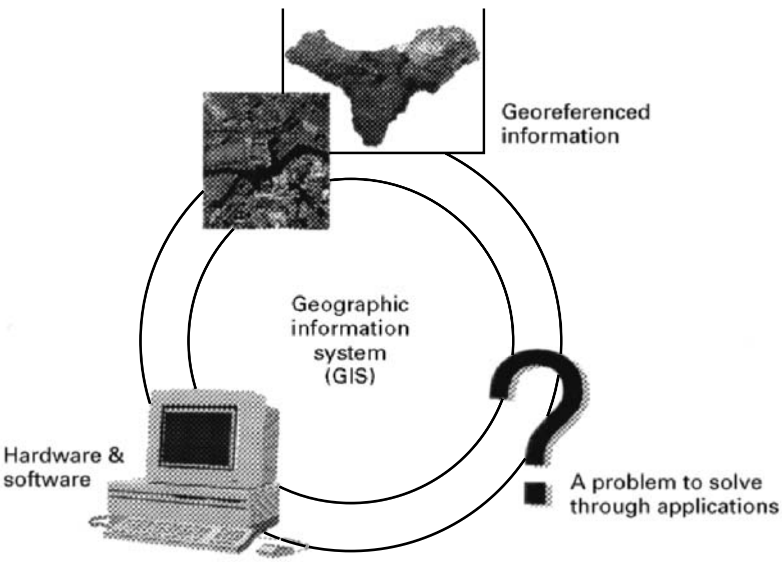


FIGURE 11.1 Schematic representation of GIS components.

shape and size is usually applied. In this way, a matrix in which every cell stores the value the selected variable takes at that point is obtained. This stored variable may be both qualitative (as in the case of land uses or type of water demand) and quantitative (elevation or flow rate demanded in every point of a network).

Several data structures have been described for permanent storage of this information. The first one would be *exhaustive enumeration.* It consists, basically, of gathering the contents of each cell in an individualized and sequential way. Obviously, when the plot is big or the cell size is too small, the amount of information starts becoming excessive; this is the main inconvenience of this structure. As an alternative to optimize the resources, a row enumeration is proposed in which the value of the variable and the first and the last column in which this value is kept are given. This structure is especially suitable in systems in which the spatial correlation principle is present, because most of the contiguous cells have the same value. Occasionally, only the value and the last cell where the variable takes this value are stored; the column where the next set of values starts is understood. Sometimes, even the value and the number of cells that keep this value are saved.

Whatever data structure is chosen to store the information contained in a plot, the main limitation of the raster model is that, for each variable, a plot is recorded. This means that for each variable it is necessary to store a *layer,* or *cover,* of information in the way described above. In Fig. 11.2, a schematic of how the information from a plot is stored in the raster model is shown. It is possible to see that a grid overlaps the plot and each cell takes a value depending on the land use it holds. In this case, a color code was used to name each land use in the plot, but for each cell the land use that takes the biggest percentage of it has been represented through an alphanumeric code (a letter).

This kind of data model is specially suitable for geographic variables, in which few regular geometric shapes are found and the exact shape of areas with the same value must be described with a lot of polygons and segments. Thus, its main application is the terrain description (digital elevation models, aspect maps, slope maps, etc.).

Since this storage method may consume a lot of memory, some alternative structures have been used to compress the raster data. For instance, in those cases in which there are some special areas of interest together with another without any interest, a quadtree structure is used. The *quadtree,* or hierarchical structure, is a data structure that subdivides any given space into four quadrants. Each of these quadrants may contain both data information and another grid with a smaller grid width, and continues to subdivide each quadrant in a similar way until they are uniform or the basic resolution of the data for the area is reached.

The Vector Model. Another way of representing spatial data is to reduce the areas where the variables take a fixed value to simple geometries (such as points, lines, or surfaces). In simplest form these geometries may be stored as a set of points or ver-

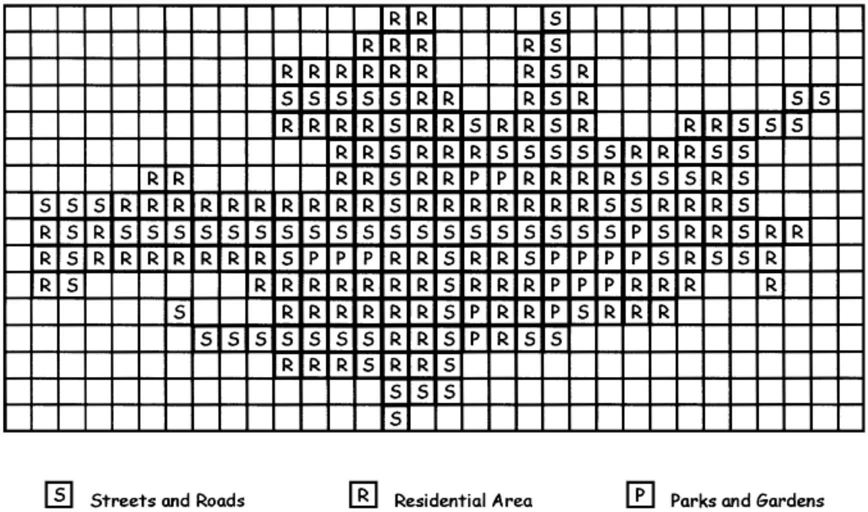


FIGURE 11.2 Scheme of the raster representation of the information contained in a map.

tices with instructions on how to connect these points. Another alternative is to store data that can be represented as lines (for example, as in the case of the layout of the water distribution network), as a mathematical function, for instance a combination of polynomials. As can be seen from the description above, there are two main differences between the vector model and the raster model. The first one is that in the case of the vector model, the boundary or limits of geographic objects in the map are explicitly stored, but the object's inner contents are not. The second difference is that only one value of the considered variable is stored for each object. For example, for a named main consisting of several pipes or segments with homogeneous characteristics, only one value of this characteristic will be recorded, while as mentioned above, in the raster model one value of the characteristic will be recorded for each cell where the main goes through.

In Fig. 11.3, a schematic of the vector model representation of the information of the same map shown in Fig. 11.2 can be observed. Streets are represented by lines of a certain color and width; meanwhile the areas covered by vegetation (parks and gardens) or buildings are represented with polygons with different shadings.

In the vector model the data are stored as objects, and each object can be classified into one of two main types depending on nature (OGC, 1999):

- *Geometric object.* Geometry provides the means for the quantitative description of an object, by means of attributes (its coordinates) or mathematical functions based on such attributes. This includes its position, dimension, size, shape, and orientation. The attributes and mathematical functions of a geometric object depend on the type of coordinate system used to define the spatial position. So, if the coordinate reference system changes, the geometric object changes too.

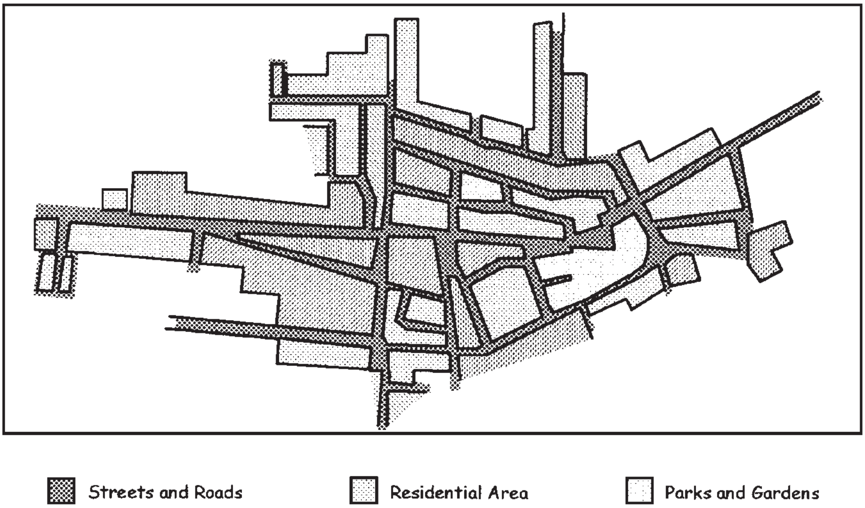


FIGURE 11.3 Schematic representation in vector mode of the information contained in a map.

- *Topological object.* Topology deals with the characteristics of the objects that remain unchanged if the space is deformed elastically and continuously. This is the case of a change in the coordinate reference system. The most common type of topology is the connectivity in a network. In this case, a change in the coordinate reference system does not affect the relations among the different objects in the topology.

The basic objects in a vector GIS are classified according to their dimensions. Three main objects are commonly used:

- *Point or node.* A node has dimension 0 and is stored in the computer with its position given by its coordinates.
- *Lines or arcs.* Depending on the type of object, a line can be represented as at least the coordinates of its end nodes (geometric object) or a function (straight or some other function) that connects two *nodes* (topological object).
- *Polygons or areas.* A polygon consists of a chain of lines where each line shares a start or end node with each adjacent line. If the coordinates are stored, then it will be a geometric object; if the lines that make up the polygon are stored, it will be a topological object.

Raster or Vector Model. When developing a data model based on a GIS, the model structure choice is important. Table 11.1 presents some advantages and disadvantages of each model (Reca, 1997).

TABLE 11.1 Comparison between Raster and Vector Models

SI G raster	SIG vectorial
Advantages	Advantages
1. Simple data structure	1. Compact data structure
2. Easy layer overlay and calculations	2. Efficient spatial data encoding
3. Suitable for high spatial variability (i.e., digital elevation models)	3. Power spatial analysis, proper for net work applications
4. Suitable for scanning, image analysis software, and remote sensing	4. High-quality output
Disadvantages	Disadvantages
1. Storage-consuming data structure, at a fixed resolution	1. Complex data structure
2. Information loss at any resolution	2. Not appropriate for terrain description or continuous surfaces
3. Difficulties with topological relations among data	3. Difficult algebraic operations as layer overlay
4. Any data transformation must be applied to every cell	4. Expensive software
	5. Image manipulation is difficult in vectorial models

The raster model is used for highly variable models, such as digital elevation models, hydrologic runoff models, or vegetative cover models. On the other hand, the vector model is used for applications where a line network analysis is required, such as for electrical, water supply, or storm water sewer networks.

Utility companies often use the vector model. A company will use the GIS to maintain an updated asset database, which frequently forms the central part of the company's performance strategies. Within the vector GIS applications, the Automated Mapping and Facilities Management (AM/FM) functions are most frequently used to manage the plant of the organization (Parsons, 1997).

11.2.3 Structures of the Geographic Databases

One of the characteristics of the georeferenced database is that the structure or the contents of this database can be hidden from the final user by setting some safety rules. Even though it is not necessary for the user of a GIS to understand in detail how the data are ordered and stored in the memory of the computer, some knowledge will help the user to understand how the system works and what the advantages and limitations of the database manager system are.

The main characteristic of any data storage system is that it allows quick access to data and cross-references. There are several ways to achieve this, from the simpler *file list* structure, to the *sequentially ordered file* and *indexed file* structures, to more complex structures used by the database manager. As the number of records in the database increases, it becomes more important to choose a structure that will provide for quick access to the data.

There are four main types of database structures commonly used in the GIS. These types are *hierarchical, network, relational,* and *object-oriented* databases. The first two types were used extensively during the seventies and eighties. The relational database is the leading type used today. The object-oriented structure is newer, and, even though it is still under development, it is prevailing in the most powerful database software.

Next we present a comparison between relational structures and object-oriented structures.

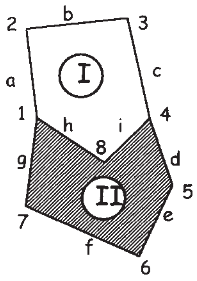
Relational Structure. The two structures defined up to this point, hierarchical and network structures, are known as navigational structures. This terminology arises from the utilization of pointers as links: the application navigates through the database using those pointers. The only paths for searching are those previously defined.

In contrast to what happens in the pointers navigation, the relational database uses the concept of two-dimensional interconnected tables. Every table has a number of rows or records, and every record has a set of predefined attributes or fields. The tables are linked through a set of primary and secondary keys. In this case it is the user who creates the links when it is suitable: when creating the database, during the data input, or later during the daily work. During a particularly complex query the database manager can create temporary links and tables. Without any doubt, the concept of *query* is the key to the relational structure of the database. The first step in the table design is to separate these in such a way that data redundancy results are reduced to a minimum.

Figure 11.4 shows the same two adjacent areas used in the example describing the hierarchical and network structures, and the resulting set formed by four tables. In order to get information from the database, the user may build any kind of query using the key elements as the selection criteria. For example, the user may form a query as: "Select all points belonging to both polygons."

Object-Oriented Structure. Today, the generalized trend in the development of information technologies is to use the object-oriented (OO) approach. Object-oriented systems are more modular and easier to maintain and describe than the traditional procedure systems. So a whole series of specialized programming languages allow for the easy implementation of object orientation in different fields. The term object oriented can be found in almost all modern software and, hence, also in the databases.

Graphical Information → Alphanumerical Data



Polygon Table

Polygon	Line 1	Line 2	Line 3	Line 4	Line 5	Line 6
I	a	b	c	i	h	-
II	d	e	f	g	h	i

Line Table

Line	Point 1	Point 2
a	1	2
b	2	3
c	3	4
d	4	5
e	5	6
f	6	7
g	7	1
h	1	8
i	8	4

Point Table

Point	X	Y	Z
1	x_1	y_1	z_1
2	x_2	y_2	z_2
3	x_3	y_3	z_3
4	x_4	y_4	z_4
5	x_5	y_5	z_5
6	x_6	y_6	z_6
7	x_7	y_7	z_7
8	x_8	y_8	z_8

FIGURE 11.4 Relational structure of database.

Usually, the term object orientation is used without a clear definition of what it actually does. One of the definitions of object orientation is: the representation of a given entity, independently of its structure and complexity, as an object in an exact way. This definition shows the contrast between what happens in the object-oriented data structure and what happens in the other data structures, where an entity must be separated into its basic geographic constituent parts. This definition also includes the fact that the database system will store and handle the object as it is and not as a collection of parts. The object is a software element that has identity, attributes, behavior (or methods), and its own functions and operations.

A typical description of a data model for an object-oriented database is shown in Fig. 11.5. We now analyze the representation of a polygon. A polygon object has, for example, a set of at least three defining points. In addition, there must be at least three sides of the polygon, defined for another set of two points. The polygon also has a function that calculates its area, and every side of the polygon has a function to calculate the side length. In order to calculate the polygon perimeter, for example, only the sum of the length of all its defining sides is necessary.

11.2.4 Topologies

Topology is one of the most important concepts to understand about the data structure in vector model GIS. It allows for great flexibility and powerful links among data. Furthermore, a topology includes its own functions for spatial querying. Each of the basic vector structures (point, line, and polygon) has intrinsic information:

- Its location, either through a geometric object or a set of coordinates

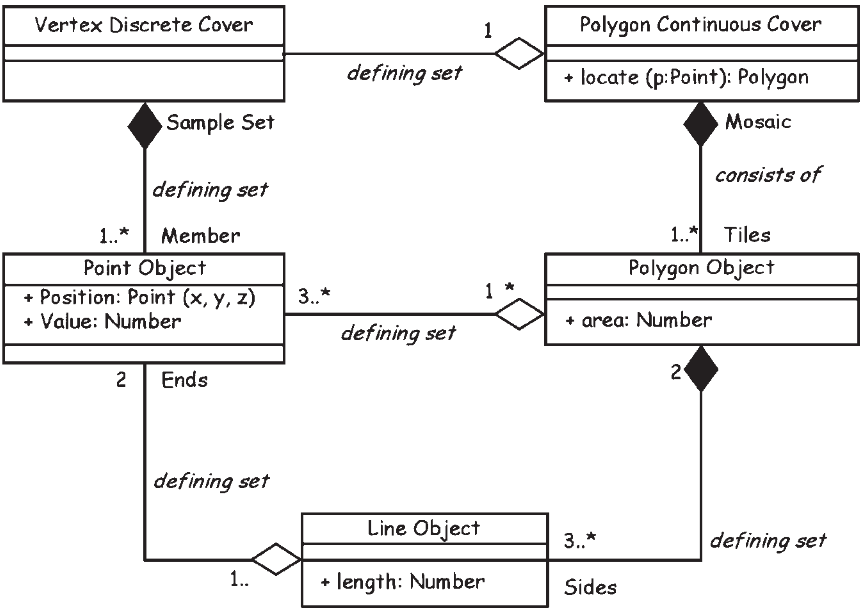


FIGURE 11.5 Context diagram (or entity-relation diagram) for a polygonal topology. (Adapted from OGC, 1999)

- Its relations with other vector structures (encoded as adjacency, connectedness, and containment)
- The ability of its attributes to be used in spatial queries and external functions

All these concepts are obtained by means of the *orientation* concept. In this sense, a line is defined with a direction and an orientation in such a manner that it can be run in direct sense, reverse sense, and both senses. A point may allow for circulation through it or not, and a polygon may have neighbors surrounding it or islands inside it.

The most important aspect of a topology consists of its capacity to understand the objects in its environment. The topology links every part in the data structure in such a way that it works as desired. The main relations (also called *spatial operators*) that may be performed by a topology are

- *Adjacency.* This operator allows the user to determine if two objects have common boundaries or not. If combined with the orientation of the objects, it also allows the user to set which object is to the right and which one is to the left.
- *Connectivity.* This is a very useful property when infrastructure networks, such as water supply, sewer systems, or telecommunications, are used. This

operator returns the connectivity of the objects in a set and can even identify which is the initial object and which is the end one.

- *Enclosing.* This is used mainly for polygon topology. This operator returns true if an object is included in another and false if they do not have this kind of relation.

Even though these are the main spatial operators of topologies, there are some other ones, like distance, perimeter, area, or costs of transportation through the topology. The last one (costs of transportation) is often used for transportation uses, but it is also used for water networks as a way of defining corrective maintenance manoeuvres and performing a closing-pipes analysis of the network.

11.3 THE USE OF GIS IN WATER SUPPLY SYSTEMS: GENERAL OVERVIEW

The ability of GIS to connect to other similar systems, together with its huge storage capabilities, allows the user to establish close control over the network for its whole life. The life cycle of any element in the network has seven stages: planning, design, project, construction, operation, maintenance, and renewal. This control ensures good-quality data, which is absolutely necessary for efficient management of the water distribution.

A GIS applied to the whole management of a water supply system has a structure as shown in Fig. 11.6. The central part of the system is a *georeferenced database manager,* which deals with alphanumerical data (such as customer information) and georeferenced data (such as customers' addresses). Depending on the objectives the company has, the GIS is often structured by services or departments. Basically, three services can be outlined:

- *Customer information system (CIS).* All the aspects concerning customers are controlled in this area: contracting, registering and dropout, billing, reclamation, and failures. This service must have full access to the customer information database.
- *Administrative management.* The water supply systems, owned by either public or private companies, must be subject to financial control. Economic and financial information of any water supply system is fundamental for its management. The possibility to deal with georeferenced information has added efficiency to the service. The department responsible for this management must provide access to both customers and infrastructure databases.
- *Technical mangement.* Finally, this is the service that deals with the information of the network. It has to control the whole life cycle of every element in the

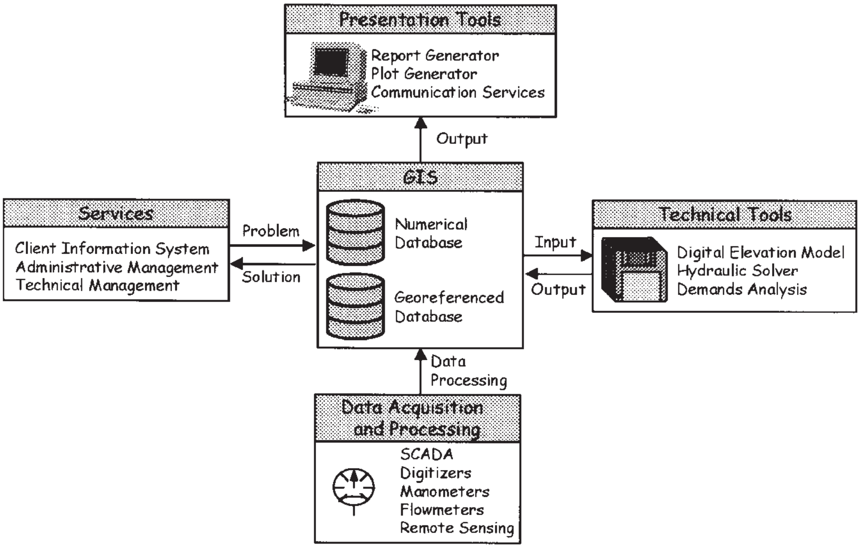


FIGURE 11.6 Configuration of a GIS in a water distribution system.

network. In this chapter, special attention will be paid to the use of GIS in the technical management of a water supply system.

Some functions of the technical management are

- Update the information existing in all the databases managed by the GIS, especially the network information.
- Plan, design, and project the new elements of the network, providing for its needs.
- Control the network operation through the creation, calibration, and continuous updating of a proper mathematical model.
- Maintain the existing infrastructure while performing a well-defined and well-updated inventory of the elements of the network. This will be useful in order to create a suitable maintenance and rehabilitation program.

External tools may aid most of these targets. This is the case of a model creator, which must simplify the whole network, and a hydraulic solver to calculate flows and pressures in pipes and nodes, respectively. Other useful tools, which sometimes are included in commercial GIS packages, are the digital elevation model manager or the demand analyzer. Usually, GIS must provide the data for these tools and be able to get the results into the results database.

Since the GIS is used by people, it has to include some interface in order to present results and introduce the data. Hence, there must be a set of presentation tools, such as report generators, plot, and graph generators, and communication systems. Finally, the GIS must provide some data acquisition and processing tools.

11.4 PLANNING, DESIGN, AND PROJECT

An optimal design module must accomplish two main objectives:

- Minimize the construction costs.
- Maximize reliability of the final solution.

These objectives must be subjected to some constraints, such as proper pressure levels at demand nodes. This complicates the calculations.

In the following paragraphs a methodology for planning new infrastructures in the network will be described. In order to increase the transport capacity of a network, new mains are to be installed. During the planning of such new mains, two steps are to be developed. The first one is the choice of the optimal layout of the new pipes. Secondly, it is necessary to calculate the expected demand of the users the new pipe will supply. Both steps can be done with the help of a GIS. A sketch of the overall design procedure is presented in Fig. 11.7.

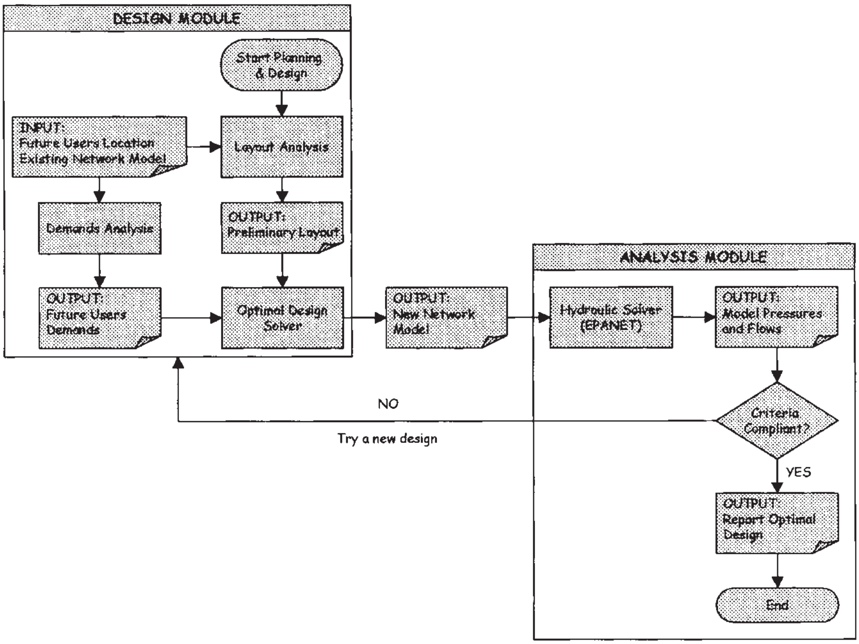


FIGURE 11.7 Flowchart for an optimal design module in a GIS.

11.4.1 Network Layout

The first step in designing a network is to decide where the new pipes are to be installed. This problem is relatively easy to solve if a GIS is available. In fact, a spatial operation has been described above. The problem can be presented as follows:

> Given a number of future users of the water supply systems and an existing water distribution network, find the cheapest layout in order to get all new users supplied from the existing network.

In other applications, mainly transportation applications, this problem is efficiently solved. An overlay between network spatial information and background cartography must be done. Usually, pipes are to be installed on the roads and rarely under the buildings. Hence, the problem consists of finding the shortest network that can link the new users' locations and the existing network.

Once this layout has been obtained, the design problem gets the usual statement:

> Given a number of future users of the water supply systems, an injection point and a distribution network, find the cheapest sizes for the pipes in order to get all new users supplied from the existing network.

This problem has been sufficiently studied over time (Pérez García, 1993; Taher and Labadie, 1996; Iglesias et al., 1999).

11.4.2 Demand Pattern Prediction with the Help of a GIS

Another important aspect (maybe the first to be taken into account) in the analysis of the volumetric efficiency of the network is the knowledge of the demands distribution on it. In the same way, to board an optimal strategy for control and tracking of uncontrolled flow rates, it is vital to have a suitable estimation of the demand evolution in the network along the day.

This estimation can be obtained using two different techniques:

- Extrapolating existing data
- Using economic and demographic models

In the first case, a unique database should be enough help provided it contains historic records of registered water demands along a given period of time. For the second method, the GIS capabilities may be crucial. The GIS allows the user to distinguish, for example, areas with different demographic and economic characteristics, which notably simplifies the use of demographic models for predicting future water duties and easily distinguishes even among areas with different trends. In this sense, the basic data that should be available are

- Geographical location of customers
- Identification of customers
- Quantifying of the average water consumption depending on the type of customer

This information is relatively easy to get if a GIS has been implemented in the WSS where the study is being developed. In fact, the demand estimation through a combination of GIS and demographic models was already used successfully in other fields within economic activities, such as the shop location depending on the profile of the possible consumer or the estimation of the stock size for each product.

To determine the average water duty for inhabitant or house connection (occasionally both duties must be superposed), some estimation based on socioeconomic criteria and technical criteria must be done. Among the socioeconomic criteria, income level of the customer, number of inhabitants per domicile, etc., could be included. The pressure level in the area and the diameter of the installed water meter are commonly used as technical criteria. So, combining water duty estimation and demographic trends in areas with different growth, it is possible to get satisfactory results from the water demands estimation.

Another aspect for which the use of a GIS may be helpful is for the distinction of the demand type depending on the land use. One must start from a geographic distribution of the different types of demands (ordered by areas with majority demand or by superposition of layers with percentages of each one):

- Domestic
- Industrial
- Leisure
- Agriculture

For the classification of land uses, either techniques based on the statistical analysis of historic records or analysis of aerial images collected by remote-sensing techniques can be used. In general, it is advisable to combine both techniques to limit the built-in uncertainty in the estimation of water demands with these methods.

As a consequence of this analysis, customer, house connection, street, or zone demands can be estimated. Therefore, we can obtain the value of the volumetric efficiency of the network comparing these estimations with the injected flow rate. It is also possible to estimate daily variations of the demands (this is known as the demand pattern). This will allow us to distinguish between the performance of the network during the peak and the off-peak hours.

Planning a network consists of laying out the pipes and sizing all the elements taking part in the new implantation zone. The new implantation network planning begins with needs' detection and locating.

The information management system has to provide the urban background cartography of the new zone, known or estimated from city urban plans. Once the urban background topology is known, optimal layout of the network is studied using the GIS. For this issue, the system evaluates the different alternatives considered, using available information on the database for each one and any established criteria, and chooses the better one.

The knowledge of water demand on the zone is needed for pipe dimension. As it is a new implantation zone, water consumption is estimated based on water duty or historical data of similar zones. Then, a forecast demand subsystem will be run, from the GIS, to do a forecast of major water consumption expected in the zone during the lifetime of the pipes.

Once water demand is estimated, the pipe diameter is calculated using optimization techniques that may be integrated into the GIS. In case there were no optimization techniques integrated into the GIS, sizing is checked with a hydraulic analysis of the network. A simulation, with the new data, is run from the system. The network analysis shows whether the layout and the diameter of pipes are enough to guarantee the service. If not, they will have to be modified. In fact, design techniques developed with GIS (Taher and Labadie, 1996) automatically runs a trial-and-error method.

11.4.3 Projects

Since generated information, dimensions, and topology belong only to the network design, they are not to be consolidated on the database. However, availability of this information is important to ensure that it is going to be updated afterward. New data introduced are stored in a nonconsolidated database, taking part in the global database but separated from real data for basic query operations. As far as the design becomes an executed project, and the customer is registered, data may be consolidated in the main database.

The need to introduce new information in the GIS must be taken into account from the moment the new elements of the network are planned. This will make the updating process of the information in the database easier and more reliable.

The implantation of new execution zones and the rehabilitation of some of the existing elements are considered in the project phase. The problems are similar, only distinguished by the credibility of base cartography data. For a new implantation network, base cartography is an estimation of reality. However, urban background is already consolidated in the database for rehabilitation.

The project can be developed by the water supply company or by another company, but in most of the cases the execution will be carried out by a construction company. In the project phase, which water supply company will maintain the network must be determined. If the project has any fault or the materials are not suitable for working conditions, failures would happen frequently and, as a consequence, network operation costs will increase and efficiency will decrease.

During the project phase, the GIS will be used to

- Get urban background information (from urban plans of the zone).
- Detect interactions with other urban infrastructures and services.
- Select elements from commercial databases.
- Provide automatical measures and budgets for the works.
- Update the new information on the database.
- Check if the project verifies the standards.

Establishing a standard on the project format will make updating the data in the GIS easier. The content of the project must include elements of the main and secondary network or house connections. In the first case, the standard ensures updating of graphic and thematic information. In the second one, the standard also includes the urban background of the parcel involving the building so that it may be consolidated in the database.

Project drawings at the standard format are introduced in the consolidated database. Later, control and pursuit of project execution allows the GIS to validate existing and modified data, if produced during execution.

To legalize the works, updated parcel information with all the modifications, is required at the standard format. This information is sent to the *management system* and, in case the work supervisor checks it as certain, it is consolidated.

When a customer is registered, the CIS checks if its house connections exist on the database. If the information is correct, graphical data are completed with a thematic one about the customer, improving the database.

Planning, design, and project system users manage only a little part of the whole information. They have access permission to this data exclusively. Access to other information not concerned with this system is not allowed. This is not only a security measure but also a way of keeping data reliable.

Access to the forecast demand application, used to estimate future water demands; hydraulic analysis, used to check network sizing; and pipe commercial databases, for pipe diameter selection, have to be allowed too.

Access to nonconsolidated information on the database is exclusive to this project system. No other system on GIS may have access permission to this data.

11.5 WATER SUPPLY NETWORK OPERATION

The normal operation of a water supply system involves some tasks to keep the parameters of the system, such as pressure, flow, leakage levels, and water quality, under control to ensure a high-quality service for customers and to reduce exploitation costs for the utility.

The water supply company staff has many control systems available to do the task. SCADAs, hydraulic software for modeling and analyzing the network, mapping facilities, maintenance management software, and others, are usual tools used to carry out the network operation. For a long time, all these systems have been used separately, even by different departments in the utility.

But most of these systems share the same data and are very related. For example, for hydraulic modeling, an updated database of the network layout is needed. So, a change in a pipe diameter due to a reliability task on the water supply system modifies the network behavior and must be automatically considered for the hydraulic model.

How is GIS involved in network operation? GIS is a powerful tool to manage most applications related to network operation to make them work as an integrated system.

11.5.1 Remote Control and Telemetry

On the operation of the network, pressure and pipe flows vary as a consequence of the change on demands. The goal on the regulation is to keep pressures on a range to ensure customers' demands and reduce water leakage levels and pipe failures on the network. The goal is achieved by acting on regulating elements on the network: pumps or pumping stations, valves, tank levels, chlorine injections.

Control on the network is done by means of a supervisory control and data acquisition (SCADA). The SCADA is a hardware and software system remotely linked to the regulated elements under control on the water system in order to monitor and control them. Data received is used for decision making on how the system must be operated.

The SCADA is a system that works by its own and does not need a GIS to operate, but it is usual that both systems are linked. As part of its capabilities, the SCADA system sends data that can be managed by the GIS.

The simplest SCADA system is composed of a communication system, linking the control center and every remote sensor, that transfers data on-line to be displayed at panel meters. Graphical interface software enhances the system as the operator has powerful access to information. The GIS may operate as the graphical interface of the SCADA as a temporal data storage.

GIS query capabilities give the water system operator better access to information because the information is located more quickly over the network and data can be statistically analyzed. Typical queries, as peak values of pressure and flow over the network, to the control points can be easily displayed with the GIS. The operator has a global vision on the water supply system behavior that a simple SCADA system does not provide.

The decision making on how to operate the regulating elements on the WSS is better and quicker with the aid of the GIS. Because real-time data may also be stored in the GIS database, other applications integrated to GIS, such as hydraulic modeling, can take advantage of that information. Nevertheless, the amount of real-time data is high enough. Most of those informations are only needed for network operation. Most GIS integrated applications only manage average or peak values at every point, so no more than these simplified values may be stored.

The GIS cannot directly manage control operations on WSS regulating elements, and communication hardware is needed. The SCADA or other special facility may be developed to integrate a remote control system to GIS. Control orders generated by the decision-making systems, system operators in most cases, are transferred to the SCADA system to be sent to the regulated elements by the control units.

11.5.2 Mathematical Model Update

A mathematical model of the network is used for hydraulic analysis in most of the systems. Model creation and calibration is fundamental for reliability of analysis results. Permanent growth of the network and pipe deterioration in time necessitate maintenance and a temporary update of the model. We will not comment on how to do this here, because it will be discussed in another chapter, but it is obvious that graphical information management possibilities offered by the GIS database and application of this information are quite useful to automating the model developing and updating process. The better the data are, the more reliable the results of the model are. That is why the GIS is so important in the creation of a mathematical model of a water distribution network.

The GIS can help basically in three stages of the model usage. During its creation, the GIS provides all the data referred to the elements of the network. The updating of the information is crucial at this point. All the data stored from the telemetry and remote control will be used for continuous calibration of the model. Finally, a reliable demand analysis module will help to simulate real or future scenes of the network.

11.5.3 Network Management

Network management is basically regulating operations on systems hydraulic elements to achieve an efficient and optimal water supply service. There are two different regulating operation types based on the purpose of regulation: normal network operation and network incident management.

Normal network operation is the on-line regulation to ensure water demands at consumption points are met. The regulation involves pressure and water quality control. During normal network operation, tank levels, pumps, and valve status are properly modified to keep parameters within range.

On-line management of hydraulic parameters is needed for control. The SCADA telemetry provides this kind of data, as mentioned above. Network elements information is needed as well. The GIS provides this other information if it is stored in it. The GIS manages both on-line parameters and network elements information to establish the operation strategy.

Water demand records are stored in GIS or CIS databases. A query to the database provides needed information for the demand-forecast model to generate a day-by-day scenario on future water consumption. The GIS uses these data for hydraulic model demand analysis. It can also generate the network hydraulic model, with a regulating strategy, and export data to the network analysis software to run the simulation, as commented. Simulation results are used to calculate exploitation costs. The system repeats the process with another regulating strategy if the previous one was not efficient enough. The process is based on GIS as a principal tool, but most specific applications must be developed to link GIS with other related software packages, like the demand forecast and the hydraulic model and analysis ones.

As described above, the process is very automated, but this is not the usual. A manual network operation is done in most cases. An experienced operator decides which regulating strategy to carry out. On-line regulation is done based upon the difference between network regulating goals and real-time hydraulic values on network control points, displayed on GIS or SCADA. For manual network operation, the GIS is just a powerful monitoring tool. The more automated regulation, the more the GIS improvement.

On the other hand, *network incident management* consists of all operations not generated by the normal regulation of pressures, water quality, and flows on the network, such as pipe breakdowns and valves closed up for maintenance.

There are two different types of incidents: scheduled and nonscheduled. Scheduled incidents are related to maintenance operations. Nonscheduled incidents are related to failures on the water supply system. From the network's operation point of view, both incidents have some similarities, but they also have some differences:

- Maintenance operations are planned on time so that a previous analysis of the influence of the maneuvers on the network can be done.
- Failures need an immediate response and the previous, and time-consuming, analysis cannot be done.

The incident requires isolation from a part of the water supply system. In both cases, an analysis on which valves and elements to close is needed. GIS network topology tools are used to calculate the elements to operate for isolation and the pipes affected. As the pipes are closed up, the customers supplied from these pipes have an interruption in water service. So, some closing alternatives are studied to identify the less problematic ones.

A working order to maintenance staff is generated. The staff closes up the marked elements. The status of those elements on GIS databases changes until the repair is completed.

A query to the GIS database for the pipes closed up provides a list of the customers affected by the shut off. When the incident is scheduled, customers affected by the shutoff can be notified. The GIS queries the CIS for customer telephone numbers, or even e-mail addresses, and an automated phone application can do the call.

In maintenance operations, the GIS generates the hydraulic model, with the closed-up elements, to analyze the more efficient regulating strategy for those customers affected, searching for alternative supplies that will ensure service, when possible.

11.5.4 Performance Indicators

Performance indicators, as suggested by their name, are parameters that enable access to the WSS management behavior. Availability of management and processing information tools to work with large amounts of information, which GIS integrates in its applications, is needed for the identification of performance indicators.

For example, the number of incidences on the network, water service pressures and their daily variation, water quality parameters, and leakage levels are parameters stored on the GIS database or easily calculated with basic query operations to the database.

Frequent GIS operations made by means of performance indicators are

- Infrastructure database queries
- Register of historic data of incidences
- Graphical linking of incidences and infrastructures
- Performance indicator linking to customers, network elements, network zones, etc.
- Performance indicator calculation

11.5.5 System Maintenance and Rehabilitation

Corrective Maintenance. *Corrective maintenance* consists of maintenance operations related to failed element replacement on the water supply system. A corrective maintenance operation is always the response to a warning on the network. The warning is detected by a customer complaint (for failures on small pipes), an abnormal pressure fall (detectable only on main pipes), or a notification of the leakage detection staff.

In the first case, the customer reports low pressure or water turbidity to the utilities operators. The operator is trained to query GIS for the reasons of the warning. The operator locates, based on the customer description, the warning in the GIS, and searches for any maintenance operation at the zone. If there is a maintenance operation, the customer is notified of the reasons for the warning and, if available, its length in time.

When there is no apparent reason for the warning, this is entered in the GIS database. If any other customer reports a service failure, the operator knows there has been a previous one. Performance indicator applications of the GIS also use this information to analyze water supply service quality.

The operator generates a working order. Maintenance staff checks for the warning and proceeds to the repair if that is the case. Sometimes the failure is difficult to locate and the working order is transmitted to the leakage detection department to determine the precise location. Once leakage is detected, it is repaired. The customer complaint is permanently stored in the GIS and graphically marked until the network incident is finished.

If failure is the consequence of a pipe breakdown, the pressure on the area of the leakage falls suddenly. When there is a control point on the area, the SCADA, or GIS, records this low-pressure value. It is compared to hydraulic simulation results, and a warning is generated when there is a significant difference. This kind of warning is limited to main pipes as a control point is needed and a significant low pressure must occur. On small pipes, maximum flows are limited so pressure falls are not large enough to be detected by the SCADA.

As this situation is related to a main pipe breakdown, water service on the area suffers the consequence. The GIS is used to analyze regulating strategies to keep the service on the system, as described before.

When the warning is sent by leakage detection personnel, a working order repair is simply generated to proceed. The process is similar to a rehabilitation one, described later.

In any case, the maintenance staff queries the GIS and gets an updated map of the area, including all network elements. Information on the GIS database may not be updated during failure repair, and this must be marked somehow at the system. Once the repair has been executed, the working crew reports all the network modifications to be updated in the GIS database. Later these data are verified to release the mark.

The GIS reduces the response time between the warning and the failure repair. Thus, better quality is achieved, and, if the failure is due to water leakage, the shorter the repair time, the lower the water loss and the better WSS efficiency.

Preventive Maintenance and Renewal. Preventive maintenance is the supervision of the network elements to prevent their failure. When the element replacement is done because of its poor status, this is called renewal. Preventive maintenance and renewal programs tend to guarantee not only water service but system efficiency, as the adequate behavior of every element on the network is ensured.

Maintenance programs are scheduled with the information given by the network element manufacturers. The manufacturer establishes review frequencies at the first stage. As time goes by, maintenance programs are modified by the utility staff based on its experience and failure historical records.

The GIS database stores maintenance dates for every element. A query to the database provides a list of the elements to review during a period of time. A GIS-based application schedules maintenance tasks to optimize available resources.

The *network management* department analyzes every planned operation, as described, and the better period for the maintenance task is introduced as a restriction in the scheduling application.

Once the maneuver is scheduled, it is stored in the GIS database available for the CIS. Thus, when the CIS receives a customer complaint about water service, the operator has the available information about all maintenance tasks.

Sometimes, preventive maintenance requires element replacement. While planning the works, the GIS database is queried to determine if all the elements needed for replacement are in stock. In either case, an order is sent to the shopping department to acquire the needed elements.

The maintenance staff reports the modified elements to the GIS administrator to update the database with the new ones. Replaced elements are not completely removed from the database because they provide useful information about the elements' lifetimes. These data are statistically analyzed to get information about the quality of elements and manufacturers, about suitable materials, etc.

Many other procedures, such as element location, are similar to corrective maintenance ones, and therefore will not be discussed here.

Network renewal is a programmed replacement of elements whose lifetime has extinguished. This renewal is limited by utility annual founds availability.

Element data needed for replacement scheduling are installation date, hydraulic behavior, working conditions, element status, maintenance reviews, etc. A GIS database query provides, based on this data a list of elements needed to be replaced every year. If the GIS has a pricing database of network elements, an application can prioritize the elements to be replaced using available annual funds for renewal and replacement costs.

Once it is known which elements are to be replaced, the process is similar to a maintenance one involving element replacement. The elements are located on the GIS. A closing pipe analysis is performed to isolate the zone. A hydraulic model, not including the out-of-service elements, is generated. A network management strategy is analyzed and the customers affected are notified about the work and its estimated length on time for completion. Then the tasks can start.

During the work, the elements to be replaced are marked on the GIS database so it is known that they are to be modified. When the work is finished, the GIS database is updated and then the marks are released.

11.6 FUTURE TRENDS IN THE DEVELOPMENT OF THE GIS

It is evident, after all the facts presented in this chapter, that we still are in a basic stage of the application of the GIS. Apart from some exceptions, most of the WSS that have faced the challenge of implementing a GIS are still developing preliminary phases of data introduction or design of applications to manage them. It can be concluded, therefore, that a first short-term objective for the application of this technology is the fulfillment of the initial phases of GIS implementation, at least in those systems whose size justifies the investment (Cubillo, 1997). From this point on, more ambitious targets could be planned, even a review of the trends of the different water supply companies may point to some characteristics GIS should have in the near future.

So, a constant feature in all the WSS that try to use the GIS is the trend to unify technical data with other administrative data in one information system. Some people even suggest the sharing of data with other utilities, like gas, electricity, or sewage companies, to create a huge corporate database that would allow tasks such as works management or predictive maintenance.

Another area in which some authors (Martín, 1997; Zaragoza, 1997) agree is the need to converge to open platforms in graphical object-oriented environments which make the use of these systems easier, independent of the user's skill. This will have repercussions on the user motivation and the increase of the reliability of the works made by them.

Some tendencies toward decentralization of the information system can be noted. This decentralization will allow for work tracking and file updating with the help of portable computers and Internet and Intranet connections. The future in this case points to client/server-like architectures.

Since the amount of data in a GIS is continuously growing, the techniques for data mining and spatial data mining are getting more and more important in GIS implementation.

Finally, other trends that may be enumerated are

- Joint modeling of spatial and temporal variations in a unique information system.
- Implementation of teledetection and real-time control systems for networks
- Inclusion of knowledge-based systems and expert systems for the decision-making support
- Three-dimensional ordination of the spatial information, in order to allow accessibility studies

It can be concluded that the future presented by GIS applied to the WSS is broad. In the case of network management, it could bring about a more efficient use of resources and an improvement in the service quality.

In final conclusion to what has been presented in this chapter, it could be said that the technical solutions when the water shortage has been reached are always desperate and in this sense the GIS does not bring miracles. It is possible to look for solutions to emergencies, but avoiding these situations through a suitable management of the existing resources is where the GIS could be of the greatest help.

11.7 REFERENCES

Cubillo, F. et al. *Guía para la implantación de Sistemas de Información en la gestión de redes de suministro de agua.* CENTA, Seville (Spain), September 1997.

Hernández Rodríguez, F. *Modelización de Información Espacial mediante Tecnología orientada a Objetos.* PhD thesis. Seville, March 1995.

Iglesias Rey, P. L., J. Izquierdo Sebastián, G. López Patiño, and F. J. Martínez Solano. "Geographical Information Systems Applied to Water Supply Systems." In E. Cabrera and J. García-Serra (eds.), *Drought Management Planning in Water Supply Systems.* Kluwer Academic Publishers, Dordrecht, Holland, 1999.

Martín, A. *Aplicación de los SIG a la Gestión Integral de los Sistemas de Distribución de Agua.* Curso de Diseño, Análisis, Operación y Mantenimiento de Redes Hidráulicas a Presión. Máster en Gestión y Uso Eficiente del Agua. Valencia (Spain), June 1997.

Open GIS Consortium (OGC). *The OGC Abstract Specification.* Wayland, Mass., 1999. (Available at *http://www.opengis.org/techno/specs.htm*)

Parsons, E. *The Essential Guide to GIS. A Hands-on GIS Workshop.* Internet Seminar, Kingston Center for GIS. Kingston (United Kingdom), July 1997. (*http://www.future-geomatics.com/esguide/start.html*)

Pérez García, R. *Dimensionado Optimo de Redes de Distribución de Agua Ramificadas considerando los Elementos de Regulación.* PhD thesis. Polytechnical University of Valencia, 1993.

Reca, J. *Estado de los SIG en España. Estado de los SIG en aplicaciones agrícolas.* Curso de Sistemas de Información Geográfica aplicados a Redes Hidráulicas. Máster en Gestión y Uso Eficiente del Agua. Valencia (Spain), November 1997.

Rossiter, David G. *Special Topics in Soil, Crop & Atmospheric Sciences: Land Evaluation, with Emphasis on Computer Application.* Notes from a course at Cornell University, 1994.

Taher, Saud A., and John W. Labadie. "Optimal Design of Water-Distribution Networks with GIS," *Journal of Water Resources Planning and Management* 4:301–311, 1996.

Zaragoza, J. *Gestión de Redes utilizando el SIG. El SIG en Aguas de Alicante.* Curso de Sistemas de Información Geográfica aplicados a Redes Hidráulicas. Máster en Gestión y Uso Eficiente del Agua. Valencia (Spain), November 1997.

CHAPTER 12
OPTIMAL OPERATION OF WATER SYSTEMS

Fred E. Goldman
GTA Engineering, Inc.
Phoenix, Arizona

A. Burcu Altan Sakarya
Department of Civil Engineering
Middle East Technical University
Ankara, Turkey

Larry W. Mays
Department of Civil and Environmental Engineering
Arizona State University
Tempe, Arizona

12.1 INTRODUCTION

12.1.1 Background

Until 1974, government regulators and water system operators concentrated their efforts on meeting water quality *maximum contaminant levels* (MCLs) at the water source or water treatment plant. However, with the implementation of the Safe Water Drinking Act in 1974 and amendments in 1986 and 1996, water system operators are being required to meet standards throughout their distribution systems and at the points of delivery. The *Surface Water Drinking Rule* requires that a detectable level of chlorine be maintained in the system. The *Total Coliform Rule* requires that total coliform standards be met at the customer's tap. The *Lead and Copper Rule* requires testing at the point of delivery (Clark 1994b; Clark et al. 1994, 1995; Pontius, 1996).

Water quality can undergo significant changes as the water travels through the water distribution system from the point of supply and/or treatment to the point of delivery. For example, chlorine concentration will decrease with time in pipes and tanks through bulk decay and through reaction with the pipe wall. Also, chlorine mass leaves the system at demand nodes or can be added to the system at chlorine booster nodes. Mixing of water from different sources can also affect water quality (Clark et al., 1995).

Pumps are usually operated according to an operating policy, which includes the scheduling of pump operation, which can affect water quality if the system pumps draw water from sources with differing water quality. Pump operation will affect the turnover rate of storage tanks (Ormsbee et al., 1989). Chlorine dosage may differ at each pump depending on the water source quality. For example, consider a pump that operates for 6 h divided into 1-h periods. The pump can either be on or off for any period. The number of combinations is $2^6 = 64$. A pump which operates for 24 h divided into 1-h periods has $2^{24} = 16{,}777{,}216$ combinations. The 6-h example can be solved by trial and error, but the 24-h example is prohibitively large to solve by trial and error. For large combinatorial optimization problems, simulated annealing provides a manageable solution strategy. These considerations have generated a need for computer models and optimization techniques that consider water quality in the water distribution system as well as system hydraulics.

Computer models that simulate the hydraulic behavior of water distribution systems have been available for many years (Wood, 1980). More recently these models have been extended to analyze the quality of water as well as the hydraulic behavior (Grayman et al., 1996; Rossman, 1993; Rossman et al., 1993; Elton et al., 1995; Boulos et al., 1995). These models are capable of simulating the transport and fate of dissolved substances in water distribution systems. The reason behind this is the governmental regulations and consumer-oriented expectations (Rossman and Boulos, 1996). As mentioned, the passage of the Safe Drinking Water Act in 1974 and its amendments in 1986 (SDWAA) changed the manner in which water is treated and delivered in the United States. The U.S. Environmental Protection Agency (USEPA) is required to establish maximum contaminant level goals for each contaminant that may have an adverse effect on the health of persons. These goals are set to the values at which no known or expected adverse effects on health can occur, by allowing a margin of safety (Clark et al., 1987; Clark, 1994).

The objective of this chapter is to describe optimization models that can be used to determine the optimal operation times of the pumps in a water distribution system for a predefined time horizon for water quality purposes while satisfying the hydraulic constraints, the water quality constraints, and the bound constraints of the system. Two solution methodologies are discussed. The first (described in Sec. 12.2) is based upon a mathematical programming approach (Sakarya, 1998; Sakarya and Mays, 2000) that links the nonlinear optimization code GRG2 by Lasdon and Waren (1986) with the water distribution simulation code EPANET by

Rossman (1994). The second approach (described in Sec. 12.3) for optimal pump operation is based upon a simulated annealing approach interfaced with the EPANET model (Goldman, 1998). The simulated annealing approach can also be used to determine the optimal operation of chlorine booster stations as described in Sec. 12.4.

12.1.2 Previous Models for Water Distribution System Optimization

Ormsbee and Lansey (1994), and more recently, Goldman et al. (2000), reviewed methods used to optimize the operation of water supply pumping systems to minimize operation costs. Ostfeld and Shamir (1993) developed a water quality optimization model that optimized pumping costs and water quality with both steady-state and dynamic conditions that was solved using GAMS/MINOS [general algebraic modeling system (Brooke et al., 1988)/mathematical in-core nonlinear optimization systems (Murtagh and Saunders, 1982)]. Ormsbee (1991) suggested that on-off operation of pumps poses a difficulty for optimization techniques that require continuous functions. Computer-based water quality models exhibit the same difficulties. Most modeling methods include numerical methods rather than continuous functions. Linear superposition was used by Boccelli et al. (1998) for the optimal scheduling of booster disinfection stations by developing system-dependent discretized impulse response coefficients using EPANET and linking with the linear programming solver (MINOS by Murtagh and Saunders, 1987).

Dynamic programming has been used for the optimal scheduling of pumps in a water distribution system by Coulbeck and Orr (1982) and Coulbeck, Brdys, and Orr (1987). The method is sensitive to the number of reservoir states and the number of pumps. In general, the increase in number of discretizations and number of state variables increases the size of the problem dramatically, which is known as the curse of dimensionality (Mays and Tung, 1992). This has restricted the application of this method to small systems. Other optimization approaches include Coulbeck et al. (1988a,b), Chase and Ormsbee (1989, 1991), Lansey and Zhong (1990), Lansey and Awumah (1994, Nitivattananon et al. (1996), Ormsbee and Reddy (1995), Ostfeld and Shamir (1993a,b), Sadowski et al. (1995), and Zessler and Shamir (1989).

The nonlinear programming approach by Brion and Mays (1991) evaluates gradients using finite differences or a mathematical approach to provide derivatives required by GRG2, which solves nonlinear optimization problems by using the generalized reduced gradient method (Lasdon et al., 1978; Mays, 1997). The method has been used for scheduling pump operations to minimize energy and improve water quality by Sakarya (1998) and Sakarya and Mays (2000), who considered three objective functions: (1) the minimization of the deviations of actual concentrations from a desired concentration, (2) minimization of total pump duration times, and (3) the minimization of total energy while meeting water quality constraints. The model by Sakarya and Mays (2000) is discussed in Sec. 12.2.

12.2 OPTIMAL PUMP OPERATION—MATHEMATICAL PROGRAMMING APPROACH

12.2.1 Model Formulation

Different objective functions can be considered to solve the optimal operation of water distribution systems for water quality purposes. In this model three objective functions are considered: case I is the minimization of deviations of the actual concentrations of a constituent from the desired concentration values; case II is the minimization of the total pump operation times; and case III is the minimization of the total energy cost. The constraints that have to be considered at all time periods for all cases are basically the hydraulic constraints; the water quality constraints; and the bound constraints on pump operation times, pressures, and tank water storage heights. In addition to the constraint set which case I considers, the bound constraints on the substance concentrations have to be satisfied for cases II and III.

Consider a water distribution system with M pipes, K junction nodes, S storage nodes (tank or reservoir), and P pumps, which are operated for T time periods. The objective function for case I is

$$\text{Min } Z_{\text{I}} = \text{minimize} \sum_{t=1}^{T} \sum_{n=1}^{N} \min\,[0, \min\,(C_{nt} - \underline{C}_{nt}, \overline{C}_{nt} - C_{nt})]^2 \qquad (12.1)$$

where N = total number of nodes (junction and storage)

C_{nt} = substance concentration at node n at time t, mass/ft^3

$\underline{C}_{nt}$, $\overline{C}_{nt}$ = lower and upper bounds, respectively, on substance concentration at node n at time t, mass/ft^3

The objective function for case II considers the minimization of the total pump durations expressed as

$$\text{Min } Z_{\text{II}} = \text{minimize} \sum_{p=1}^{P} \sum_{t=1}^{\text{T}} D_{pt} \qquad (12.2)$$

where D_{pt} is the length of time pump p operates during time period t in hours.

The objective function for case III considers the minimization of the total energy cost expressed as

$$\text{Min } Z_{\text{III}} = \text{minimize} \sum_{p=1}^{P} \sum_{t=1}^{T} \frac{\text{UC}_t 0.746 \text{ PP}_{pt}}{\text{EFF}_{pt}} D_{pt} \tag{12.3}$$

where UC_t = unit energy or pumping cost during time period t, \$/kWh)
PP_{pt} = power of pump p during time period t, horsepower (hp)
D_{pt} = length of time pump p operates during time period t, h
EFF_{pt} = efficiency of pump p in time period t

The efficiency of the pump and the unit energy cost are considered to be constant for all time periods.

The distribution of flow throughout the network must satisfy the conservation of mass and the conservation of energy, which are defined as the hydraulic constraints. The conservation of mass at each junction node, assuming water is an incompressible fluid, is

$$\sum_i (q_{ik})_t - \sum_j (q_{kj})_t - Q_{kt} = 0 \qquad k = 1, \ldots, K; t = 1, \ldots, T \tag{12.4}$$

where $(q_{ik})_t$ is the flow in pipe m connecting nodes i and k at time t in cubic feet per second and Q_{kt} is the flow consumed (+) or supplied (−) at node k at time t in cubic feet per second. The conservation of energy for each pipe m connecting nodes i and j in the set of all pipes M is,

$$h_{it} - h_{jt} = f(q_{ij})_t \qquad \forall\, i, j, \in M; t = 1, \ldots, T \tag{12.5}$$

where h_{it} is the hydraulic grade line elevation in feet at node i (equal to the elevation head E_i plus the pressure head H_{it}) at time t, and $f(q_{ij})_t$ is the functional relation in feet between head loss and flow in a pipe connecting nodes i and j at time t.

The total number of hydraulic constraints is $(K + M)T$, and the total number of unknowns is also $(K + M)T$, which are the discharges in M pipes and the hydraulic grade line elevations at K nodes. The pump operation problem is an extended period simulation problem. The height of water stored at a storage node for the current time period y_{st} is a function of the height of water stored from the previous time period which can be expressed as

$$y_{st} = f(y_{s,t-1}) \qquad s = 1, \ldots, S; t = 1, \ldots, T \tag{12.6}$$

The water quality constraint which is the conservation of mass of the substance within each pipe m connecting nodes i and j in the set of all pipes M is

$$\frac{\partial(C_{ij})_t}{\partial t} = -\frac{(q_{ij})_t}{A_{ij}} \frac{(\partial C_{ij})_t}{\partial x_{ij}} + \theta(C_{ij})_t \qquad \forall\, i, j \in M;\ t = 1, \ldots, T \tag{12.7}$$

where $(C_{ij})_t$ = concentration of substance in pipe m connecting nodes i and j as a function of distance and time, mass/ft^3

x_{ij} = distance along pipe, ft

A_{ij} = cross-sectional area of pipe connecting nodes i and j, ft^2

$\theta(C_{ij})_t$ = rate of reaction of constituent within the pipe connecting nodes i and j at time t, mass/ft^3/day

All the constraints considered up to now are equality constraints. In addition, there also exist inequality constraints, which are the bound constraints. The lower and upper bounds on pump operation time are given as

$$\Delta t_{\min} \leq D_{pt} \leq \Delta t_{\max} \qquad p = 1, \ldots, P;\ t = 1, \ldots, T \tag{12.8}$$

where D_{pt} is the length of the operation time of pump p at time t, and $\Delta t_{\min}$ and $\Delta t_{\max}$ are the lower and upper bounds on D_{pt}, respectively. $\Delta t_{\min}$ can be zero in order to simulate the pump closing, and $\Delta t_{\max}$ is equal to the length of one time period.

The nodal pressure head bounds are

$$\underline{H}_{kt} \leq H_{kt} \leq \overline{H}_{kt} \qquad k = 1, \ldots, K;\ t = 1, \ldots, T \tag{12.9}$$

where $\underline{H}_{kt}$ and $\overline{H}_{kt}$ are the lower and upper bounds, respectively, on the pressure head at node k at time t, H_{kt}. There is no universally accepted values for the lower and upper bound values. The range of 20 to 40 psi is acceptable for minimum pressure for average loading conditions but may be lowered during emergency situations such as a fire.

The bounds on the height of water storage are

$$\underline{y}_{st} \leq y_{st} \leq \overline{y}_{st} \qquad s = 1, \ldots, S;\ t = 1, \ldots, T \tag{12.10}$$

where $\underline{y}_{st}$ and $\overline{y}_{st}$ are the lower and upper bounds, respectively, on the height of water stored at node s at time t, y_{st}. These limits are due to physical limitations of the storage tank.

Cohen (1982) stated that, if there is not a requirement for the periodicity in operation of a network, then the optimization of the operation has no meaning. Hence, to achieve this, in the previous studies made on the pump operation problem, final storage bounds are tightened so that the storage in the tanks will be more or less the same as the initial states. However, a system reaches (approaches) steady state if the daily pump operation schedules repeat themselves for a certain period of time. A water distribution never actually reaches these steady-state conditions in reality. The tank water levels at the beginning and at the end of the day are equal to each other, when the steady-state conditions are reached. Hence, the constraint that forces the tank water levels at the end of the day to be more or less equal to the initial state is not considered. Instead of this, a constraint, which forces the water levels in the tanks to be within the desired limits, is considered.

Cases II and III consider the same constraint set defined by Eqs. (12.4) to (12.10) as case I, with an additional constraint for the substance concentration values to be within their desired limits. The bounds on substance concentrations are

$$\underline{C}_{nt} \leq C_{nt} \leq \overline{C}_{nt} \qquad n = 1, \ldots, N; t = 1, \ldots, T \tag{12.11}$$

where $\underline{C}_{nt}$ and $\overline{C}_{nt}$ are the lower and upper bounds, respectively, on substance concentration at node n at time t, C_{nt}.

The above formulation results in a large-scale nonlinear programming problem with decision variables $(q_{ij})_t$, H_{kt}, y_{st}, D_{pt}, and $(C_{ij})_t$. In the proposed solution methodology, the decision variables are partitioned into two sets; control (independent) and state (dependent) variables. The pump operation times are the control variables. The problem is formulated above as a discrete time optimal control problem. Equations (12.1) to (12.3) are the objective functions for cases I, II, and III, respectively, and the hydraulic and water quality constraints, Eqs. (12.4) to (12.7) define the simulator equations. Equation (12.8) is the control variable bound constraint. Similarly, Eqs. (12.9) to (12.11) are the state variable bound constraints.

12.2.2 Solution Methodology—Mathematical Programming Approach

The solution algorithm that is used in this study is a reduction technique, which is similar to the algorithms used for a groundwater management model by Wanakule et al. (1986), for water distribution system design by Lansey and Mays (1989), for operation of pumping stations in water distribution systems by Brion and Mays (1991), for optimal flood control operation by Unver and Mays (1990), and for optimal determination of freshwater inflows to bays and estuaries by Bao and Mays (1994a, b). In all these studies, the optimal solution of the problem is obtained by using a hydraulic simulation code together with a nonlinear optimization code.

Mathematical formulation of the pump operation problem is a large-scale nonlinear programming problem. The problem is reformulated in an optimal control framework where an optimal solution to the problem is arrived at by linking a simulation code, EPANET (Rossman, 1994), with an optimization code, GRG2 (Lasdon and Waren, 1986). The decision variables are partitioned into control variables and state variables in the formulation of the reduced problem. The control variables are determined by the optimizer and used as input to the simulator, which solves for the state variables. Hence, the state variables are obtained as implicit functions of the control variables. This results in a large reduction in the number of constraints as both the hydraulic constraints and the water quality constraints are solved by the simulator. Only the bound constraints remain to be solved by the optimizer. Figure 12.1 shows the linkage between the optimization and the simulation codes.

Improvements in the objective function of the nonlinear programming problems are obtained by changing the control variables of the reduced problem. NLP codes restrict the step size by which the control variables change so that the control variable bounds are not violated. The state variables, which are implicit functions of control variables, are not taken into consideration in the determination of step size. If the bounds of the state variables are violated, more iterations would be needed to obtain a feasible solution. The *penalty function method* is used to overcome this problem. The state variable bound constraints are included in the objective function as penalty terms. The application of this technique is also beneficial from the size point of view of the problem. Including the state bounds into the objective function results in a significant reduction in the size of the problem, such that the number of the constraints is also reduced.

There are many kinds of penalty functions that can be used to incorporate the bound constraints into the objective function. In this research, two different penalty functions, the bracket and the augmented lagrangian, are used. The bracket penalty function method (Reklaitis et al., 1983; Li and Mays, 1995) uses a very simple penalty function expressed as

$$\mathrm{PB}_j\,(V_{j,i}, R_j) = R_j \sum_i \left[\min\,(0, V_{j,i})\right]^2 \tag{12.12}$$

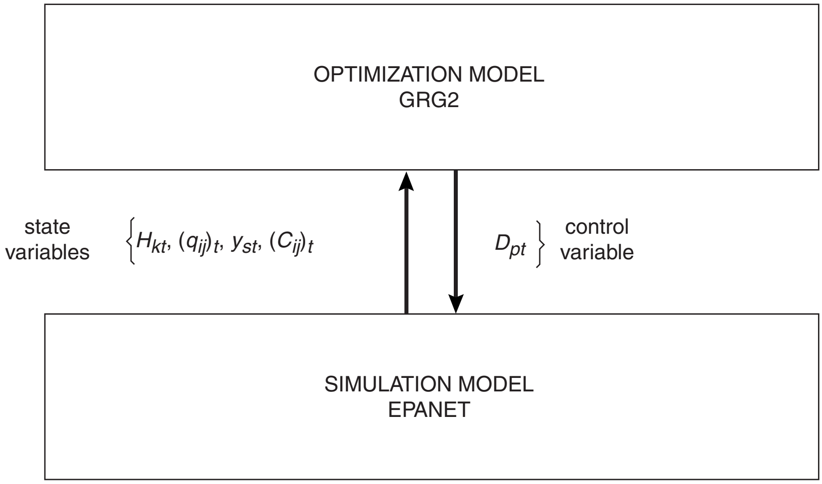


FIGURE 12.1 Optimization-simulation model linkage.

The augmented lagrangian method (Fletcher, 1975) uses the following penalty function

$$\mathrm{PA}_j(V_{j,i}, \mu_{j,i}, \sigma_{j,i}) = \frac{1}{2}\sum_i \sigma_{j,i} \min\left[0, \left(V_{j,i} - \frac{\mu_{j,i}}{\sigma_{j,i}}\right)\right]^2 - \frac{1}{2}\sum_i \frac{\mu_{j,i}^{\;2}}{\sigma_{j,i}} \qquad (12.13)$$

where the index j is the representation of H for pressure head, C for concentration, and y for water storage height bound constraints. The index i is a one-dimensional representation of the double index (k, t) for the pressure head penalty term, (n, t) for the concentration penalty term, and (s, t) for the storage bound penalty term. PB_j and PA_j define the bracket and the augmented lagrangian penalty functions for bound constraint j, respectively. $V_{j,i}$ is the violation of the bound constraint j, R_j is a penalty parameter used in the bracket penalty method; and $\sigma_{j,i}$, and $\mu_{j,i}$ are, respectively, the penalty weights and lagrangian multipliers used in the augmented lagrangian method. A detailed description of the bracket and the augmented lagrangian penalty function methods and the methods to update the penalty function parameters can be found in Sakarya (1998).

The violation of the pressure head constraint is defined as

$$V_{H,kt} = \min\,[(H_{kt} - \underline{H}_{kt}), (\overline{H}_{kt} - H_{kt})] \qquad (12.14)$$

Similarly the violations of the substance concentration and the water storage height bound constraints can be defined.

The reduced problem for case I is

$$\text{Min } L_{\mathrm{I}} = P_C(V_{C,nt}, F_C) + P_H(V_{H,kt}, F_H) + P_y(V_{y,st}, F_y) \qquad (12.15)$$

subject to

$$0 \le D_{pt} \le \Delta t \qquad p = 1, \ldots, P;\ t = 1, \ldots, T \qquad (12.16)$$

where P_C, P_H, and P_y define the bracket or the augmented penalty terms associated with the concentration, pressure head, and storage bound constraints, respectively, depending on the penalty function method used. Similarly, F_C, F_H, and F_y define the penalty function parameters which are the penalty parameters for the bracket or the penalty weights and the lagrangian multipliers for the augmented lagrangian penalty method, associated with the concentration, pressure head, and storage bound constraints, respectively.

For case II, the reduced objective function is subjected to the same constraint defined by Eq. (12.16) as case I, and has the following form

$$\text{Min } L_{\mathrm{II}} = \sum_{p=1}^{P}\sum_{t=1}^{T} D_{pt} + P_C(V_{C,nt}, F_C) + P_H(V_{H,kt}, F_H) + P_y(V_{y,st}, F_y) \qquad (12.17)$$

The reduced objective function for case III is defined as

$$\text{Min } L_{\text{III}} = \sum_{p=1}^{P}\sum_{t=1}^{T} \frac{\text{UC}_t 0.746\text{PP}_{pt}}{\text{EFF}_{pt}} D_{pt} + P_C(V_{C,nt}, F_C) + P_H(V_{H,kt}, F_H) + P_y(V_{y,st}, F_y) \quad (12.18)$$

subject to Eq. (12.16).

The solution of the final form of the problem is obtained by the two-step optimization procedure described in Wanakule et al. (1986), Lansey and Mays (1989), Brion and Mays (1991), and Mays (1997). The finite-difference approximations were used to calculate the derivatives of the objective function with respect to pump operation times, which are the reduced gradients that the optimization code needs to find the optimal solution. Initially, the original objective function and the penalty method are chosen depending on the case being considered. The penalty function parameters, which are lagrangian multipliers and penalty weights for the augmented lagrangian penalty method, and penalty parameters for the bracket penalty method associated with three terms in the objective function are fixed, and the inner level optimization is made. The control variable, D_{pt}, is solved by the optimizer. After the optimal values of D_{pt} for the fixed penalty function parameters are obtained, the outer-level loop is carried out to update the penalty function parameters, if necessary. A new problem is formed with these updated penalty function parameters, and the inner-level optimization is carried out again to solve for the control variables. The procedure is repeated until the penalty function parameters are not updated (i.e., overall optimum solution is found), the iteration limit of the outer loop is reached, or no improvements are achieved for some predefined consecutive iterations. The flowchart of the optimization model is given in Fig. 12.2.

During the solution procedure, GRG2 needs the values of reduced objective function and the reduced gradients while searching for the optimum solution. The calls made to calculate the values of the reduced gradients are named PARSH calls, and the calls made to calculate the reduced objective function are named FUNCTION calls. For each iteration step in the inner loop, GRG2 changes the control variables, depending on the values of the reduced objective function and the reduced gradients of the objective function.

A simplified method is used to reduce the number of EPANET calls. The main idea of this simplified method is that, if the maximum change in the control variables between consecutive iterations is small, the change that occurs in the state variables is also small. Hence, if the maximum change in the control variables between consecutive iterations is smaller than a specified limit, as the change in the values of the state variables will also be small, EPANET will not be called at that iteration to calculate the state variables; instead the previous values will be used.

12.2.3 North Marin Water District

The North Marin Water District (NMWD) serves a suburban population of approximately 53,000 people who live in the northern portion of Marin County,

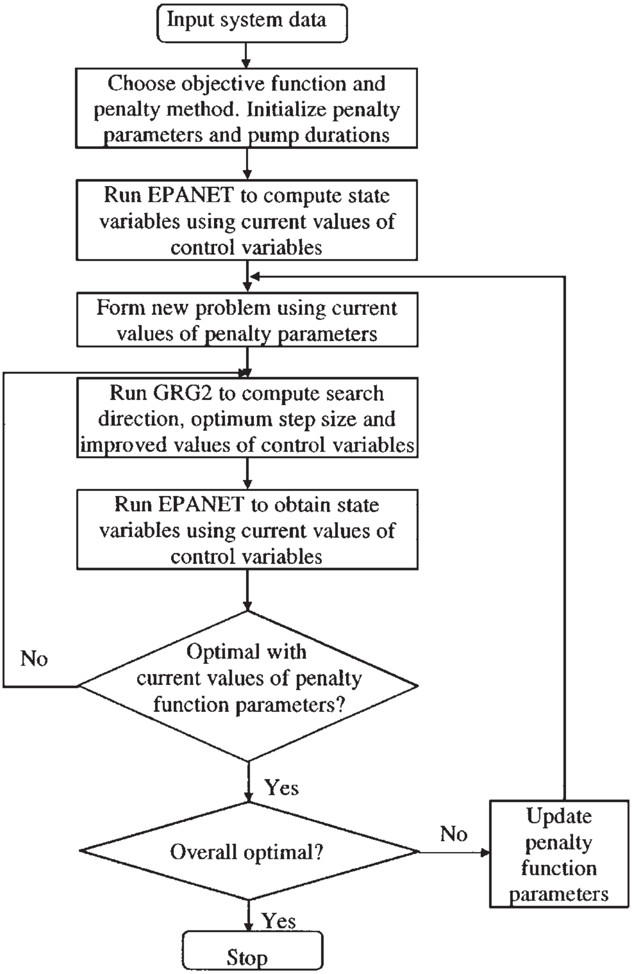


FIGURE 12.2 Flowchart of the optimization model.

California. The NMWD system is divided into a series of zones. In this research, only the primary (zone I) zone of the NMWD is used, after making some changes. Figure 12.3 shows the water distribution system of the North Marin Water District zone I, used in this study. The network has 91 junction nodes, 115 links, 2 pumps, 3 storage tanks, and 2 reservoirs. Sixty of the junction nodes are demand nodes. The diameters of the pipes in the NMWD system range between 8 to 30 in. The minimum and maximum pressures at demand nodes were set at 20 and 100 psi, respectively. The desired minimum water storage heights at all tanks were 5 ft. The minimum and maximum allowable concentration limits at all demand nodes were set at 50 μg/L and 500 μg/L, respectively. The bulk and the wall rate coeffi-

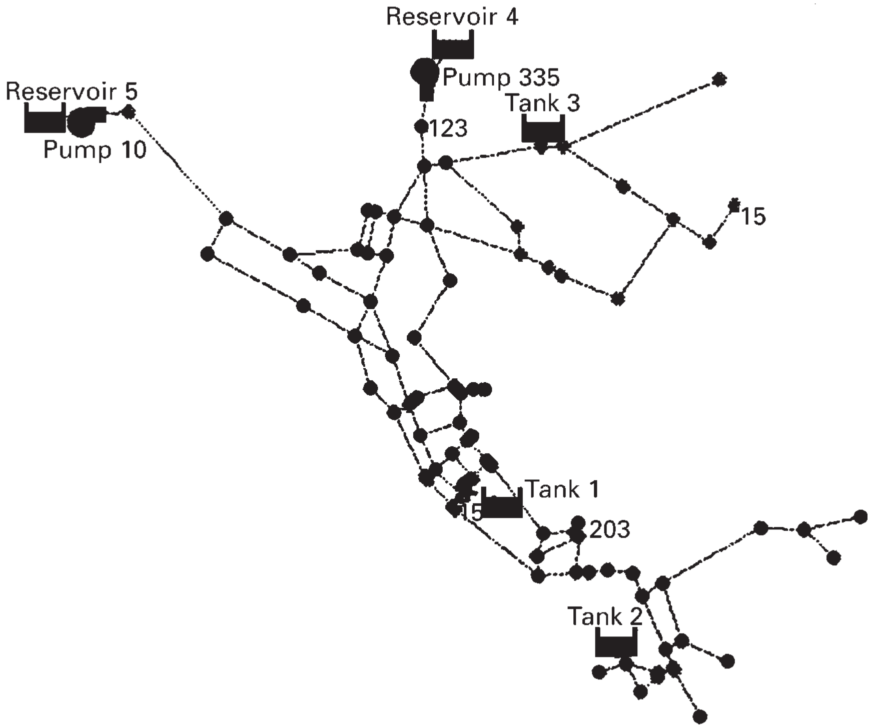


FIGURE 12.3 Water distribution system of North Marin Water District zone 1. (Rossman, 1994)

cients used in the simulation were −0.1/day and −1 ft/day, respectively. The simulation was conducted for a total of 12 days with 2-h time intervals where the values at the last day were used. The unit cost of energy was assumed to be constant at 0.07 $/kWh for all time periods, and the efficiency of both pumps was 0.75.

In this example, two different objective functions were used. The hypothetical example has shown that using a substance concentration violation term as a bound constraint gives better results than using it as the original objective function; therefore, case I which considers the minimization of substance concentration from its desired limits was not considered in this example.

The optimum pump operation schedules were determined using different concentration values at reservoirs 4 and 5. Table 12.1 summarizes the results obtained for cases II and III. Run number 1 in these tables shows the results obtained when the concentration values at both reservoirs are 500 μg/L. Run number 2 was based upon 200-μg/L and 400-μg/L concentrations at reservoirs 4 and 5, respectively. For run number 3, a 300-μg/L concentration was used at both reservoirs. Initially, both pumps were considered to be on for all time periods. Hence, the total pump operation time was 48 h and the total 24-h energy cost was $555.83 for all runs for

TABLE 12.1 Optimized Solutions Obtained by Using Different Solution Approaches for 500 μg/L of Concentration at Both Reservoirs

	Case II: minimize pump operation time		Case III: minimize pump operation cost	
	Mathematical programming	Simulated annealing	Mathematical programming	Simulated annealing
Original objective function	34.47	24.00	399.95	429.53
Concentration violation	0.00	0.00	0.00	0.00
Pressure violation	0.00	0.00	0.00	0.00
Tank water level violation	0.00	0.00	0.00	0.00
Operation time of pump 10, h	19.97	5.00	24.00	14.00
Operation time of pump 335, h	14.50	19.00	13.62	17.00
Total pump operation time, h	34.47	24.00	37.62	31.00
Total 24-h energy cost, $	401.06	433.63	399.95	429.53

both cases. The initial pump operation schedule resulted in a pressure violation of 5582.15 and there existed no violation of tank water surface height level for all runs for both cases. Run numbers 1 and 3 resulted in no concentration violation, whereas a violation of 10.92 existed for run number 2 for both cases.

Figures 12.4 and 12.5 show the optimal pump operation times of pumps 10 and 335 for runs 1, 2, and 3 for case II, respectively. Similarly, the optimum pump operation times of pumps 10 and 335 obtained for all runs for case III are shown in Figs. 12.6 and 12.7. For case III the optimized solutions resulted in pump 10 operating for all time periods for all runs.

Run number 1 gave the minimum pump operation time, which is the original objective function for case II. No concentration, pressure, or tank water level violation existed for this run. Note that this run used a 500-μg/L of concentration at both reservoirs. When the concentration values were reduced to 300-μg/L for run 3, the total pump operation time increased to 37.45 h, but still no violation of any penalty term existed. It should be noted that run 3 gave the minimum total 24-h energy cost of all three runs for case II. For run 2, concentrations of 200 μg/L and 400 μg/L were used at reservoirs 4 and 5, respectively. Concentration violations existed for this run.

For case III, when the total energy is considered as the original objective function, run 1 gave the best result. For the other two runs, there existed violations of the concentration and the total 24-h energy costs were higher. Even though the minimization of the total energy cost was the objective function for this case, the best result obtained from case III (run 1) had a higher value than the result obtained for case II (run 3), which considered the total pump operation as the original objective function.

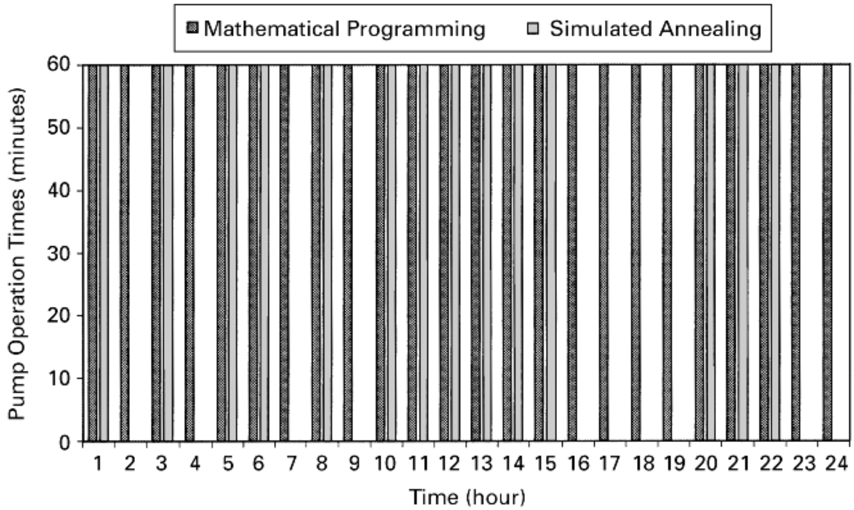


FIGURE 12.4 Optimal operation times of pump 10 used in NMWD for minimizing pump operation time.

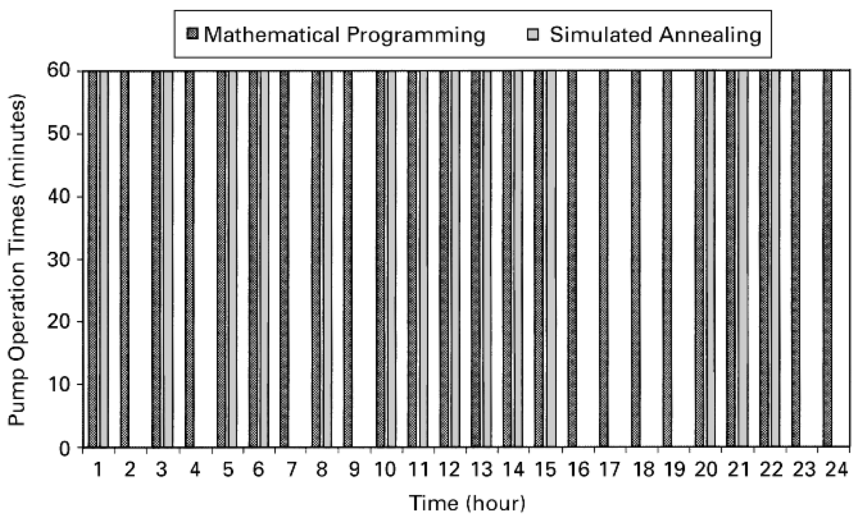


FIGURE 12.5 Optimal operation times of pump 335 used in NMWD for minimizing pump operation time.

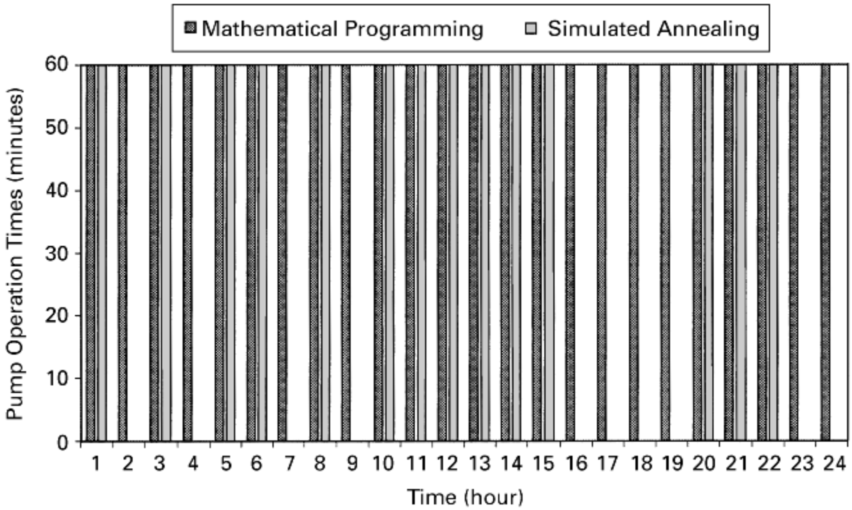


FIGURE 12.6 Optimal operation times of pump 10 used in NMWD for minimizing pump operation cost.

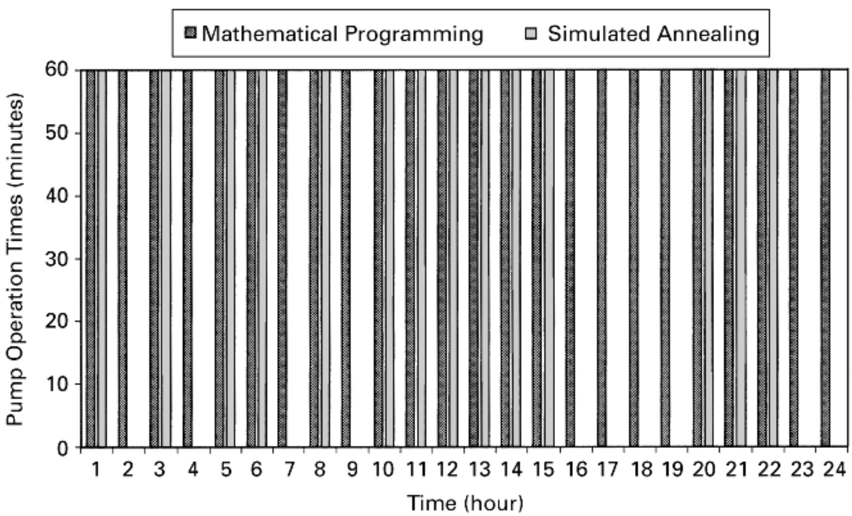


FIGURE 12.7 Optimal operation times of pump 335 used in NMWD for minimizing pump operation cost.

At the optimized solutions, pump 335 was closed and pump 10 operated for all time periods, 24 h, except run number 1 for case II. The reason for this is that pump 335 is larger than pump 10; hence, closing pump 335 for a certain time period has more effect than closing pump 10. The reduced gradients obtained for pump 335 is greater than the reduced gradients of pump 10, which forces GRG2 to close pump 335 rather than pump 10.

12.2.4 Summary and Conclusions—Mathematical Programming Approach

The optimal control methodology developed herein, which interfaces the EPANET simulation code with the GRG2 optimization code, provides an efficient methodology to determine the optimal operation schedules of the pumps to minimize the deviations of actual substance concentrations from the desired values, and to minimize the total pump operation times or the total energy cost while satisfying the hydraulic, the water quality, and the bound constraints. The resulting optimum control problem is a large-scale nonlinear programming problem, which cannot be solved by using existing nonlinear programming codes. The discrete time optimal control approach used in this research makes it possible to solve the resulting problem. The simulation code EPANET solves the hydraulic and the water quality constraints, and the optimization code GRG2 finds the optimum pump operation times by solving the unconstrained nonlinear problem. The development of the optimal control methodology and computer program to solve the optimal pump operation schedules in a water distribution system for water quality purposes is the unique contribution of this research.

This model development is the first application of the discrete time optimal control approach to find the optimal operations of pumps considering water quality. The use of different objective functions proved that considering the minimization of the total pump operation time or the total energy cost as the objective function and using the substance concentration violation term as a bound constraint gives better results than using the minimization of the substance concentration from the desired values as the objective function. This research has shown that using two different penalty function methods, the augmented lagrangian method and the bracket penalty method, resulted in similar results. The solutions obtained depend on the values of the initial penalty function parameters, which makes finding the global optimum more difficult. Global optimum solutions cannot be guaranteed, since convexity of the objective function cannot be proven. The best local optima can be used for practical purposes.

The use of the simplified method in calculation of the reduced gradients and the objective function provides reduced computational effort and gives similar results. It was found that if the daily pump operation schedules repeat themselves for a certain period of time, the system reaches steady state where the tank water levels at the beginning and at the end of the day are equal to each other. Hence, there is no need to consider a constraint that forces the tank water levels at the end of the day to be more or less equal to the initial state.

12.3 OPTIMAL PUMP OPERATION—SIMULATED ANNEALING APPROACH

12.3.1 Model Formulation

First we can define the optimization problem for a water distribution system with M pipes, K junction nodes, S storage nodes (tanks or reservoirs), and P pumps, which are to be operated for T time periods. The statement of the optimization problem is as follows: Minimize the energy cost [Eq. (12.3)] subject to:

Conservation of mass for each junction node [Eq. (12.4)]

Conservation of energy for each pipe [Eq. (12.5)]

Storage tank continuity [Eq. (12.6)]

Water quality constraint [Eq. (12.7)]

Nodal pressure head bounds [Eq. (12.9)]

Water storage tank height bounds [Eq. (12.10)]

Bounds on substance concentrations [Eq. (12.11)]

The above formulation results in a large-scale nonlinear programming problem with decision variables $(q_{ij})_t$, H_{kt}, y_{st}, D_{pt}, and $(C_{ij})_t$. In the proposed solution methodology, the decision variables are partitioned into two sets: (1) D_{pt}, the control (independent) variable, and (2) the remaining state (dependent) variables. The operation of a pump during a time period (on or off) is the control variable. The problem is formulated above as a discrete time optimal control problem.

12.3.2 Solution Methodology—Simulated Annealing Approach

The Method of Simulated Annealing. Simulated annealing is a combinatorial optimization method that uses the Metropolis algorithm to evaluate the acceptability of alternate arrangements and slowly converges to an optimum solution. The method does not require derivatives and has the flexibility to consider many different objective functions and constraints. Simulated annealing uses concepts from statistical thermodynamics and applies them to combinatorial optimization problems. Kirkpatrick et al. (1983) explains how the Metropolis algorithm was developed to provide the simulation of a system of atoms at a high temperature that slowly cools to its ground energy state. If an atom is given a small random change, there will be a change in system energy, ΔE. If $\Delta E \leq 0$, the new configuration is accepted. If $\Delta E > 0$, then the decision to change the system configuration is treated probabilistically. The probability that the new system is accepted is calculated by:

$$P(\Delta E) = \exp - \frac{\Delta E}{k_B T} \tag{12.19}$$

A random number evenly distributed between 0 and 1 is chosen. If the number is smaller than $P(\Delta E)$, then the new configuration is accepted; otherwise it is discarded and the old configuration is used to generate the next arrangement. The Metropolis algorithm simulates the random movement of atoms in a water bath at temperature T. By using Eq. (12.19), the system becomes a Boltzmann distribution.

Entropy is a measure of the variation of energy at a given temperature. At higher temperatures there is significant energy variation which reduces dramatically as the temperatures are lowered. This implies that the annealing process should not get "stuck" since transitions out of an energy state are always possible. Also, the process is a form of "adaptive divide-and-conquer." Gross features of the eventual solution appear at the higher temperatures with fine details developing at low temperatures (Kirkpatrick et al. 1983).

The requirements for applying simulated annealing to an engineering problem are (Dougherty and Marryott, 1991).

1. A concise representation of the configuration of the decision variables
2. A scalar cost function
3. A procedure for generating rearrangements of the system
4. A control parameter (T) and an annealing schedule
5. A criteria for termination

Configuration of Decision Variables. The decision variables consist of the operational schedule of each pump during discrete time periods. For this application, the 24-h day was divided into 1-h periods. The program finds the optimum pump schedule where each pump is either on or off during each time period. A configuration is a given schedule of pump operation, which determines the 1-h periods each pump is operating. EPANET is then run using that pump configuration. The program results are used to rate the performance of the simulation. The decision on whether to retain or change the pump operation configuration is based on the Metropolis algorithm as shown in Fig. 12.8, a pseudocode description of the adaptation of the simulated annealing algorithm to pump operation optimization.

Development of Cost Function. The cost function $g(C_i)$ for a given configuration C_i is used in place of energy of a system of atoms. The temperature T can be considered a control parameter, which has the same units as the cost function. The

```
INITIALIZE;
T = T0;
C = Ck0;
g = g(Ck0);
DO WHILE (stopping criterion not satisfied);
        DO WHILE (equilibrium not satisfied);
                PERTURB:
                Ck+1 = rearrangement of Ck;
                ·g = g(Ck+1) − g(Ck);
                IF · g · 0 THEN
                        ACCEPT:
                        C = Ck+1;
                ELSE
                        IF random(0,1) < exp(−·g/T) THEN
                                ACCEPT:
                                C = Ck+1;
                        ENDIF
                ENDIF
        ENDDO
        T = · T
ENDDO
```

FIGURE 12.8 Pseudocode description of the simulated annealing algorithm. T = psuedo temperature; C_k = pump operation schedule configuration; g = cost of C_k including pseudocosts for constraint violations. (After Dougherty and Marryott, 1991)

cost function should include the cost of pumping and penalties for violations of the storage tank level, pressure, and water quality bounds. The pumping cost is calculated for each period for each pump running and is a function of the pump flow rate, head, efficiency, and electricity rate during that period. The violations of the pressure and water quality bounds are calculated using penalty functions. By adjusting the penalty functions, the optimization problem can be adjusted to bias one constraint over another constraint. The cost of pumping is added to the "pseudocost" penalty functions:

$$\mathrm{COST}_{\mathrm{pumps}} = K \sum_{ij} \frac{Q_{ij}\,\mathrm{TDH}_{ij}\,P_j t_j}{\eta_{ij}} \tag{12.20}$$

where K = unit conversion factor
Q_{ij} = flow from pump i during period j
TDH_{ij} = operating point of total dynamic head for pump I during period j
P_j = power rate per kilowatt hour during period j
t_j = length of time that pump operates during period j
η_{ij} = wire to water efficiency of pump i during period j

Storage tanks have a minimum level to provide emergency fire flow storage. If the tanks are depleted below this level, then fire protection is compromised. A penalty term τ_1 has been developed to account for this constraint which is based on the constraint penalty terms developed by Brion (1990) and given by

$$\tau_1 = \sum_j \beta_s [\min(0, E_{sj} - E_{\min, s})]^2 \tag{12.21}$$

where τ_1 = penalty for violations of tank low water level bound
E_{sj} = water level for tank s during period j
$E_{\min, s}$ = lower bound or minimum level in tank s
β_s = penalty term for tank low water level constraint

Cohen (1982) stated that "optimization of a network over a limited horizon of, say, 24 hours has no meaning without the requirement of some periodicity in operation. A simple way to do that is to constrain the final states to be the same as the initial ones." A constraint τ_2 is developed to generate a cost if the tank levels do not return to their starting elevation.

$$\tau_2 = \sum_s \beta_{s2} [\min(0, 1 - |E_{s,1} - E_{s,j}|)]^2 \tag{12.22}$$

where $E_{s,1}$ = water level for tank s at beginning of simulation, during period 1
$E_{s,j}$ = water level for tank s at end of simulation, during final period j
β_{s2} = penalty term for beginning and ending tank level constraint for tank s

For a 24-h simulation the concept of returning tanks to their starting level is somewhat addressed by using a starting configuration where the pumps supply a volume of water equal to the sum of the nodal demands. Providing a total volume of pumped water equal to the total volume of the system demands will return the tanks to their original level exactly if there is only one tank. Even if pumping equals demand exactly, the tanks may not return to their original level if there are several tanks unless the tanks all start full. This is because one tank may "supply" water to another tank during the simulation and, hence, the total volume stored will be the same but shifted from one tank to another.

The Cohen condition may be unnecessarily restrictive. The second example involved modifying pump operation for a 24-h period that was repeated for 12 days to allow the water quality variations to overcome the initial conditions. It was observed that the tanks exhibited periodic behavior, adjusting themselves until the pumps were supplying a quantity equal to the demands (see Fig. 12.9). If the pumps operated for longer periods, the tanks remained closer to full and the pump heads moved to the left of the system curve reducing the flow. If the pumps operated for shorter periods, the tank levels lowered until the flows increased to meet the demands. By running the simulations over several days, the pump operation can be scheduled to optimize efficiency and perhaps reduce cost.

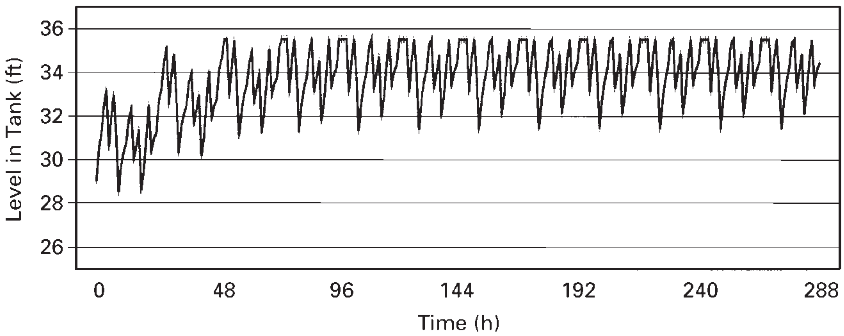


FIGURE 12.9 Periodic behavior.

A water distribution system needs to deliver water to its customers at sufficient pressure to service the water system customers but at a pressure that will not damage water systems or customers facilities. The Uniform Plumbing Code sets the normal pressure range as 15 to 80 psi (IAPMO, 1994). While a city may have a pressure range in a zone from 40 to 80 psi with a 20 psi residual during a fire flow (Malcolm Pirnie, 1996), the system operation needs to operate between two extreme pressures, $p_{\min}$ and $p_{\max}$.

The penalty functions for minimum and maximum pressure bounds at each node are

$$\tau_3 = \sum_{k,j} \gamma_1 [\min(0, p_{k,j} - p_{\min})]^2 \tag{12.23}$$

$$\tau_4 = \sum_{k,j} \gamma_2 [\min(0, p_{\max} - p_{k,j})]^2 \tag{12.24}$$

where $p_{\min}$ = minimum system pressure bound
$p_{\max}$ = maximum system pressure bound
p_{kj} = pressure at node k during time period j
γ_1, γ_2 = penalty terms for minimum and maximum pressure violations

The penalty terms for minimum and maximum free chlorine concentration are

$$\tau_5 = \sum_{k,j} \gamma_3 [\min(0, c_{k,j} - c_{\min})]^n \tag{12.25}$$

$$\tau_6 = \sum_{k,j} \gamma_4 [\min(0, c_{\max} - c_{k,j})]^n \tag{12.26}$$

where $c_{\min}$ = minimum free chlorine concentration
$c_{\max}$ = maximum free chlorine concentration

c_{kj} = chlorine residual concentration at node k during time period j
β_3 = penalty term for minimum chlorine residual pressure bound violations
γ_4 = penalty term for maximum chlorine residual pressure bound violations

The value of n will usually be 2. In cases where the lower concentration bound is more important, a value of $n < 1$ will place a higher penalty on minimum free chlorine violations. There will also be a term for the amount of chlorine used. The goal will be to meet the chlorine residual bounds while using the smallest amount of chlorine. Not only will the operational cost be decreased, but the creation of total trihalomethane (TTHM) will also be reduced.

Development of Software. The basic strategy was to use EPANET as the simulator and to develop new functions around EPANET that calculate penalties or costs, generate pump operation configurations, and evaluate the acceptance of potential configurations using the Metropolis algorithm. The EPANET program was developed for a single simulation as described by the flowchart in Fig. 12.10. In order to carry out the numerous simulations required for annealing, several modifications were required. New functions were written in the C programming language to carry out penalty calculations, generate pump operation configurations, modify the temperature according to the annealing schedule, and analyze each result using the Metropolis algorithm. Special routines were developed to collect information useful in understanding the progress of the annealing method.

Figure 12.11 is a flowchart that shows the operation of the annealing program. The chart shows that the number of loops and iterations per loop are set before the process begins. Each loop carries out a number of iterations at a given temperature. If a pump configuration is held for 50 times, it is considered to be the optimal configuration. If the program finishes normally, the final pump configuration is considered to be optimal. Frustration between constraints in a water distribution system provides many good configurations to recommend to the system operators. There may be pump configurations that provide better solutions than the final solution. A routine was added to save the 20 best configurations for examination by the user.

12.3.3 Applications—Simulated Annealing Approach

Austin, Texas, Example. Brion (1990) and Brion and Mays (1991) applied their model to the City of Austin Northwest B Pressure Zone, which consisted of 3 pumps, 126 pipes of approximately 38.5 mi in total length, 98 junction nodes of which 5 were pressure "watch points," and one storage tank (Fig. 12.12). The 24-h simulation was broken down into twelve 2-h time periods. The system served about 31,000 residents located in about 32,000 acres.

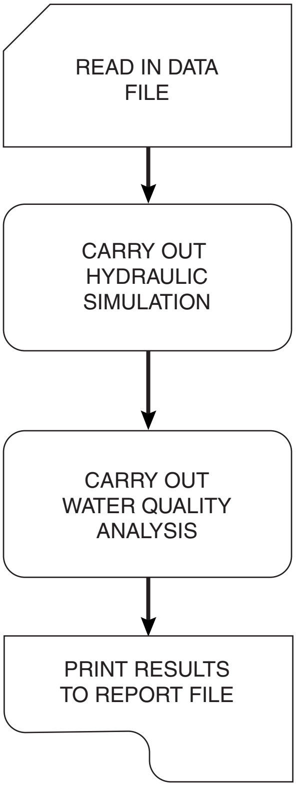


FIGURE 12.10 EPANET Flowchart.

The pipes, junctions, pumps, tanks, and diurnal flow distribution can be found in Goldman (1998) or Brion (1990). Brion and Mays (1991) collected information on the actual pumping times used on September 29, 1988, resulted in a pumping cost of $231 during the day. Using NLP with an augmented lagrangian description of the pressure and tank constraints and linking the GRG2 program with KYPIPE, the cost was reduced to $219, a savings of 5.2 percent. The resulting pump operation schedule is shown in Fig. 12.13.

The series of trial pump configurations was generated by starting with the last trial configuration, and then using a random number generator to choose a pump and a period, and switching the operation. If the chosen pump was operating during the chosen period in the previous trial, it is turned off. If it was off during the chosen period, it is changed to operating. The stopping condition was having a pump schedule withstand 20 challenges from new pump schedules. Because of the high level of frustration, with many pump schedules having very similar costs, the ending condition was not met.

The best result was $221.38, which occurred in trial 67 of loop 3. This resulted in a savings of 4.1 percent. The annealing method did not isolate one solution because of the high level of frustration. The program found 173 solutions, which met the pressure and tank constraints and resulted in a cost less than $235; 50 pump schedules had a cost less than $230. The feasible solutions were reviewed to find the least costly pump operation schedule that met the pressure and tank constraints. The optimum pumping schedule found using simulated annealing and NLP is shown in Fig. 12.13.

A review of Fig. 12.13 shows the NLP solution had a 15-min sliver of operation occurring for pump 7 during the period between hours 18 and 20, which is undesirable for operating pumps and most likely would need to be adjusted by the operator. Short-period results can be expected from the NLP method, which is optimizing the duration of pump operation for a 2-h time period. Simulated annealing resulted in a minimum 2-h pumping period since it is a combinatorial optimization method where the pump has to be on or off during the entire period; thus the minimum operating time is 2 h.

The tank trajectory for the simulated annealing optimization pump schedule is shown in Fig. 12.14. The minimum height in the tank was 1 ft. A penalty term was used to return the tank level to the original tank level, within a tolerance of 1 ft.

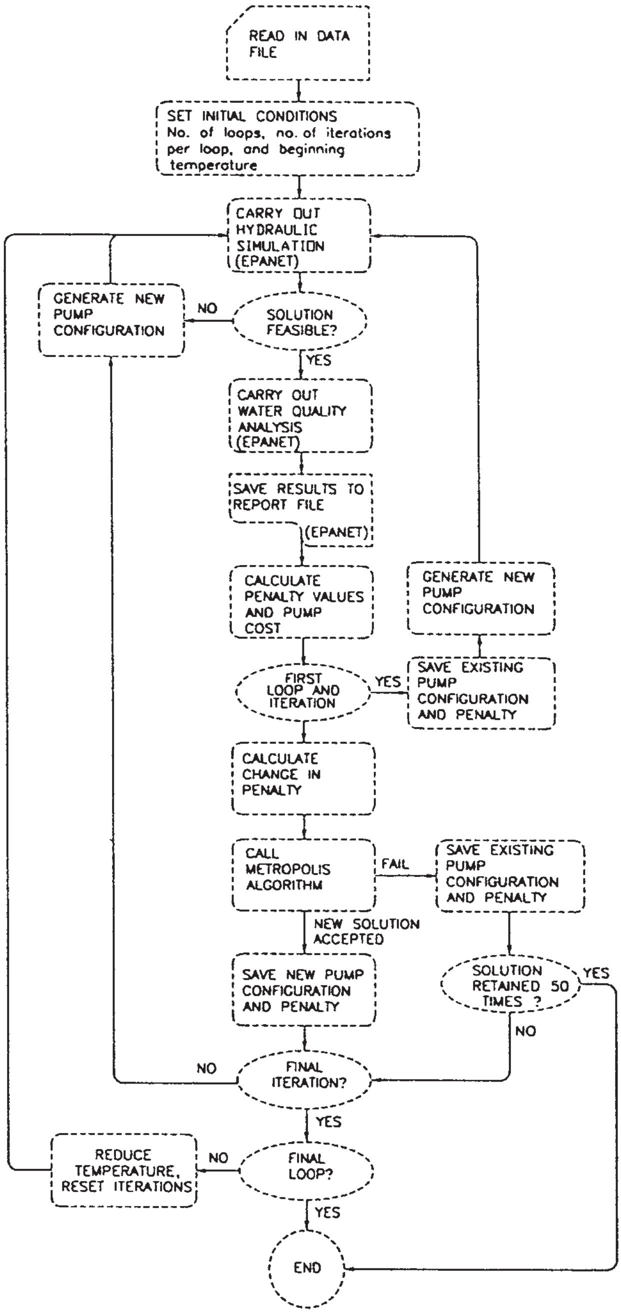


FIGURE 12.11 Simulated annealing flowchart.

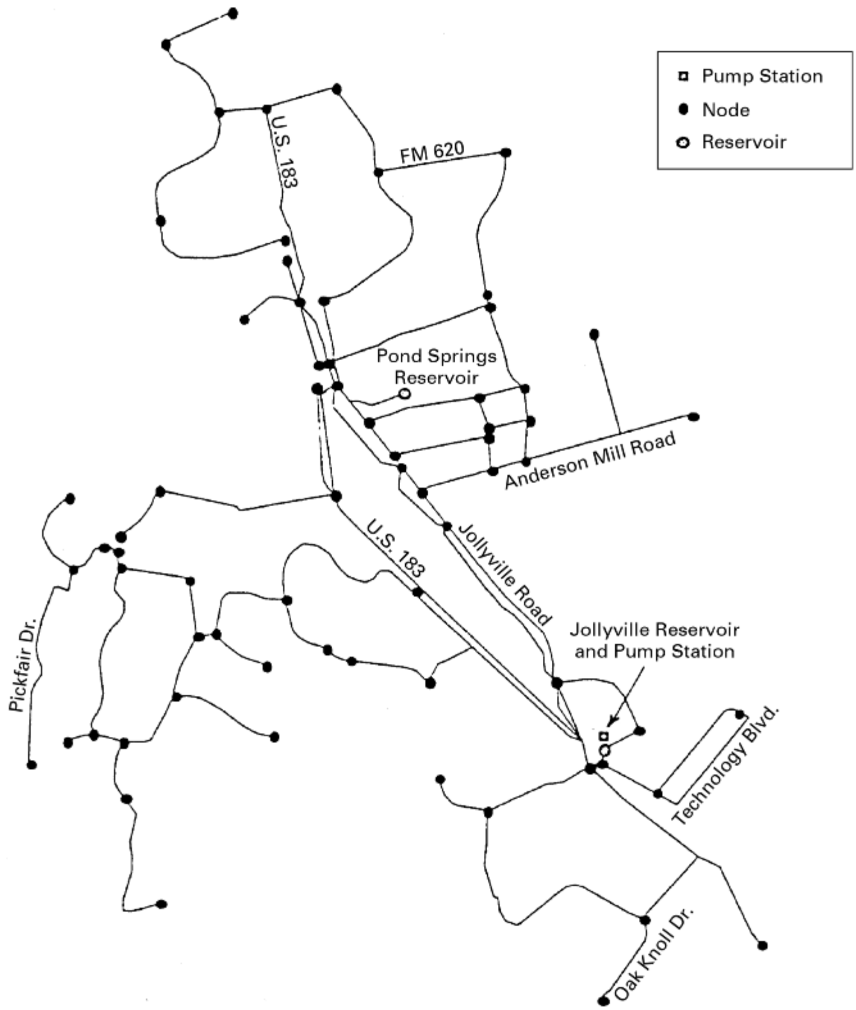


FIGURE 12.12 Water distribution system for City of Austin Northwest B Pressure Zone. (Brion and Mays, 1991)

North Marin Example. The piping system consists of 125 pipes, 102 nodes, 3 tanks, and 2 pumps as shown in Fig. 12.3. Twenty-four 1-h periods were considered with the pumping schedule repeated every 24 h. The goal of the optimization is to find the optimal operation of two pumps with differing capacity characteristics pumping from two different sources of water. Two strategies were studied. In the first, both sources had free chlorine concentrations of 0.5 mg/L, and the minimum free chlorine concentration at any demand node was taken as 0.05 mg/L.

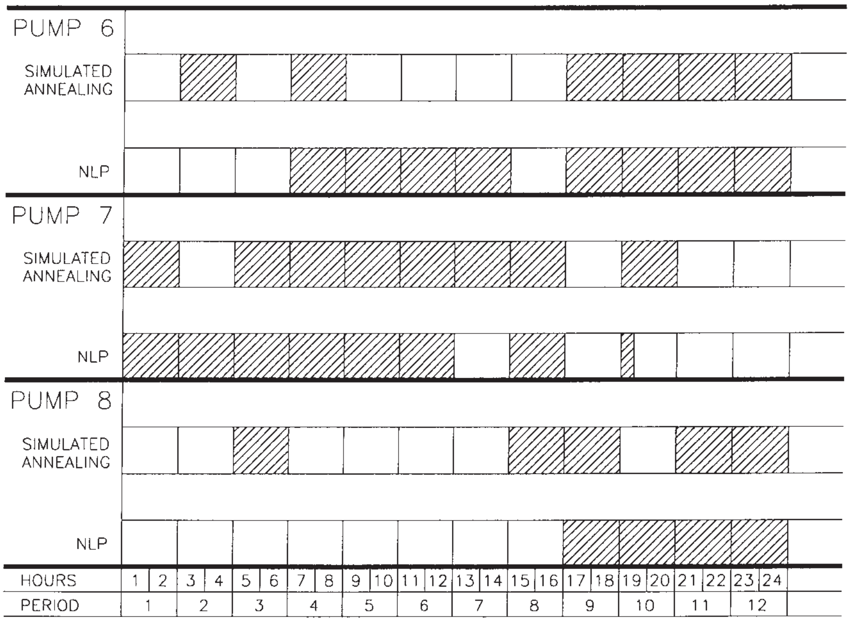


FIGURE 12.13 Austin example: simulated annealing and NLP pump operation schedules.

Pressures at demand nodes were to be kept between 20 and 100 psi. Global bulk decay was taken as 0.1/day, and global wall decay was taken as 1.0 ft/day. The second strategy had two different chlorine residuals. The free chlorine for the larger pump was 0.4 mg/L and for the smaller pump was 0.2 mg/L. The minimum free chlorine concentration was taken as 0.05 mg/L.

It takes many days for initial chemical concentrations to be overcome and repeatable steady-state behavior to be observed. This periodic behavior was discussed previously and was reported for water quality by Boccelli et al. (1998). Figure 12.15 shows the periodic behavior of the chlorine concentration in tank 1 of this example. To reach steady-state conditions, each iteration was run for 288 h. This resulted in each iteration taking 10 s on average. The results for the last 24 h were used to calculate pumping cost and constraint violation penalties.

The two optimization techniques, simulated annealing and the mathematical programming approach (NLP using reduced gradients with lagrangian multipliers for constraints and GRG2), obtained similar results. For the case where the source chlorine concentration for both sources was 0.5 mg/L, there were no pressure or chlorine violations and the minimum cost found by NLP was $399.95 per day versus $429.53 for simulated annealing. For the case where water qualities for the two sources were 0.2 and 0.4 mg/L, the NLP cost was $408.39 versus $407.66 for simulated annealing. There were no pressure violations, but the total chlorine violations measured in

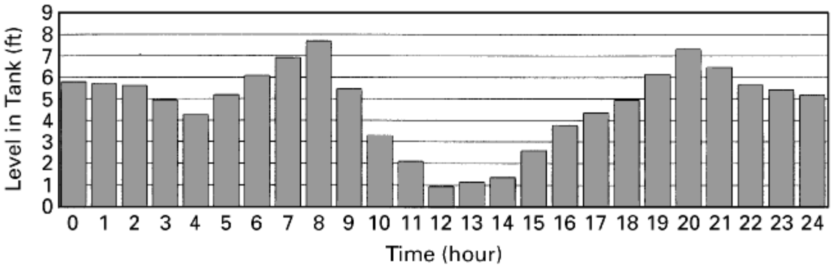


FIGURE 12.14 Austin example: simulated annealing and tank levels.

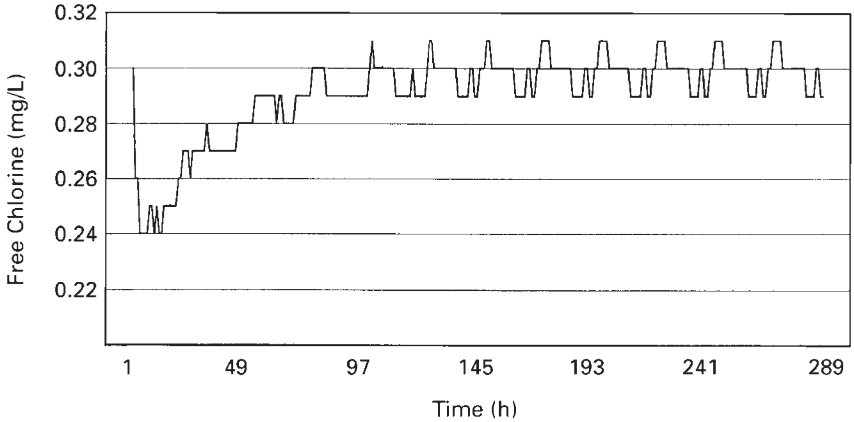


FIGURE 12.15 North Marin example: chlorine concentration, tank 1.

total chlorine excursions for all nodes during the 24 h was 69.76 μg/L for NLP solution as compared to 5.89 μg/L for simulated annealing solution.

There were distinct differences in computation times to arrive at a solution. The mathematical programming model was solved on a PC with a Pentium processor, while the simulated annealing was solved on a UNIX computer cluster consisting of five IBM computers, two RS/6000 Model 390 interactive servers, and three RS/6000 Model 590 mathematical computational processing. The Arizona State University machines run the IBM AIX operating system, which is IBM's version of UNIX. In both instances the time to run a 12-day-long EPANET simulation was about 10 s. The NLP optimization required about one-third of the iterations as the simulated annealing optimization, although there was no attempt to streamline the simulated annealing and it turned out that the best solutions occurred before the final loop started. It is believed that even with some improvement in efficiency the simulated annealing would require at least twice the iterations that NLP requires.

Simulated annealing has been shown to be more flexible and adaptable than NLP optimization. The requirement that many parts of the distribution system such as node quality, pressures, tank levels, and pump operation must have derivatives with respect to pump operation time during each period restricts the flexibility of the method to accept changes to the system or consider a variable pump efficiency. On the other hand, simulated annealing can accept changes to the distribution system since it is concerned with cost calculation and deriving a Markov chain of operation schedules.

The NLP method's efficiency in finding optimum solutions appears to be very sensitive to the lagrangian coefficients used in the outer loop of the optimization. Also, the significant differences in the pump capacities resulted in much more rapid changes in pressures and water quality than changes in operation of the larger pump. This resulted in the preferred changes in the larger pump, which reduced its operating time until any more restrictions would have resulted in an unbalanced solution.

A fundamental *tenant* in using NLP optimization linked to a simulator such as EPANET is the implicit function theorem (Brion, 1991), which states that if D_t is the set of pump durations during periods D_t (the control variable), then $H(D_t)$, the node pressure matrix, and $E(D_t)$, the tank level matrix (the dependant or state values), exist in the neighborhood of D^*, the set of optimum pump operations. This implies that the solutions exist in the neighborhood. However, in running the simulated annealing routines many unbalanced nonfeasible solutions existed in the vicinity of the optimum solution. This becomes important if continuity is assumed in the vicinity of the NLP optimums. This problem does not reveal itself in NLP that includes water quality since the water quality portion of EPANET does not include the solving of equations that could be unbalanced.

Another important issue is global versus local optimum and the concept of frustration. A total of 36 pump operation configurations were found by simulated annealing that have total water quality violations less than 5.89 μg/L and energy costs less than \$420 per day, where the water quality violations are the sum of the violations at all demand nodes for the last day of the simulation. There are many nearly equal optimal pump operation schedules that have nearly the same penalty values. This is most likely the reason that simulated annealing found optimum solutions in the next-to-the-last loop. The actual pump operation schedules developed by the two optimization techniques are shown in Fig. 12.16 for the case where both sources had chlorine concentration of 0.5 mg/L.

As previously mentioned, the NLP solution where the sources had identical 0.5-mg/L concentrations resulted in a cost of \$399.95 versus \$429.53 which appears to be a preferred solution. However, a review of Figs. 12.15 and 12.16 reveals a weakness with the NLP optimization. NLP is optimizing the amount of pumping time for a given period. This can result in several periods where the pumps operate only part of the time. This occurs during periods 2, 3, 8, 10, 12, and 14. In order to operate pumps more continuously, there would need to be some

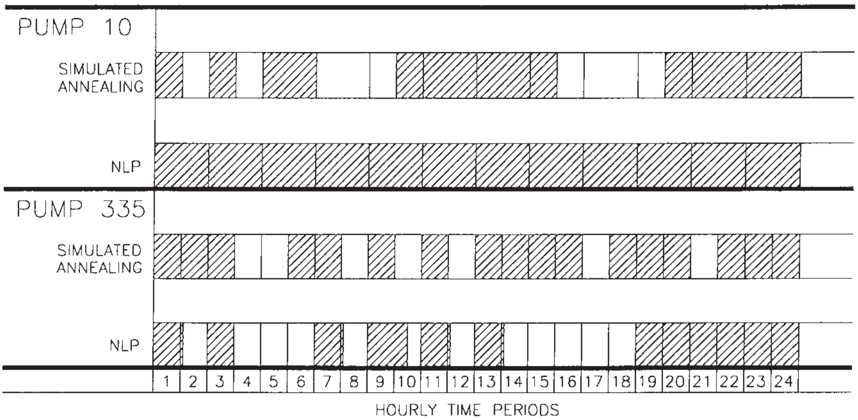


FIGURE 12.16 System 1: simulated annealing and NLP pump operation schedules. 0.5 mg/L free chlorine at both sources and minimum free chlorine at nodes as 0.05 mg/L.

adjustment of the pumping schedule shown in Fig. 12.15, which could impact the ultimate cost of pumping.

12.3.4 Summary and Conclusions—Simulated Annealing Approach

Simulated annealing has been successfully linked with EPANET to optimize the operation of a water distribution system for hydraulic behavior and water quality. The programs developed have successfully considered a variety of optimization objectives. This sharply differs from other optimization techniques which require derivatives and are restricted to one objective function and may need to be modified for each data set. The annealing programming uses the global variables and dynamically allocated memory utilized by EPANET to make the computer memory requirements match the size and scope of the distribution system. The programming utilizes the data set read during the first iteration of EPANET to size the vectors required to store the data. Each implementation imports the data set and carries out the optimization.

The broadness of the scope of applications and the flexibility to assimilate changes to the distribution system character allow the programming to be adapted to supervisory control and data acquisition (SCADA) systems. For example, let us suppose that a fire event has depleted the storage of water in a system reservoir and it was desired to find the optimal operation of pumps to meet demands and optimally recover storage in the tanks. Simulated annealing could suggest several pump operation schedules that could meet the needs of the system.

The water quality example showed the ability of simulated annealing to suggest pumping schedules to meet water quality objectives. The objective of the example was to obtain desired chlorine residuals in a system served by two water sources using an efficient pump operation. Simulated annealing provides many pump schedules that meet the needs of the system. If conditions change, the basic data file can be modified, perhaps automatically by the SCADA system and a new schedule could be found. If a coliform test sample came back positive, the pump operation schedule can be found to restore chlorine levels. With a minor modification to the program, a variation in chlorine concentration at the source could be considered as part of the effort to find an operating schedule that restored the system to water quality standards as quickly as possible. If flushing is carried out, the only change would be increasing the demand at the node where the flushing occurred in the data set. Simulated annealing could suggest a pumping schedule that would meet the increased demands within the system pressure constraints. Again, with a minor revision, simulated annealing could suggest the optimum flushing flow or could evaluate the efficiency of locating the flushing at alternate locations, again, generating a group of pumping schedules that would meet the problem conditions.

12.4 OPTIMAL OPERATION OF CHLORINE BOOSTER STATIONS

12.4.1 Model Formulation

First we can define the optimization problem for a water distribution system with M pipes, N nodes, K junction nodes, S storage nodes (tanks or reservoirs), and B booster pumps, which are to be operated for T time periods. The statement of the optimization problem is as follows:

Minimize the amount of chlorine used.

$$\text{Minimize } Z = \sum_{b}^{K} \sum_{t} q_{bt} C_b \Delta t_t \qquad \text{g/day} \tag{12.27}$$

where q_{bt} = flow of booster chlorination station b during time period t (gal/h), C_b = concentration of chlorine being injected into the system by station b (mg/L), Δt_i = length of time period t (h), K = (3.785 L/gal)/(1000 mg/g), and $\sum \Delta t_i = 24$ h, subject to

Conservation of mass for each junction node [Eq. (12.4)]

Conservation of energy for each pipe [Eq. (12.5)]

Water quality constraint [Eq. (12.7)]

Bounds on substance concentrations [Eq. (12.11)]

The above model formulation results in a large-scale nonlinear programming problem with decision variables $(q_{ij})_t$, H_{kt}, y_{st}, q_{pt}, and $(C_{ij})_t$. In the proposed solution methodology, the decision variables are partitioned into two sets: control (independent) and state (dependent) variables. The operation of the chlorine booster station(s) during a time period is the control variable.

The penalty terms for minimum and maximum free chlorine concentration are Eqs. (12.25) and (12.26). The goal is to meet the chlorine residual bonds using the minimum amount of chlorine. This not only decreases the operational costs, but also reduces the amount of trihalomethanes. The solution methodology used is the simulated annealing approach described in Sec. 12.3.2.

12.4.2 Application

The network shown in Fig. 12.17 (Cherry Hill–Brushy Plains, South Central Connecticut Regional Water Authority) is used to illustrate the simulated annealing approach to solve the model formulated in Section 12.4.1 to determine the optimal operation of chlorine booster stations. Boccelli et al. (1998) used this same network to illustrate their methodology. They used linear supposition and linear programming models generating impulse response coefficients using EPANET. This network has six chlorine booster stations designated as 101, 102, … , 106 on the layout in Fig. 12.17. The booster stations were modeled as sources with a fixed concentration of 2000 mg/L and varying flow. This arrangement is similar to a hypochlorinator with a tank holding a chlorine solution at 2000 m/L and a metering pump. The flow rates were in the range of a metering pump, and although they add a small amount of water to the system, they do not impact the system hydraulics.

The hydraulic conditions of the system are constant, with demands and flow rates repeated every 24 h. Simulations were for 7 days (168 h) so that the system could reach steady state, and results of the last day were polled for violations. Because the hydraulics were fixed for this example, the pump and demand characteristics were inputted as pattern multipliers. EPANET designates chlorine booster pumps as sources. Using the [SOURCES] section of the data file, the concentration of the sources is set as 2000 mg/L. The CHG_OP.C subprogram changes the source flow multipliers to match flow rates to the pump operation schedule testing during the trial. A zero flow rate is set as a 0.0000001 multiplier since the EPANET program does not accept a source with zero flow.

Two applications are explained, with the first consisting of one booster station located at the pumping station node 101. The control variable is the flow rate for each of four 6-h periods. Ten flow rates were considered by using a range of 0.0 to 2.0 gallons per minute (gpm) for each of the periods with increments of 0.2 gpm. The chlorine booster station schedule is given in Table 12.2, and the chlorine violations are given in Table 12.3. The one booster station application requires 2813 g/day of chlorine.

The second application considers the six booster station locations, having 4.74×10^{18} possible combinations. This application also set the flow rate for each booster pump for each of the four 6-h time periods. The chlorine booster station

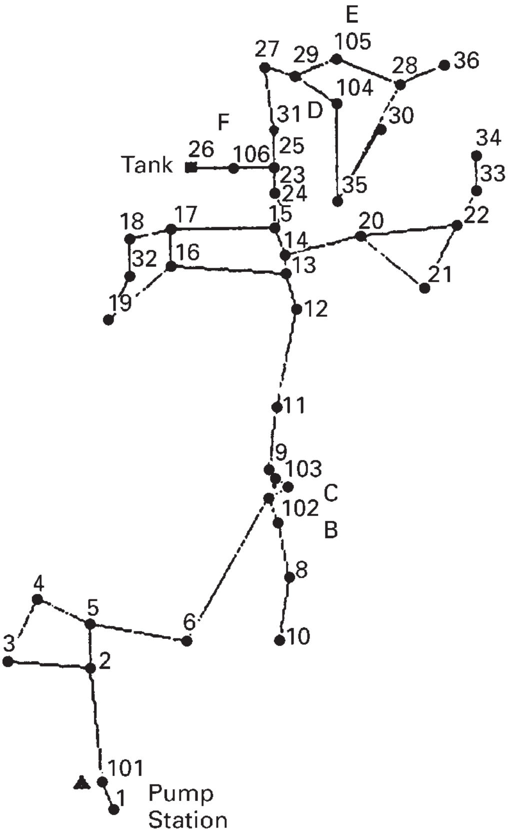


FIGURE FIGURE 12.17 System layout.

TABLE 12.2 One Chlorine Booster Application: 24-h Pumping Schedule

Zoom loop	6-h 1	2	3	4	Total chlorine, g/day
1	0.6 gram (4542 mg/min)	0 (0)	0.4 gpm (3028 mg/min)	0 (0)	2725
2	0.56 gpm (4239 mg/min)	0.04 gpm (330 mg/min)	0.44 gpm (3331 mg/min)	0.008 gpm (61 mg/min)	2866
3	0.584 gpm (4421 mg/min)	0 (0)	0.44 gpm (3331 mg/min)	0.008 gpm (61mg/min)	2813

schedule is given in Table 12.4, and the chlorine violations are given in Table 12.5. The chlorine usage using six booster stations is 1504 g/day, which is considerably less (a 46 percent reduction) than for one booster station.

12.4.3 Summary and Conclusions—Chlorine Booster Station Operation

The broadness of the scope of applications and the flexibility to assimilate changes to the distribution system character allows the methodology to be applied to SCADA systems. For example, let us suppose that a fire event has depleted the

TABLE 12.3 One Chlorine Booster Application: Data Points outside of 0.2- to 4.0-mg/L Range

Node	Time period, h	Chlorine concentration, mg/L
23	146	6.1
27	149	5.3

TABLE 12.4 Six Chlorine Booster Application: 24-h Pumping Schedule

Booster	6-h period 1	6-h period 2	6-h period 3	6-h period 4	Chlorine used, g/day
101	0.0816 gpm (617.7 mg/min)	0.0120 gpm (90.8 mg/min)	0.0960 gpm (726.7 mg/min)	0.072 gpm (545.0 mg/min)	712
102	0.0032 gpm (24.2 mg/min)	0.0120 gpm (90.8 mg/min)	0.0032 gpm (24.2 mg.min)	0.0012 gpm (90.8 mg/min)	83
103	0.0280 gpm (212.0 mg/min)	0.0280 gpm (212.0 mg/min)	0.0432 gpm (327.0 mg/min)	0.0032 gpm (24.2 mg/min)	279
104	0.0032 gpm (24.2 mg/min)	0.0032 gpm (24.2 mg/min)	0.0032 gpm (24.2 mg/min)	0.0032 gpm (24.2 mg/min)	35
105	0.0032 gpm (24.2 mg/min)	0.0032 gpm (24.2 mg/min)	0.0032 gpm (24.2 mg/min)	0.0120 gpm (90.8 mg/min)	59
106	0.0 gpm (0.0 mg/min)	0.060 gpm (454.2 mg/min)	0.0032 gpm (24.2 mg/min)	0.060 gpm (454.2 mg/min)	336
Total					1504

TABLE 12.5 Six Chlorine Boosters: Data Points outside of 0.2- to 4.0-mg/L Range

Node	Time period, h	Chlorine concentration, mg/L
36	144	15.6
36	145	6.6
36	146	5.8
36	147	11.4
36	148	6.88
34	154	0.187
34	155	0.180
8	159	4.07
8	160	4.03
34	164	0.160
34	165	0.188
36	167	5.3

storage of water in a system reservoir and it is desired to find the best operation of pumps to meet demands and recover the storage tanks. Simulated annealing could easily suggest several pump operation schedules that could meet the needs of the system. If conditions change (such as new demands or closed pipes), the basic data file can be modified, perhaps automatically by the SCADA system, and a new schedule could be computed quite easily. The method described by Boccelli et al. (1998) requires derivations of a new set of impulse functions, which can be quite cumbersome and timely.

12.5 CHALLENGES FOR THE FUTURE

Some of the challenges in model development for the future are listed here.

- In the past few years there have been reported in the literature (1) several models for the optimal design of water distribution systems, (2) several models for reliability analysis, (3) a few models for the optimal operation of water distribution systems, and (4) a few models for optimal reliability and availability based design. See Mays (1989, 2000) for a review of these models. None of these models have adequately interfaced the design, operation, and reliability aspects together.
- There is currently no universally accepted definition or measure of the reliability of water distribution systems. Many models have been reported in the litera-

ture for the reliability analysis of water distribution systems. Unfortunately the water distribution industry (engineering community) has not incorporated these methodologies into practice. There are many reasons for this, two of which are that engineers are not familiar with the methodologies and the proper software has not been developed to make the application an easy task.

- An optimization model is needed that would include objective functions such as minimizing energy costs, minimizing water age, minimizing deviations of actual concentrations of a constituent from the desired concentration values, minimization of total pump operation times, and maximizing reliability and availability.
- The model needs to be developed so that it is flexible enough that the user can select the objective function(s), but yet is interfaced with a widely used simulation model such as the EPANET model. A model that selects the optimal pump operation that minimizes water age during operation has never been attempted.
- Utilities have several strategies available to correct problems with excessive loss of disinfectants in distribution networks including: (1) switching to a more stable secondary disinfecting chemical, such as chloramines; (2) pipe replacement, flushing, and relining; (3) making operational changes to reduce the time that water spends in the system; and (4) using booster disinfection. The model suggested above needs to incorporate these strategies in the optimization framework.
- Very few of the previous modeling efforts for optimization have interfaced the design and reliability aspects in a usable form for the practicing community, and none have interfaced the operation and reliability aspects in any kind of form.
- Almost all research efforts to develop optimization models for design and operation have failed to use and rely upon the state-of-the-art hydraulic and water quality simulation models such as EPANET. Many have simplified the hydraulic and water quality processes to fit the optimization methodology selected.
- The biggest challenge of all is to convince the water distribution community that the types of models that have been suggested in the literature and expanded upon above, the operation models and the reliability and availability models, are valid tools that can be used to better design and operate their systems.

12.6 REFERENCES

Bao, Y. X., and L. W. Mays, "New Methodology for Optimization of Freshwater Inflows to Estuaries," *Journal of Water Resources Planning and Management,* ASCE, 120(2):199–217, 1994a.

Bao, Y. X., and L. W. Mays, "Optimization of Freshwater Inflows to the Lavaca-Tres Palacios, Texas, Estuary," *Journal of Water Resources Planning and Management,* ASCE, 120(2):218–236, 1994b.

Boccelli, D. L., M. E. Tryby, J. G. Uber, L. A. Rossman, M. L. Zierolf, and M. M. Polycarpou, "Optimal Scheduling of Booster Disinfection in Water Distribution Systems," *Journal of Water Resources and Planning and Management,* ASCE, 124(2):99–111, 1998.

Boulos, P. F., T. Altman, P. Jarrige, and F. Collevati, "Discrete Simulation Approach for Network-Water-Quality Models," *Journal of Water Resources Planning and Management,* ASCE, 121(1):49–60, 1995.

Brion, L. M. "Methodology for Optimal Operation of Pumping Stations in Water Distribution Systems," Ph.D. thesis, University of Texas, Austin, 1990.

Brion, L. M., and L. W. Mays, "Methodology for Optimal Operation of Pumping Stations in Water Distribution Systems," *Journal Hydraulic Engineering,* ASCE, 117(11):1551–1569, 1991.

Brook, A., D. Kendrick, and A. Meeraus, *GAMS: A User's Guide,* The Scientific Press, Redwood City, Calif., 1992.

Chase, D., and L. Ormsbee, "Optimal Pump Operation of Water Distribution Systems with Multiple Storage Tanks," *Proceedings of the American Water Works Association Comput. Specialty Conference, American Water Works Association,* 1989, pp. 205–214.

Chase, D., and L. Ormsbee, "An Alternate Formulation of Time as a Decision Variable to Facilitate Real-Time Operation of Water Supply Systems," *Proceedings of the 18th Annual Conference of the ASCE Water Resources Planning and Management Division,* ASCE, 1991, pp. 923–927.

Clark, R. M., "Development of Water Quality Models," in M. H. Chandhry and L. W. Mays (eds.), *Computer Modeling of Free-Surface and Pressurized Flows,* Kluwer Academic Publishers, Dordrecht, The Netherlands, 1994a, pp. 553–580.

Clark, R. M., "Applying Water Quality Models," in M. H. Chandhry and L. W. Mays (eds.), *Computer Modeling of Free-Surface and Pressurized Flows,* Kluwer Academic Publishers, Dordecht, The Netherlands, 1994b, pp. 581–612.

Clark, R. M., J. Q. Adams, and R. M. Miltner, "Cost and Performance Modeling for Regulatory Decision Making," *Water,* 28(3):20–27, 1987.

Clark, R. M., W. M. Grayman, J. A. Goodrich, R. A. Deininger, and K. Skov. "Measuring and Modeling Chlorine Propagation in Water Distribution Systems," *Journal of Water Resources Planning and Management,* ASCE 120(6):871–887, 1994.

Clark, R. M., L. A. Rossman, and L. J. Wymer, "Modeling Distribution System Water Quality: Regulatory Implications," *Journal of Water Resources Planning and Management,* 121(6):423–428, 1995.

Cohen, G., "Optimal Control of Water Supply Networks," Chap. 8 in S. G. Tzafestas (ed.), *Optimization and Control of Dynamic Operational Research Models,* Vol. 4, North-Holland Publishing Company, Amsterdam, 1982, pp. 251–276.

Coulbeck, B., M. Brdys, and C. H. Orr, "Exact Incorporation of Maximum Demand Charges in Optimal Scheduling of Water Distribution Systems," *IFAC 10th Triennial World Congress,* Munich, Federal Republic of Germany, 1987, pp. 359–364.

Coulbeck, B., M. Brdys, C. Orr, and J. Rance, "A Hierarchical Approach to Optimized Control of Water Distribution Systems: Part I. Decomposition," *Journal of Optimal Control Applications and Methods,* 9(1):51–61, 1988a.

Coulbeck, B., M. Brdys, C. Orr, and J. Rance, "A Hierarchical Approach to Optimized Control of Water Distribution Systems: Part II. Lower Level Algorithm," *Journal of Optimal Control Applications and Methods,* 9(2):109–126, 1988b.

Coulbeck, B., and C. H. Orr, "Optimized Pump Scheduling for Water Supply Systems," *IFAC 3rd Symposium, Control of Distributed Parameter Systems,* Toulouse, France, 1982.

Dougherty, D. E., and R. A. Marryott, "Optimal Groundwater Management. 1. Simulated Annealing," *Water Resources Research,* 27(10):2493–2508, 1991.

Elton, A., L. F. Brammer, and N. S. Tansley, "Water Quality Modeling in Distribution Networks," *Journal AWWA,* 87(7):44–52, 1995.

Fletcher, R., "An Ideal Penalty Function for Constrained Optimization," *Journal of Institute of Mathematics and Its Applications,"* 15:319–342, 1975.

Goldman, F. E., "The Application of Simulated Annealing for Optimal Operation of Water Distribution Systems," Ph.D. dissertation, Arizona State University, Tempe, 1998.

Goldman, F. E., A. B. Sakarya, L. E. Ormsbee, J. G. Uber, and L. W. Mays, "Optimization Models for Operation," Chap. 16, in L. W. Mays (ed.), *Water Distribution Systems Handbook,* McGraw-Hill, New York, 2000.

Grayman, W., R. A. Dieninger, A. Green, P. Boulos, R. W. Bowcock, and C. Godwin, "Water Quality and Mixing Models for Tanks and Reservoirs," *Journal AWWA,* 88(7):60–73, 1996.

IAPMO, *Uniform Plumbing Code,* International Association of Plumbing and Mechanical Officials, 1994.

Kirkpatrick, S., C. D. Gelatt, and M. P. Vecchi, "Optimization by Simulated Annealing," *Science,* 220(4598):671–680, 1983.

Lansey, K. E. and K. Awumah, "Optimal Pump Operations Considering Pump Switches," *Journal of Water Resources Planning and Management,* ASCE, 120(1):17–35, 1994.

Lansey, K. E., and L. W. Mays, "Optimization Model for Water Distribution System Design," *Journal of Hydraulic Engineering,* 115(10):1401–1418, 1989.

Lansey, K., and Q. Zhong, "A Methodology for Optimal Control of Pump Stations," *Water Resources Infrastructure (Proceedings of 1990 ASCE Water Resources Planning and Management Specialty Conference*), ASCE, 1990, pp. 58–61.

Lasdon, L. S., and A. D. Waren, *GRG2 User's Guide.* Department of General Business. The University of Texas at Austin, 1986.

Lasdon, L. S., A. D. Waren, A. Jain, and M. W. Ratner, "Design and Testing of a Generalized Reduced Gradient Code for Nonlinear Programming," *ACM Transactions on Mathematical Software,* 4(1):34–50, 1978.

Li, Guihua, and L. W. Mays, "Differential Dynamic Programming for Estuariane Management," *Journal of Water Resources Planning and Management,* ASCE, (121)6:455–462, 1995.

Little, K., and B. McCrodden, "Minimization of Raw Water Pumping Costs Using MILP," *Journal of Water Resources Planning and Management,* ASCE, 115(4):511–522, 1989.

Malcolm Pirnie, *Peoria Water Master Plan,* Technical Memorandum No. 3, "Distribution and Flow Model" (draft), 1996.

Mays, L. W. (ed.), *Reliability Analysis of Water Distribution Systems,* American Society of Civil Engineers, New York, 1989.

Mays, L. W., *Optimal Control of Hydrosystems,* Marcel Dekker, New York, 1997.

Mays, L. W. (ed.), *Water Distribution Systems Handbook,* McGraw-Hill, New York, 2000.

Mays, L. W., and Y. K. Tung, *Hydrosystems Engineering and Management.* McGraw-Hill, New York, 1992, pp. 140–157.

Metropolis, N., A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller, "Equation of State Calculations by Fast Computing Machines," *Journal of Chemical Physics,* 21:1087–1092, 1953.

Miller, L. H., and A. E. Quilici, *The Joy of C,* John Wiley & Sons, New York, 1997.

Mulligan, A. E., and L. C. Brown, "Genetic Algorithms for Calibrating Water Quality Models," *Journal of Environmental Engineering,* 124(3):202–211, 1998.

Murtagh, B. A., and M. A. Saunders, "A Projected Lagrangian Algorithm and Its Implementation for Sparse Nonlinear Constraints," *Mathematical Programming Study,* 16:84–117, 1982.

Murtagh, B. A., and M. A. Saunders, *MINOS 5.1 User's Guide,* Stanford Optimization Lab., Stanford University, Stanford, Calif., 1987.

Nitivattananon, V., E. C. Sadowski, and R. G. Quimpo, "Optimization of Water Supply System Operation," *Journal of Water Resources Planning and Management,* ASCE, 122(5):374–384, 1996.

Ormsbee, L. E. (ed.), "Energy Efficient Operation of Water Distribution Systems," *A Report by the ASCE Task Committee on the Optimal Operation of Water Distribution Systems,* 1991.

Ormsbee, L. E. and K. E. Lansey, "Optimal Control of Water Supply Pumping Systems," *Journal of Water Resources Planning and Management,* 120(2):237–252, 1994.

Ormsbee, L. E., and S. L. Reddy, "Nonlinear Heuristic for Pump Operations," *Journal of Water Resources Planning and Management,* 121(4):302–309, 1995.

Ormsbee, L. E., T. M. Walski, D. V. Chase, and W. W. Sharp, "Methodology for Improving Pump Operation Efficiency," *Journal of Water Resources Planning and Management,* 115(2):148–164, 1989.

Ostfeld, A., and U. Shamir, "Optimal Operation of Multiquality Networks. I: Steady-State Conditions," *Journal of Water Resources Planning and Management,* 119(6):645–662, 1993a.

Ostfeld, A., and U. Shamir, Optimal Operation of Multiquality Networks. II: Unsteady Conditions," *Journal of Water Resources Planning and Management,* 119(6):663–684, 1993b.

Pontius, F. W., "Overview of the Safe Drinking Water Act Amendments of 1996," *Journal AWWA,* 88(10):22–33, 1996.

Reklaitis, G. V., A. Ravidran, and K. M. Ragsdell, *Engineering Optimization: Methods and Applications,* Wiley-Interscience, New York, 1983.

Rossman, L. A., *EPANET Users Manual,* Risk Reduction Engineering Laboratory, Office of Research and Development, U.S. Environmental Protection Agency, Cincinnati, Ohio, 1993.

Rossman, L. A., *EPANET Users Guide.* Drinking Water Research Division, Risk Reduction Engineering Laboratory, Office of Research and Development, U.S. Environmental Protection Agency, Cincinnati, 1994.

Rossman, L. A., and P. F. Boulos, "Numerical Methods for Modeling Water Quality in Distribution Systems: A Comparison," *Journal of Water Resources Planning and Management,* 122(2):137–146, 1996.

Rossman, L. A., P. F. Boulos, and T. Altman, "The Discrete Volume Element Method for Network Water Quality Models," *Journal of Water Resources Planning and Management,* 119(5):505–517, 1993.

Rossman, L. A., R. W. Clark, and W. M. Grayman, "Modeling Chlorine Residuals in Drinking-Water Distribution Systems," *Journal of Environmental Engineering,* 120(4):803–820, 1994.

Sadowski, E., V. Nitivattananon, and R. E. Quimpo, "Computer-Generated Optimal Pumping Schedule," *Journal AWWA,* 87(7): 53–63, 1995.

Sakarya, A. B., "Optimal Operation of Water Distribution Systems for Water Quality Purposes," Ph.D. dissertation, Arizona State University, Tempe, 1998.

Sakarya, A. B., and L. W. Mays, "Optimal Operation of Water Distribution Pumps Considering Water Quality," *Journal of Water Resources Planning and Management,* 126(4):210–220, 2000.

Unver, O. L., and L. W. Mays, "Model for Real-Time Optimal Flood Control Operation of a Reservoir System," *Water Resources Management,* Vol. 4, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1990, pp. 21–46.

Vasconcelos, J. J., P. F. Boulos, W. M. Grayman, L. Kiene, O. Wable, P. Biswas, A. Bhari, L. A. Rossman, R. M. Clark, and J. A. Goodrich, *Characterization and Modeling of Chlorine Decay in Distribution Systems,* AWWA Research Foundation and American Water Works Association, Denver, 1996, pp. 258–271.

Wanakule, N., L. W. Mays, and L. S. Lasdon, "Optimal Management of Large-Scale Aquifers: Methodology and Applications," *Water Resources Research,* 22(4):447–465, 1986.

Wood, D. J., *User's Manual—Computer Analysis of Flow in Pipe Networks Including Extended Period Simulations,* University of Kentucky, 1980.

Zessler, U., and U. Shamir, "Optimal Operation of Water Distribution Systems," *Journal of Water Resources Planning and Management,* 115(6):735–752, 1989.

CHAPTER 13
OPTIMAL MAINTENANCE, REHABILITATION, REPLACEMENT SCHEDULING

Caty Werey
Cemagref—ENGEES
Strasbourg, France

13.1 INTRODUCTION

Water distribution networks in industrialized countries are aging; many of them have pipes more than 50 years old. Therefore, much rehabilitation or renewal work appears to be necessary in the future to keep this infrastructure in good condition. In addition, regulations concerning water quality have become stricter and impose heavy investment costs on water utilities. So, choices need to be made in order to meet investment needs and water regulation responsibilities while keeping costs within the available budget. Up to now, renewal decisions have been mainly made with regard to the failure rate and in relation to roadworks planning. The present decision process is essentially curative; it seems therefore necessary to propose a proactive methodology for planning the renewal decisions.

13.2 WHAT IS A PLANNING PROCESS FOR RENEWAL?

The planning process for renewal can be divided into four steps:

1. The *diagnostic* of the network's health, which is based on the network's characteristics, environment, and history. This step gives information on the aging and alteration of the network through physical factors (failures, corrosion, water quality, etc.), as well as about the pipe's close environment.

2. The second step takes into account the results of step 1, but also takes into account economic and sociologic factors in order to optimize the renewal date. This *optimization* is made without constraints. So, on one hand it leads to a classification of the pipes and the identification of the most risky ones; on the other hand it gives the optimal budgets. The cost of potential needs will probably be higher than the financial resources that the utility can accommodate and also will be irregular with demand coming in peaks and troughs.
3. The third step takes into account the financial constraints in order to plan the financial resources and to smooth the expenditures for renewal. This step gives a tool for *planning budgets* in order to propose a sustainable budget evolution.
4. The fourth and last step concerns the *planning of the renewal works*. Pipes, streets, or parts of the network are chosen to be renewed at a fixed time after taking into account the results of the three preceding steps and operational constraints such as the geographic area, road planning, or works on other networks.

13.3 OPTIMIZATION OF THE RENEWAL DATE

This chapter focuses on the optimization step of the renewal process. It proposes a method to evaluate the budgets that would be economically optimal, taking into account the aging process but also other technical, social, and economic issues.

This method, a model called RENCANA, was developed by Werey (2000a,b). It is a decision tool that can be used to decide when to replace a pipe by answering the question: Is it cheaper to renew and pay, with certainty, the renewal cost than to keep in place the old pipe and accept the risk of failures given by the expected cost of maintenance? This cost includes the direct costs of reparation and social costs due to the social effects of the failure (water loss, damages, interruption in industrial activity, etc.). Hydraulic reliability approaches are also included to take into account the hydraulic function of a pipe in a network (Fujiwara, and de Silva, 1990; Berthin, 1994; Bremond and Berthin, 2001).

The first step for this model was led by Shamir and Howard (1979) who minimized the total cost including both renewal and repair costs. An exponential law was used for the failure evolution. Then this first model was developed for the cost valuation on the one hand, and for the failure prediction on the other.

Walski and Pellicia (1982) proposed that the reparation cost could be represented by a function of the pipe diameter and material and labor costs. Walski (1987) introduced the idea of social costs due to the failure and the repair. Further studies in France (Kennel, 1992; Elnaboulsi, 1993) applied social cost valuations on different scales of the network and showed clearly the effect of these indirect costs in an economic optimization. Walski and Male (2000) provide a fairly detailed evaluation of the various issues related to maintenance and rehabilitational replacement and an extensive list of related literatures.

Concerning the failure evaluation, survival functions were introduced by Andreou (1987) for the failure probability estimation, taking into account both the pipe's data and its environment and failure history. This approach was applied on French networks by Eisenbeis (1994) and Eisenbeis and Legat (2000). Optimization can then consider expected numbers of failures provided by such approaches (Elnaboulsi and Alexandre, 1997).

A more accurate approach is possible by using dynamic programming which allows us to take into account the growing information on the failure occurrence with the time, such as is already used in investment choice models (Dixit and Pindyck, 1994). Other approaches have been introduced previously by Lansey et al. (1992) and Kim and Mays (1999).

Dynamic programming divides the decision process of a system into decision steps and hazard steps. Each step makes the system change from one state to another. An uncertainty step is characterized by a probability of transition. In RENCANA, the different states are characterized by the cost of repair or of renewal and the probabilities of transition are the probabilities of occurrence of the next failure given by survival models based on the pipe's characteristics, its local environment, and its failure history (Andreou, 1987; Eisenbeis, 1994; Eisenbeis and Legat, 2000). These probabilities represent the *aging process.*

The RENCANA model also takes into account:

- The effect of a failure on the consumption point by comparing the water consumption "required" in a certain consumption point with the consumption "delivered" when one pipe element fails. This shows the *hydraulic function* of the pipe and needs a hydraulic simulation where the failed piece of pipe is cut (Berthin, 1994).
- The *consumer specificity* considers that consumers differ not only by their consumption volume, which gives the traditional classification (domestic, industrial, institutional consumer), but also by their sensibility toward a water delivery interruption (e.g., hospital, person on dialysis).
- The relation with the *environment of the pipe.*

13.4 EXAMPLE STUDY NETWORK

An example network with four consumption nodes and five pipe elements, as given in Fig. 13.1, is used to explain the objectives of the modeling. ($^1/_3$ and $^2/_3$ mean that more delivery goes through pipe element IV than through pipe element II; we choose the simplified sharing in $^1/_3$ and $^2/_3$.)

Pipe element j will have homogeneous physical (material, diameter, lining date, etc.) and environmental (traffic, position under road or pavement, corrosiveness, etc.) characteristics. The division of the network into pipe elements must also

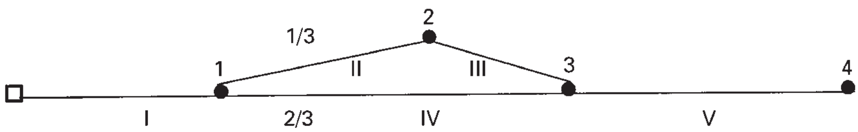


FIGURE 13.1 Example network.

take into account the hydraulic function, so each junction of pipes determines a new element.

Consumption node *i* will depend on the pipe element definition but also on the consumer specificity. For instance, several domestic consumers can be bundled at the same node, but a hospital or a dentist's office must be clearly identified because of its sensitivity to interruption of water.

Examples of renewal ordering without any modeling are as follows: First consider the example network from only the hydraulic point of view (all pipe elements and nodes are similar). The renewal order could be

I, V, IV, II, III

If both hydraulic function and consumer specificity are taken into account, for instance if node 2 is a sensitive consumer, then the renewal order could be

I, II, V, IV, III

If node 4 is sensitive, the renewal order could be

I, V, IV, II, III

Now we will include hydraulic function and aging. For instance, if pipe element III is very damaged, the proposition for renewal order could be

I, III, V, IV, II

The environmental effect on cost and hydraulic function is illustrated by the case where pipe element IV is a main road with high maintenance costs. We could propose the following renewal order:

I, IV, V, II, III

13.5 THE RENCANA MODEL

The model construction uses dynamic programming in order to profit from the growing information that appears as time moves forward (Dixit and Pindyck, 1994). The model is applied to each pipe element *j*. At each time, two decisions are possible: *to replace* or *not to replace.* Replacement here means to change the pipe at the very same time, that is, at the same diameter but with current material and lining conditions. Rehabilitation techniques could be taken into account in a further development of RENCANA. If the decision taken is *to replace,* the new pipe is supposed to have no failure on the simulation horizon, and the only cost will be the renewal cost R_j. On the other hand; when the decision is *not to replace,* two hazard possibilities are considered: *failure* on the pipe element *j* or *no failure* on the pipe element *j*.

The time step is the year, and we assume that a pipe element fails at most once a year. We consider also that only one pipe element is failing at one time, which means that when a pipe element is damaged, the next failure will not occur before the damaged pipe is repaired.

In this case the cost of maintenance m_j is as follows:

$$m_j = r_j + s_j$$

where r_j is the repair cost and s_j is the social cost. The cost flow could be as shown in Fig. 13.2. For any pipe element, the maintenance cost of any failure is considered equal to m_j. The pipe replacement (renewal) is made at the beginning of the period, while the failure in the modeling is supposed to have occurred at the middle of the period.

13.5.1 Valuation of the Decision Tree

We assume that the pipe element will be renewed at least by time Ω_j, the end of the simulation. At each node by reverse induction we compare the cost of renewing now or waiting and renewing later with the risk of a failure occurring and a repair being needed (see Fig. 13.3).

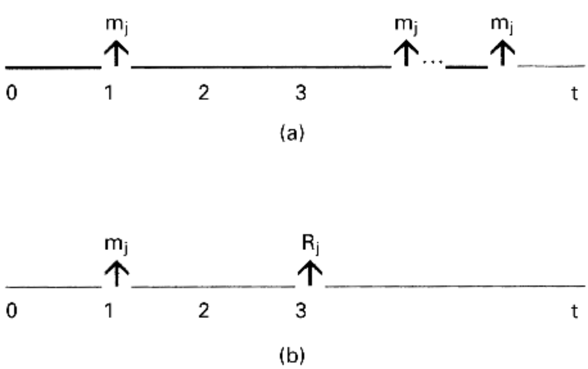


FIGURE 13.2 (*a*) Pipe kept in place, (*b*) replacement.

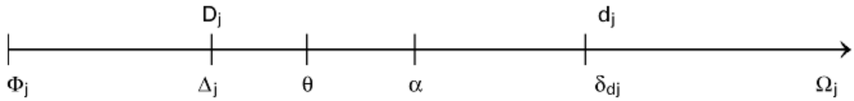


FIGURE 13.3 Φ_j, lining date of pipe element j; D_j, number of failures on pipe element j before beginning of simulation; Δ_j, date of Dth failure on j; α, beginning of simulation (January 1); θ, year of beginning of failure registration; ω, end of simulation; Ω, simulation horizon for pipe element j, $\Omega_j = \omega_j - \alpha$; t, number of years since α; d_j, number of failures since beginning of simulation; δ_{dj}, date of dth failure; a, discount rate.

- *Moving up a branch.* The probability of transitioning from state x_j[date: $\alpha + t - 1$, number of failures $D_j + d_j, \mu_j$] to state $x_j[\alpha + t, D_j + d_j + 1, \delta_{d+1)j} = \alpha + t - 1/2]$ is given by $p_j(t, D_j + 1, \mu_j)$.
- *Moving down a branch.* The probability of transitioning from state $x_j(\alpha + t - 1, D_j + d_j, \mu_j)$ to state $x_j[\alpha + t, D_j + d_j, \mu_j]$ is given by $[1 - p_j(t, D_j + d_j + 1, \mu_j)]$, with $\mu_j = \Delta_j$ if $d_j = 0$, and $\mu_j = \delta_{dj}$ otherwise.

$p_j(t, D_j + d_j + 1, \mu_j)$ is also the probability that on pipe element j, the $(D_j + d_j + 1)$th failure occurs during period t, knowing that the $(D_j + d_j)$th failure occurred at date μ_j, with $\mu_j = \Delta_j$ if $d_j = 0$, and $\mu_j = \delta_j$ otherwise.

The RENCANA model gives for each state $x_j(\alpha + t - 1, D_j + d_j, \delta_{dj})$, that is, at the beginning of period t, the cost $C_j^*\,[(\alpha + t - 1, D_j + d_j, \delta_{dj})]$ of the best decision concerning state $x_j(\alpha + t - 1, D_j + d_j, \delta_{dj})$ and so back to $C_j^*\,(\alpha, D_j, \Delta_j)$ ($t \in [1, T]$)

$$C_j^*\,(\alpha + t - 1, D_j + d_j, \delta_{dj}) = \min$$

$$\Bigg\{ R_j;\, p_j(t, D_j + d_j + 1, \delta_{dj})\, \frac{m_j}{\sqrt{1 + a}}$$

$$+\, p_j(t, D_j + d_j + 1; \delta_{dj})\, \frac{C_j^*\,[(\alpha + t, D_j + d_j + 1, \delta_{(d+1)j} = \alpha + t - 1/2)]}{1 + a}$$

$$+\, [1 - p_j(t, D_j + d_j + 1, \delta_{dj})]\, \frac{C_j^*\,[(\alpha + t, D_j + d_j, \delta_{dj})]}{1 + a} \Bigg\}$$

and

$$C_j^*\,(\alpha, D_j, \Delta_j = \min$$

$$\Bigg\{ R_j;\, p_j(t, D_j + 1, \Delta_j) \left(\frac{m_j}{\sqrt{1 + a}} + \frac{C_j^*\,[(\alpha + 1, D_j + 1, \alpha + 1/2)]}{1 + a} \right)$$

$$+\, [1 - p_j\,(t, D_j + 1, \Delta_j)]\, \frac{C_j^*\,[(\alpha + 1, D_j, \Delta_j)]}{1 + a} \Bigg\}$$

If C_j^* $(\alpha, D_j, \Delta_j) = R_j$, pipe element j must be replaced now, that is, during period 1. If C_j^* $(\alpha, D_j, \Delta_j) < R_j$, the renewal decision can be delayed for 1 year and the optimal decision will depend on the state that is reached at $\alpha + 1$ and the result of the valuation at this point.

Figure 13.4 presents an example with $\Omega_j = 3$, where it has been assumed, without modeling, that when a failure has occurred, the probability of having another failure grows faster than when no failure occurred.

13.6 RESULTS GIVEN BY THE MODEL

13.6.1 On the Example of Three Periods

The numeric example gives the following information:

- At time α, C_j^* $(\alpha, D_j, \Delta_j) = 194.11$ means that the pipe element can be let in place and the renewal delayed.
- At time $\alpha + 1$, C_j^* $(\alpha + 1, D_j + 1, \alpha + 1/2) = 199.26$ and C_j^* $(\alpha + 1, D_j, \Delta_j) = 197.27$, so here again renewal can be delayed.
- At time $\alpha + 2$, only C_j^* $(\alpha + 2, D_j, \Delta_j)$ is different from R_j, which means that renewal should be done at $\alpha + 2$, except if no failure has occurred since $\alpha + 2$. In this last case we can wait one more year.

We have stressed the optimal decisions on the decision tree. We now apply the RENCANA model to the little study network.

13.6.2 On the Little Study Network

RENCANA was applied to the example network (Figure 13.1), with subjective numerous values. Nodes 1 and 3 represent domestic consumers in individual houses, and node 2 represents domestic consumers in a block of flats. These three nodes aren't really sensitive to water delivery disruption, but the delivery in node 2 must be higher than in node 1 or 3, so pipe element II will be repaired with priority compared to pipe element IV: this is expressed by $m_{\mathrm{II}} = 50 > m_{\mathrm{IV}} = 30$. Node 4 will be an industrial consumer, sensitive to a water cut, so m_{V} will be important ($m_{\mathrm{V}} = 70$). According to the hydraulic function, m_{I} is high ($m_{\mathrm{I}} = 100$) and $m_{\mathrm{III}} = 50$. The five pipe elements are in different aging stages, $p_{\mathrm{I}} = p_{\mathrm{IV}} = p_{\mathrm{V}} = 0.2$, $p_{\mathrm{III}} = 0.3$, $p_{\mathrm{II}} = 0.5$, so pipe element II is more damaged than the others.

For environmental reasons, pipe elements are placed under a main road, and also due to the hydraulic function, we consider that R_{I} and R_{III} are higher than the other renewal costs: $R_{\mathrm{I}} = 500$, $R_{\mathrm{II}} = 200$, $R_{\mathrm{III}} = 350$, $R_{\mathrm{IV}} = 200$, $R_{\mathrm{V}} = 200$.

All the values discussed are displayed in Table 13.1, and the results of RENCANA for $\Omega = 3$ are given in Table 13.2. To summarize, the model results are as follows:

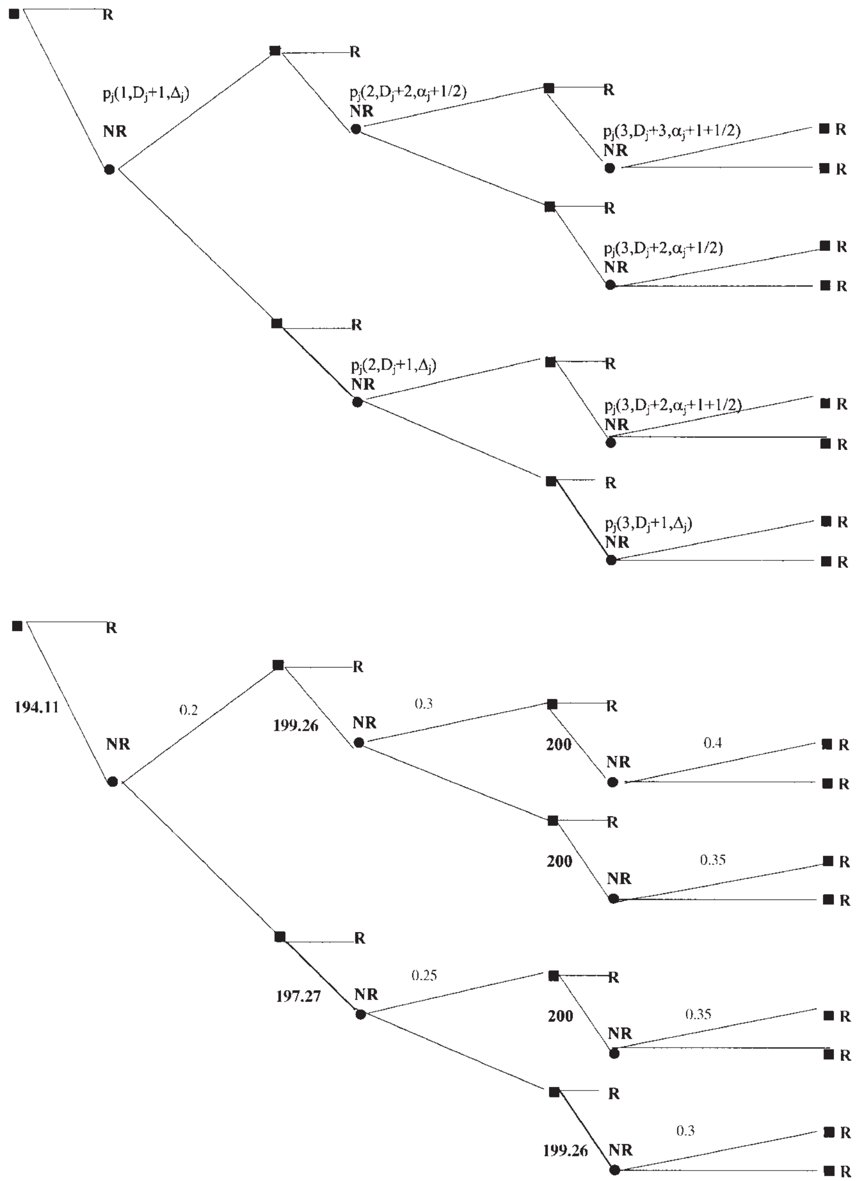


FIGURE 13.4 Example with $\Omega = 3$. Numeric application: $R_j = 200$ kF, $m_j = 50$ kF, $a = 0.05$, $D_j = 0$, $\alpha = \Phi_j$.

TABLE 13.1 Values of the Different Parameters

Pipe element	m_j	R_j	p_j
I	100	500	0.2
II	50	200	0.5
III	50	350	0.3
IV	30	200	0.2
V	70	200	0.2

TABLE 13.2 Results for RENCANA on the Little Study Network

Pipe element I	495.71	500 500	500 500 500 500
Pipe element II	200	200 200	200 200 200 200
Pipe element III	347.97	350 350	350 350 350 350
Pipe element IV	194.11	199.26 197.27	200 200 200 199.26
Pipe element V	200	200 200	200 200 200 200

- Pipe elements II and V should be renewed in the coming year.
- Pipe elements I and III can wait until the second year for renewal.
- Pipe element IV can be delayed to the fourth year for renewal, unless a failure occurs during the next 2 years, in which case the pipe must be renewed during the third year (see the numerical example above, Figure 13.4).

The aging stage of pipe element II and the importance of the consumer at node 4 are expressed in these results.

How can we also take into account the hydraulic function of the pipe? We consider this in Sec. 13.7.

13.7 CONSIDERATION OF THE HYDRAULIC FUNCTION OF THE PIPE IN THE NETWORK

A hydraulic simulation will be necessary to determine the demand (flow in the pipes hydraulic head at the nodes) at all consumption nodes when a certain pipe element is failing. The assumption is made that in this case no water can flow through the pipe element and that only one pipe can break at one time.

Several reliability indexes can be used (Berthin, 1994; Bremond and Berthin, 2001); we propose to use the following two are proposed

Relative reliability from pipe element j *to node* i: impact of a break on link j on consumption at node i

$$f_{ij} = \frac{v_{ij}}{V_i}$$

where V_i = desired water demand
v_{ij} = consumption at node i when pipe element j has a failure

Nonsatisfaction index: expected reliability at all nodes when pipe element j has a break (Werey, 2000a)

$$f_j^* = p_j \prod_{k \neq j} (1 - p_k) \left(1 - \frac{\sum_{i=1}^{I} w_i V_i f_{ij}}{\sum_{i=1}^{J} w_i V_i} \right)$$

with
$$W_i = \frac{w_i V_i}{\sum_{i=1}^{I} w_i V_i}$$

where p_j = probability that pipe element j has a failure

w_i = weighting factor representing the economic sensitivity of consumers at node i to a reduction in water use
I = number of consumption nodes

This index has been valued on the little study network, with the input values given in Table 13.3. The results are presented in Table 13.4.

The nonsatisfaction order is II, III, I, V, IV:

- II has the highest nonsatisfaction index, expressing the high probability of failure and that consumer 2 is a big consumer.
- III is in second position, owing to the high probability of failure and the hydraulic rule.
- For I it is essentially the hydraulic rule that is stressed.
- V has an important consumer to supply.
- IV has the lowest nonsatisfaction index, owing to the low probability of failure and the meshing.

If we consider RENCANA results, we see that RENCANA gave more importance to the sensitivity consumer at node 5, so pipe element V must be renewed immediately. The nonsatisfaction index f_j^* will help take into account the hydraulic rule of a pipe element and so rank the pipes identified by RENCANA for the following year's renewal program in regard to budget restrictions.

TABLE 13.3 Input Values for f_j^*

Node i	w_i	V_i	Pipe element j	p_j
1	1	5	I	0.2
2	1	20	II	0.5
3	1	5	III	0.3
4	2	20	IV	0.2
			V	0.2

TABLE 13.4 Results of the Little Study Network for f_j^*

Hypothosis for v_{ij}					Result for f_j^*
$v_{1I} = 0$	$v_{1II} = V_1$	$v_{1III} = V_1$	$v_{1IV} = V_1$	$v_{1V} = V_1$	$f_I^* = 0.03$
$v_{2I} = 0$	$v_{2II} = 1/3V_2$	$v_{2III} = 2/3V_2$	$V_{2IV} = 2/3V_2$	$V_{2V} = 3V_2$	$f_{II}^* = 0.07$
$v_{3I} = 0$	$v_{3II} = 2/3V_3$	$v_{3III} = 1/3V_3$	$V_{3IV} = 2/3V_3$	$V_{3V} = V_3$	$f_{III}^* = 0.04$
$v_{4I} = 0$	$v_{4II} = 2/3V_4$	$v_{4III} = 1/3V_4$	$V_{4IV} = 2/3V_4$	$V_{4V} = 0$	$f_{IV}^* = 0.01$
					$f_V^* = 0.02$

13.8 COST VALUATION

The valuation of the repairing cost and the renewal cost can be provided by a cost accounting approach. Data analysis can be used to link costs and environmental or pipe characteristics. Standard cost can be established based on a standard repair, for example, putting in place a sleeve, and thus give a valuation of all the principal costs concerned (material, staff, machinery, etc.). Average cost results from an analysis of previous repair costs. In the following section, standard cost was used for the repairing cost and average cost for the renewal cost, which also includes house connections renewal.

The social effects of a break will be evident when the social costs to different parties are taken into account: the utility first of all (water losses, quality of service, reputation, etc.), the water consumers (discomfort, loss of economic activity, cost of substitute water delivery, damage, etc.), residents (discomfort, flooding, damage, etc.), and pedestrians and drivers (loss of time, discomfort, etc.). These costs have already been identified as having an influence on the renewal data, an idea that has been perceptively put forward (Clark, 1982; Kennel, 1992; Elnaboulsi, 1993; Drane, 1997). Several methods of cost valuation can be identified using contingency approaches, introducing valuation of substitutions or inquiries results.

The following testing focuses on introducing a consumers' differentiation by valuing the loss of economic activity for industry and trades and the discomfort to domestic consumption due to an interruption of water delivery during the 2 hours after a failure:

- For domestic consumers, each family was considered to incur a cost of 150 francs corresponding to the cost of a meal taken outside for three persons (Werey, 2000a). A similar method was used to determine costs incurred by schools and kindergartens.
- For trades and services, such as small supermarkets, dentists, hairdressers, or laboratories, results from a study on an interruption of drinking water delivery during 110 hours in Tours, in 1988, due to pollution of the water supply, were used to reduce the duration to 2 working hours.
- For industry, ratios edited by the Banque de France for each specific activity sector allow for the calculation of loss in business profit, relating to the number of employees, when the activity is stopped for 2 hours.
- For people involved in hospital or home dialysis, the price of a tank used as a water delivery substitute could be introduced.

The social cost approach needs to give a valuation of all the effects to be included in the renewal process, and as RENCANA is applied at a pipe element level, each piece of information must be available at this level and connected with the right pipe element.

13.9 TESTING RENCANA ON REAL DATA FROM WATER NETWORKS

13.9.1 Example 1: Urban Network

A valuation was made on 14 pipe elements corresponding to 6 streets in a district with houses, buildings, trades, and services on one side and industrial activity on the other side.

Data are presented in Tables 13.5 and 13.6. Table 13.5 gives the valuation of the different costs: maintenance cost (col. 7) = social cost (cols. 2 to 5) + repair costs (col. 6), and renewal cost (col. 10) including connection renewal as detailed in Sec. 13.8. Table 13.6 gives the technical data of the pipe and the failure history. These data have been used for the aging modeling with survival functions in order to determine the different probabilities used for the decision tree. Modeling on 10 years gave the results for $C_j^* (\alpha, D_j, \Delta_j)$ ($a = 5\%$) presented in Table 13.7.

RENCANA suggests that only pipe 42-43 has to be renewed now. Examination of the whole decision tree over 10 years indicates that there is no need to renew the 13 other pipes. That means that the renewal can be postponed for at least 10 years. The same result was given by a modeling on 20 years. This result can be explained by the slow aging process and also, except for pipe 42-43, the low value of the social cost with regard to the renewal cost. In fact, industrial consumers have a much higher social cost than other consumers. But damages to an underground parking lot or tramway traffic, or even causing a traffic jam, can also lead to high social costs. Here, however, only social costs for water consumers were included.

We now focus only on the results for two pipes: pipe 42-43, because of the high social cost and because it is a pipe with no previous failure; and pipe 76-99 which already has had four failures (Tables 13.8 and 13.9).

Tables 13.8 and 13.9 present an extract of the modeling over 10 years, corresponding to the beginning of the tree. The left side presents the probabilities, and the right side presents the cost valuation. Notice the difference of the aging of the two pipes and also the fact that the probabilities on the lowest branches are the highest. This traduces the effect of the period without failure from the last real failure: the risk for the next failure to occur is higher than in the top of the tree where the delay between two failures is 1 year, but where the number of failures grows each year. So the delay between the failures seems to have much more of an effect on the probabilities than the number of failures.

13.9.2 Example 2: Rural Network

Example 1 presented a probability distribution that was different from the distribution that was expected and used for the little study network, and so the renewal decision appears at first not in the top of the tree but in the bottom. Example 2 will

TABLE 13.5 Urban Network: Social and Economic Data

1	2	3	4	5	6	7	8	9	10
Pipe	Social: domestic, kF	Social: trade, kF	Social: industrial, kF	Total social cost, kF	Repair cost, kF	Mainten. cost, kF	Length, m	No. of connections	Renewal cost, kF
75.102	7.35			7.35	12.50	19.85	220.00	4	343.00
75.103	0.15	12.24		12.39	12.50	24.89	132.00	2	196.80
77.78	18.45	3.94		22.39	8.70	31.09	312.00	24	639.00
76.77	0.30			0.30	8.70	9.00	100.00	2	105.00
74.76	23.75	1.36		25.11	12.25	37.36	308.00	11	471.00
69.70	6.90	20.18		27.08	12.25	39.33	340.00	44	675.00
34.71	3.15			3.15	8.70	11.85	112.00	5	174.00
71.72	4.80	1.48		6.28	8.70	14.98	100.00	10	240.00
34.70	29.50			29.50	12.25	41.75	360.00	28	849.00
96.99	1.95			1.95	8.70	10.65	312.00	3	279.00
76.99a	8.05			8.05	8.70	16.75	404.00	15	543.00
48.38			41.49	41.49	12.25	53.74	520.00	14	615.00
39-42			54.78	54.78	12.25	67.03	460.00	10	525.00
42-43			474.65	474.65	12.25	486.90	252.00	5	279.00

Note: kF = 1000 francs.

TABLE 13.6 Urban Network: Technical Data

Pipe j	No. of failures, D_j	Failure date, Δ_j	Lining year	Nominal diameter	Pipe thickness	Pipe length	Pipe material
75.102	0		1993	150	6.3	220	Dc
75.103	0		1964	150	8	132	Sc
77.78	2	July 1, 1985	1904	100	9.6	312	P
76.77	0		1934	100	9.6	100	P
74.76	1	Feb. 8, 1989	1960	100	7.2	308	Sc
69.70	0		1944	200	8.8	340	S
34.71	2	Oct. 1, 1993	1917	100	9.6	112	P
71.72	0		1930	100	9.6	32	P
34.70	0		1917	200	11.9	360	P
96.99	2	Dec. 21, 1991	1930	100	9.6	100	P
76.99	4	Jan. 14, 1992	1930	100	9.6	404	P
48.38	0		1965	150	8	520	Sc
39-42	0		1922	150	9.6	460	P
42-43	0		1913	150	9.6	252	P

Note: P = pit iron; S = spun iron; Sc = spun coated iron; D = ductile iron; Dc = ductile coated iron.

show a difference between pipes that have already failed before the modelling and the other pipes.

The data used and the results of the optimization are presented in Tables 13.10 and 13.11. Costs have been valued simply on two levels: low and high. Table 13.11 shows that pipe 4 has to be renewed immediately, but that the others can wait.

All pipes are made of spun iron. For the urban network example one survival function, which mixed pipes with failures and those with no failures, was evaluated. Here two survival functions were considered, one for pipes with no failure and one for pipes with failures. Tables 13.12 and 13.13 show an example for each category.

Tables 13.12 and 13.13 show that when a pipe has already had failures, it has a different probability evolution than when no failure has occurred. In the first case RENCANA puts an emphasis on the number of failures, so the decision tree gets worse at the top, whereas in the second case, the emphasis is put on the delay to get the first failure, so the decision tree is worse at the bottom.

TABLE 13.7 Urban Network: Results

Pipe	75,102	75,103	77,78	76,77	74,76	69,7	34,71	71,72	34,70	96,99	76,99	43,38	39,42	42,43
C_j^*, kF	240.4	149.1	426.9	77.14	332.3	444.3	116.1	162.9	551.9	188.0	376.0	459.9	407.0	279.0
Decision	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	R

Note: kF = 1000 francs; NR = no renewal; R = renewal.

TABLE 13.8 Urban Network: Pipe 42-43

Pipe	**42-43**	**0 failure(s)**		**R = 279 kF**		**m = 487.15 kF**	
Proba				**Optimization cost (10 years)**			
			0.028211				279
		0.026062	0.031228			278.1	279
	0.022802		0.024669		276.22		277.43
		0.027331	0.030301			278.66	279
0.082217			0.024669	279			277.43
		0.215567	0.025854			274.96	277.95
	0.083487		0.020411		279		273.32
		0.084698	0.085857			279	279
[0, 1]	[1, 2]	[2, 3]	[3, 4]	$t = 0$	$t = 1$	$t = 2$	$t = 3$

Note: kF = 1000 francs.

TABLE 13.9 Urban Network: Pipe 76-99

Pipe	**76-99**	**4 failure(s)**		**R = 543 kF**		**m = 16.75 I205 kF**	
Proba				**Optimization cost (10 years)**			
			0.138732				419.49
		0.139501	0.165149			402.37	420.19
	0.138141		0.130829		385.92		418.78
		0.164729	0.181156			402.9	420.51
0.293567			0.130829	376.02			418.78
		0.130369	0.154494			401.45	419.36
	0.302698		0.122223		391.39		417.99
		0.310872	0.318269			407.64	424.81
[0, 1]	[1, 2]	[2, 3]	[3, 4]	$t = 0$	$t = 1$	$t = 2$	$t = 3$

13.10 CONCLUSION

RENCANA gives, at the pipe level, an answer to the following question: Should the pipe be repaired or renewed and, if so, when? It gives the expected cost of renewal for the different years and shows the effect of taking into account social cost in order to take greater care of social aspects in the renewal decision. It is a step toward a general planning model helping management to know the condition of its network and to decide which pipe element should economically be renewed.

The nonsatisfaction index helps to prioritize the pipes to be renewed according to a proposed budget, which can be lower than the budget resulting from the optimization.

This approach shows the importance of data gathering in considering the different parts of a network. On the one hand, databases give a failure prediction that takes into account the characteristics of the pipe and its environment and the failure history. On the other hand, they include consumer specificity by social cost

TABLE 13.10 Rural Network: Technical Data

Pipe j	No. of failures, D_j	Failure date, Δ_j	Lining year, Φ_j	Nominal diameter	Pipe length, m	Traffic*	Acidity*	Wetness*	Rocky*
3	5	Nov. 22, 1996	1927	125	2500	1	0.5	1	0
4	3	July 25, 1996	1931	40	300	0	0	0	1
5	10	Aug. 28, 1992	1927	60	2000	1	0.5	1	0
7	0		1963	300	2100	1	0	0	0
8	0		1927	100	370	0	0.5	1	0
9	0		1961	60	2600	0	0	0	0.5
11	0		1963	300	3600	0	0	0	0
12	1	Feb. 20, 1991	1950	100	1350	0	0	0	0
13	0		1931	175	400	0	0	1	0

*1 = yes, 0 = no.

TABLE 13.11 Rural Network: Economical Data and Results

	Low costs				High costs			
Pipe *j*	R_j, kF	$C_j^*(\alpha, D_j, \Delta_j)$, kF	$C_j(\alpha, D_j, \Delta_j)$, kF	Decision	R_j, kF	$C_j^*(\alpha, D_j, \Delta_j)$, kF	$C_j(\alpha, D_j, \Delta_j)$, kF	Decision
3	2250.00	1561.79	1561.79	NR	4750.00	3730.88	3730.88	NR
4	270.00	**270.00**	*291.46*	**R**	420.00	420.00		**R**
5	1800.00	1303.79	1303.79	NR	2375.00	2275.49	*2392.58*	NR
7	1890.00	1220.00	1220.00	NR	2490.00	1619.81	1619.81	NR
8	333.00	216.43	216.43	NR	500.50	333.29	333.29	NR
9	2340.00	1510.93	1510.93	NR	3065.00	1991.02	1991.02	NR
11	3240.00	2089.36	2089.36	NR	4215.00	2721.99	2721.99	NR
12	1215.00	793.46	793.46	NR	1627.50	1110.72	1110.72	NR
13	360.00	232.79	232.79	NR	535.00	349.29	349.29	NR

Note: kF = 1000 francs; NR = no renewal; R = renewal.

TABLE 13.12 Rural Network: Pipe 5

Pipe	**76-99**	**4 failure(s)**		**R = 1 800 kF**		**m = 68.70 kF**	
Proba				**Optimization cost (10 years)**			
			0.424799				1471.28
		0.424799	0.331822			1423.33	1459.65
	0.424799		0.424799		1377.53		1471.28
		0.331822	0.286918			1411.46	1452.42
0.240866			0.424799	1303.79			1471.28
		0.424799	0.331822			1423.33	1459.65
	0.224569		0.424799		1343.94		1471.28
		0.211606	0.200946			1387.22	1433.74
[0, 1]	[1, 2]	[2, 3]	[3, 4]	$t = 0$	$t = 1$	$t = 2$	$t = 3$

TABLE 13.13 Rural Network: Pipe 7

Pipe	**7**	**0 failure(s)**		**R = 1890.00 kF**		**m = 68.70 kF**	
Proba				**Optimization cost (10 years)**			
			0.001354				1411.12
		0.001378	0.001849			1344.13	1411.24
	0.001405		0.001354		1280.34		1411.12
		0.001883	0.002223			1344.26	1411.34
0.004354			0.001354	1220.76			1411.12
		0.001378	0.01849			1344.13	1411.24
	0.004537		0.001354		1281.51		1411.12
		0.00471	0.004875			1345.26	1412.19
[0, 1]	[1, 2]	[2, 3]	[3, 4]	$t = 0$	$t = 1$	$t = 2$	$t = 3$

valuation. Of course, these databases must be maintained in order to actualize and enrich the information available on a network.

The next step in the decision process for renewal works is to consider the operational constraints due to the other underground assets and to road planning, for instance, and to include budget constraints relying on a sustainable water price evolution.

This chapter focused on renewal decisions. A similar approach could be developed for rehabilitation decisions, but rehabilitation effects on the "renewal cost," on the social benefits of trenchless solutions, and on the evolution of the rehabilitated pipe will have to be considered.

13.11 REFERENCES

Andreou, S. A., et al., "A new methodology for modeling break failure patterns in deteriorating water distribution systems," *Advanced Water Resources,* March 10, 1987, pp. 2–20.

Berthin, S., "Indices de fiabilité d'un système de distribution d'eau potable," DEA "Mécanique et ingénierie," U.L.P. Strasbourg, 1994, 79 pp.

Bremond, B., and S. Berthin, "Reliability of drinking water supply system," *IWA Specialised Conference, Brno, Czech Republic,* May 16–18, 2001, pp. 167–174.

Clark, R. M., et al., "Water distribution systems: a spatial and cost evaluation," *Journal of the Water Resources Planning and Management Division,* 109:243–256, October 1982.

Dixit A. K., and R. S. Pindyck, *Investment under Uncertainty,* Princeton University Press, Princeton, N.J., 1994, 455 pp.

Drane X., Evaluation des coûts sociaux liés à une interruption de l'alimentation en eau potable, 1997, 74 pp.

Eisenbeis, P. "Modélisation statistique de la prévision des défaillances sur les conduites d'eau potable," Thése de doctorat "Génie de l'environment," U.L.P. Strasbourg, 1994, 156 pp. + annexes.

Eisenbeis, P., and Y. Legat, "Using Maintenance Records to Forecast Failures in Water Networks," *Urban Water* 2:173–181, 2000.

Elnaboulsi, J., "Le renouvellement des réseaux d'eau potable: approche économique d'optimisation, modèle des risques proportionnels," DESS "Economie Industrielle et Décentralization," U.F.C. Besançon, 1993, 171 pp. + annexes.

Elnaboulsi, J., and O. Alexandre, "Détermination de la date optimale de renouvellement des réseaux urbains d'eau potable," *Canadian Water Resources Journal,* 22(3):327–349, 1997.

Fujiwara, O., and A. de Silva, "Algorithm for reliability-based optimal design of water networks," *Journal of Environmental Engineering,* 116(3):575–587, May/June, 1990.

Kennel, S., "Optimisation du renouvellement des canalizations d'eau potable: approche méthodologique," DESS "Economie Industrielle et Décentralization," U.F.C. Besançon, 1992, 79 pp.

Shamir, U., and C. Howard, "An analytic approach to scheduling pipe replacement," *Journal of AWWA,* 71:248–258, 1979.

Walski, T. M., "Replacement rules for water mains," *Journal of AWWA,* 79:33–37, 1987.

Walski, T. M., and A. Pellicia, "Eocnomic analysis of water main breaks," *Journal of AWWA,* 74:140–147, 1982.

Werey, C., "Politiques de renouvellement des réseaux d'eau potable," Ph.D. dissertation in management science, Thèse de Doctorat "Sciences de gestion," U.L.P. Strasbourg, 2000a, 162 pp.

Werey, C., "Ageing water mains: evaluating the optimal date of renewal," *AWWA 2000 Infrastructure Conference Baltimore,* March 12–15, 2000b, 11 pp.

CHAPTER 14
COMPARING MICROBIAL AND DBP RISK TRADEOFFS IN DRINKING WATER: APPLICATION OF THE CRFM

Robert M. Clark
Glenn Rice
U.S. Environmental Protection Agency
Cincinnati, Ohio

14.1 INTRODUCTION

A major objective of drinking water treatment is to provide potable water free from chemical and microbial contaminants. Conventional drinking water treatment combined with disinfection has proven to be effective in achieving this objective and is often considered to be one of the major public health advances in modern times. In the United States, chlorine is most often the final disinfectant added to treated water for microbiological protection before the water is discharged into a drinking water distribution system. However, it is now common knowledge that chemical disinfectants, especially chlorine, react with *natural organic matter* (NOM) to form *disinfection by-products* (DBPs), which are considered to be of public health concern from a chronic exposure point of view.

Drinking water chlorination, therefore, poses the dilemma of a risk tradeoff. Chemical disinfection reduces the risk of infectious disease, but the interaction between chemical disinfectants and precursor materials in source water results in the formation of DBPs. In the early 1970s the identification of chloroform in drinking water (Rook, 1974; Bellar et al., 1974) raised questions about possible health risks posed by DBPs. Since 1974, additional DBPs have been identified and concerns have intensified about health risks resulting from exposures to DBPs although a causal link between DBP exposures and these health risks has not been conclusively established. Nevertheless, risk managers have responded, in the interest of protecting public health, by developing alternative treatment systems and issuing rules and regulations designed to maintain protective levels of disinfection while reducing potentially harmful levels of DBPs.

Many environmental exposures pose these types of tradeoffs. Often the objective of the risk assessor working in collaboration with the risk manager is to clarify choices between alternatives or to find the "optimal" decision point that minimizes the overall risk to exposed populations taking cost into consideration. In an attempt to explore these risk tradeoff issues, the EPA's National Center for Environmental Assessment (NCEA) and National Risk Management Research Laboratory (NRMRL) have initiated research into the use of integrated risk assessment and risk management techniques for environmental decision making. An early product of this collaborative effort was the development of a concept called the *Comparative Risk Framework Methodology* (CRFM).

This chapter describes the basic concepts underlying the CRFM. Various metrics that can be used to make comparisons among alternative interventions are discussed, and a case study is presented using one of the metrics, quality adjusted life years.

14.2 THE CRFM CONCEPT

The CRFM combines elements of the National Academy of Science (NAS) risk assessment paradigm with a cost-effectiveness analysis approach (NAS, 1986; Haddix et al., 1996; Gold et al., 1996). A draft document was prepared describing this effort and is available on the EPA's Office of Research and Development web site (*http://www.epa.gov/ncea/frame.htm*). It provides a systematic way to assess interdependent environmental health risks and compares the impact of alternative interventions designed to minimize these risks (Rice et al., 2001a). In this context water treatment could be considered an environmental health intervention designed to protect human populations against potentially pathogenic microorganisms in drinking water while minimizing their exposure to potentially harmful disinfection by-products. To illustrate this application, the CRFM was utilized to evaluate three water treatment strategies for protecting against microbial contamination while minimizing exposure to DBPs. *Quality adjusted life years* (QALYs) was used as a common health metric for comparing alternative strategies because of the disparate risks posed by pathogens and DBPs. In the case study the risk tradeoffs associated with applying increasing concentrations of ozone for the inactivation of *Cryptosporidium,* resulting in lower risk from microbial infection, were compared to the increased risk from exposure to bromate and other DBPs which may be a product of using ozone as a disinfectant.

The goal of the CRFM concept is to develop quantitative central-tendency estimates of the risks associated with various strategies that might be used to protect public health and to make comparisons among them. Economic costs may be estimated so that the public health gain (i.e., reduction in risk) can be compared with the cost to achieve that gain.

The CRFM concept is illustrated in Fig. 14.1. The five analytical levels shown in Fig. 14.1 are intended to reflect a process that systematically characterizes what is known and not known about the technical basis for the decision being addressed.

Level 1 assesses potential exposures by using data that describe the current environmental conditions and the intervention being analyzed or interventions being compared. Level 2 identifies potentially disparate health effects, quantifies the health risks associated with each intervention being considered, and defines the confidence in the linkage between intervention and each health effect. Level 3 translates health effects identified in level 2 into adverse human health conditions, which are defined as observable and measurable effects in humans. For the risk tradeoffs considered, level 4 uses weighting procedures to summarize and quantitatively value the morbidity and premature mortality associated with the adverse human health conditions defined in level 3. Results of the first four levels can be used to compare the public health impacts and can be integrated into cost-benefit analyses. Level 5 quantifies the monetary costs associated with the adverse human health conditions and the technology costs of the interventions. It compares the effects of the alternative interventions based on health and monetary resources.

One of the clear and most powerful benefits of using the CRFM approach is to assist in identifying areas where research needs to be conducted in order to bolster regulatory decision making. It provides a transparent and structured approach to risk analysis and establishes a basis for conducting sensitivity and uncertainty analysis. It also allows for identification of research that will reduce the quantitative uncertainty in risk estimation.

14.3 APPLICATION OF THE CRFM TO WATER SUPPLY

In the CRFM, the adverse human health conditions that are directly linked to each intervention are defined, quantified, and then valued. A single common metric (i.e., a summary outcome measure) is utilized to assess the adverse human health conditions associated with each intervention's intended and unintended impacts on the disease burden as compared to an existing intervention (i.e., the baseline intervention) or from implementation of alternative interventions.

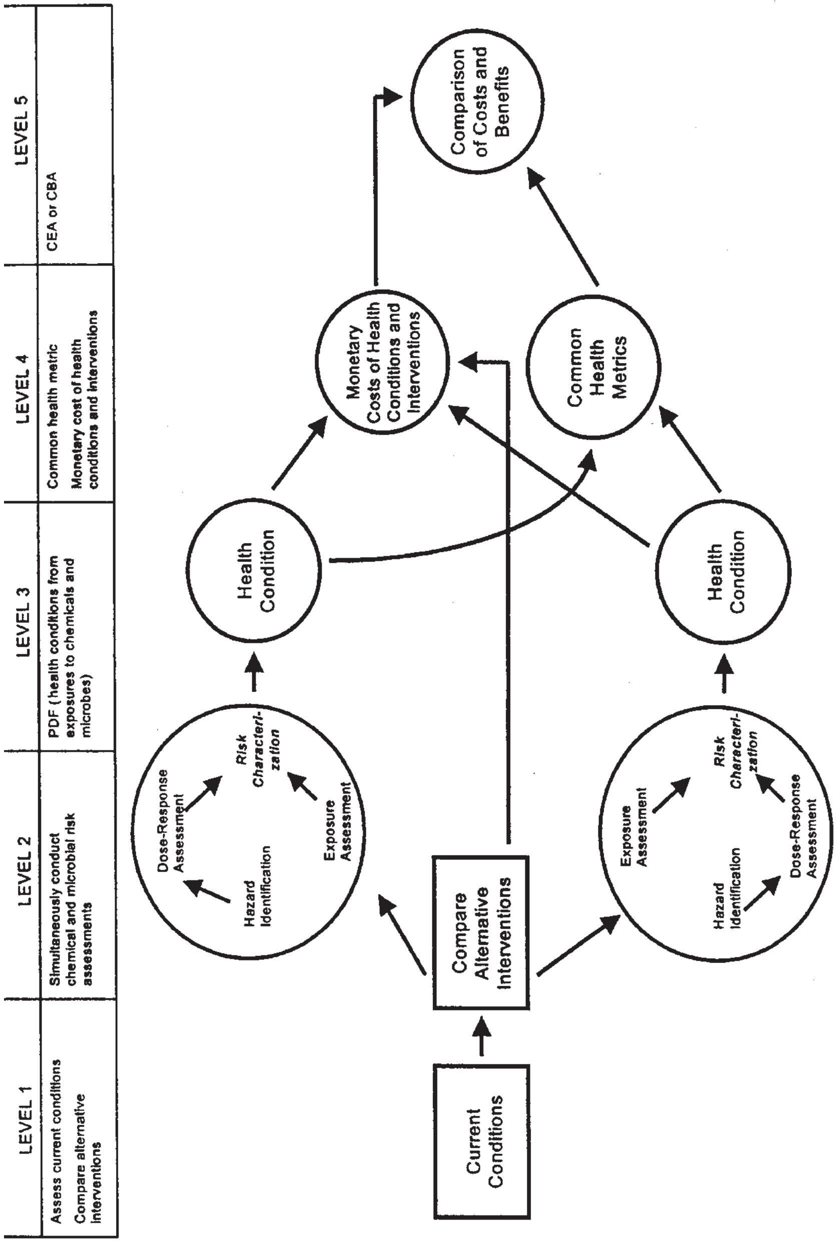


FIGURE 14.1 Schematic illustrating the comparative risk framework methodology.

14.3.1 Outcome Measures

A number of outcome measures may be used to compare adverse human health conditions associated with environmental public health interventions. These include natural events such as cases averted, deaths prevented, and life years saved. They can also be expressed as more abstract common health metrics such as health-related quality of life measures and dollars.

In levels 1 to 3 health risks are translated into specific adverse human health conditions, resulting in a description of the spectrum of morbidity and premature mortality within the population, for a series of interventions. However, it does not allow for a comparison of the impact of a given intervention on multiple health conditions, nor does it allow for a comparison of the overall health impact of alternative interventions. The initial challenge for a common health metric is to provide a meaningful way to aggregate disparate adverse human health conditions.

In level 4, the health metric may be used to compare the change in health with the resources that are required to generate the change and the impact of an intervention on the burden associated with distinct health conditions. That is, if an intervention reduces the burden of one condition but has unintended consequences that increase the burden of another health condition, then the impact of the intervention on the different conditions must be compared. When a comparative analysis of multiple interventions is made, then an aggregation and comparison of the net health impact of the different interventions is required. The comparison of health conditions within and between interventions requires that the health metric capture the relevant attributes of the different conditions in a consistent manner.

These outcome measures can then be used to assess risk tradeoffs and may be utilized in economic evaluations (level 5). Within the CRFM, health-related quality of life measures and monetary metrics can account for changes in morbidity and mortality associated with alternative interventions. In environmental health, dollar metrics have been used extensively to value changes in morbidity and mortality. In particular, they are commonly used to develop monetized benefit estimates in cost-benefit analyses (U.S. OMB, 1996a,b). Different approaches can be used for comparing costs and benefits both within and between interventions, depending on the outcome measure used to value and aggregate the adverse human health conditions. In cost-benefit analysis, monetary costs are compared to a monetary measure of the avoided disease burden by calculating a benefit-cost ratio or the net benefit. In cost-effectiveness analysis, monetary costs are compared to the disease burden by calculating a *cost-effectiveness* (CE) ratio. The CE ratio is calculated as the change in costs divided by the change in the disease burden, where the change associated with the avoided disease burden associated with an alternative intervention is calculated with respect to the disease burden attributable to the baseline intervention.

Some of the monetary-based approaches are willingness to pay in exchange for a specific improvement in health status, and revealed and stated preference methods. *Revealed preference* methods rely on the observation of individual responses related to health risks or specific health conditions. *Stated preference* methods attempt to value changes in morbidity and mortality through survey techniques that directly ask respondents about their willingness to pay for a specific change in health or a reduced risk of a health condition for each level of each attribute. This approach is used in prospective clinical trials where individuals describe their own conditions based on the classification system. In modeling applications, analysts or experts map the expected health conditions to be chosen. In both cases preferences for the conditions are based on a predetermined function.

In addition, the CRFM can also use measures of health-related quality of life, such as QALYs, because they provide a relatively transparent metric of morbidity and mortality. QALYs and similar measures are used extensively in cost-effectiveness analyses in other areas of public health. The use of monetary metrics and health-based metrics are not mutually exclusive, and in some cases a cost-benefit analysis using monetized benefits may complement a cost-effectiveness analysis using QALYs by providing decision makers with a monetary measure of health benefits along with nonhealth benefits.

14.3.2 The QALY Concept

The comparative risk problem posed by examining alternative drinking water treatment systems requires a common health metric capable of capturing the impact of reducing both morbidity and mortality associated with a range of health effects (e.g., microbial illness, cancer, and reproductive effects). Both health-related quality of life and dollars provide this type of common metric. Health-related quality of life measures changes in health status based on changes in life expectancy, and the durations of the changes are adjusted by the weight associated with the quality of life. Dollars provide a health metric for valuing changes in morbidity and mortality due to a variety of conditions. Dollar valuations of changes in health status are based on individuals preferences and their willingness to pay to avoid or be free from the specific condition.

Because monetary valuations of health outcomes have been described elsewhere for use in cost-benefit analyses (U.S. OMB, 1996a,b), a brief explanation of the potential use of QALYs for environmental cost-effectiveness analyses is given in this section. Measures of health-related quality of life such as QALYs provide a unidimensional measure of changes in health status expressed as years of life, where those years are weighted by the health-related quality of life during those years. Figure 14.2 illustrates how these measures combine morbidity and mortality into a composite measure. For example, the quality of a hypothetical individual's life over its duration is expressed on a scale of 0 (death) to 1 (perfect health). In general, individuals are not always in perfect health, and quality tends to decline with age (Putnam and Grahm, 1993). A curve is generated which relates the quality of an individual's life to his or her life span without any disease or illlness. The area under this curve represents the QALYs for this hypothetical individual.

Another curve is generated which represents a similar trajectory for the same individual if he or she were to develop a disease which resulted in a short period of morbidity and eventually premature mortality. The QALYs for this individual would be reduced to the area under the second curve. As a result, the condition would result in a loss of QALYs equivalent to the area between the two curves. This area could be divided into a portion attributable to premature mortality and a portion attributable to reduced quality of life during the person's lifetime. If the individual recovered from the condition, the measure would still capture the portion resulting from the initial morbidity. For example, the individual might suffer a short period of illness from which he or she recovers. Then he or she might become ill again and eventually die.

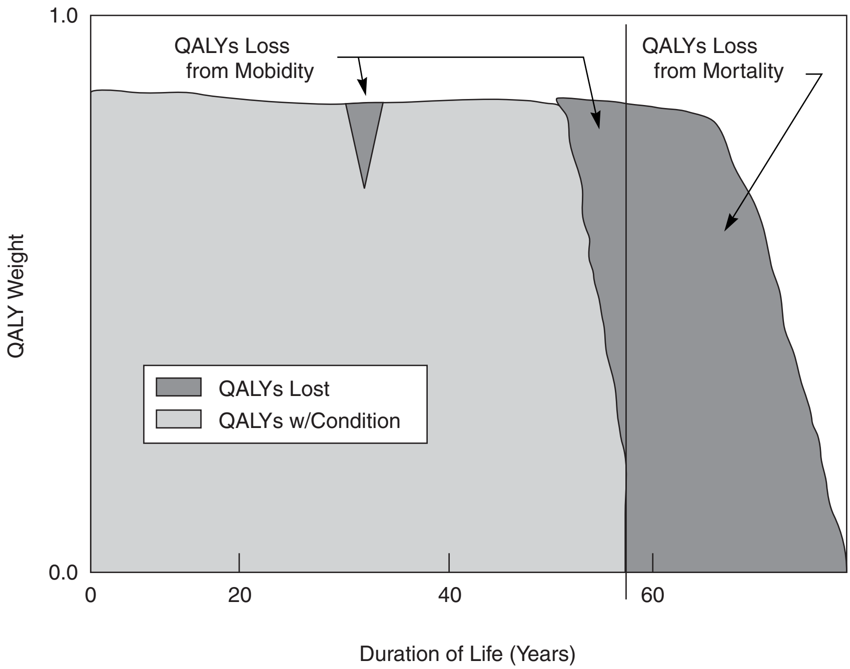


FIGURE 14.2 Quality cost of hypothetical microbial illness and death for a typical individual.

14.4 WATER TREATMENT RISK TRADEOFF EXAMPLE

The following case study was developed to demonstrate how the CRFM could be used to evaluate alternative drinking water disinfection approaches. Three treatment options are evaluated in terms of their impact on microbial risks [gastrointestinal (GI) illnesses and mortality], their impact on DBP-induced risks (cancer, reproductive toxicity, and developmental toxicity), and their financial costs. In this analysis a conventional water treatment plant (coagulation, sedimentation, sand filtration, and chlorine disinfection) is used as a baseline. It is compared to two alternatives: a conventional treatment system with preozonation added and a conventional system combined with point-of-use devices located at selected homes in the service area. Ozone pretreatment has the potential for increasing the fraction of pathogens inactivated. Although ozone pretreatment decreases the concentration of many DBPs, it may increase others, in particular, brominated compounds. When point-of-use water filters are installed in the homes of individuals with compromised immune systems, the benefits are limited to the consumers receiving the filters. It is assumed that all three options operate in a steady-state environment. Table 14.1 summarizes the initial and annual cost of the three alternatives.

Figure 14.3 illustrates the conventional treatment system utilized in this case study, which is typical of current practice for water purveyors in the United States. Figure 14.3 shows the conventional system with preozonation. Ozone pretreatment is an alternate disinfectant for which information on microbial treatment efficiency and DBP production are available. For purposes of this analysis *Cryptosporidium* is the microorganism of interest.

An advantage of the CRFM approach is that it requires the explicit identification of all the assumptions that are part of the analysis. Some of the assumptions embedded in this analysis are as follows.

14.41 Treatment Assumptions

- The in-home filters are assumed to be installed in the dwellings of individuals with compromised immune systems [the acquired immunodeficiency syndrome (AIDS) subpopulation serves as a proxy for this group].
- The in-home filters are assumed to completely remove all microbial agents but have no effect on DBP concentrations.
- While the efficacy of *Cryptosporidium* removal was assumed to be 100 percent for the point-of-use devices, the inactivation efficiency of the baseline treatment and the ozone-supplemented treatment were assumed to be less than 100 percent, based on studies conducted at a pilot facility operated by the U.S. EPA.

TABLE 14.1 Treatment Technology Costs

Treatment	Initial cost	Annual cost
Baseline (chlorination)*	$78,200,000	$7,300,000
Incremental cost with preozonation*	$4,800,000	$400,000
Incremental cost in-home reverse osmosis filter†	Purchase: $500–$800 Installation: $70–$150	$70–$100

*Initial estimated construction costs (not amortized) are given in 1997 dollars without adding interest.

†Lykins et al. (1992, p. 205) estimated the costs (in 1991 dollars) of a reverse osmosis unit for a single tap.

- Concentration data [in micrograms per liter (μg/L)] for individual DBPs in the case study (one alternative with chlorination only, and one with chlorination following preozonation) were adapted from a paper by Miltner et al. (1990), resulting from a study in which Ohio River water was treated in a pilot plant.
- To estimate the formation of bromate by ozone in the Miltner et al. (1990) study, results were employed from another study wherein raw Ohio River water was ozonated in the same pilot-scale contactor (Owens et al., 2000).
- A population of 460,000 people served is assumed.
- The concentration of *Cryptosporidium* oocysts in tap water depends on the removal efficiency of the treatment technology being considered.

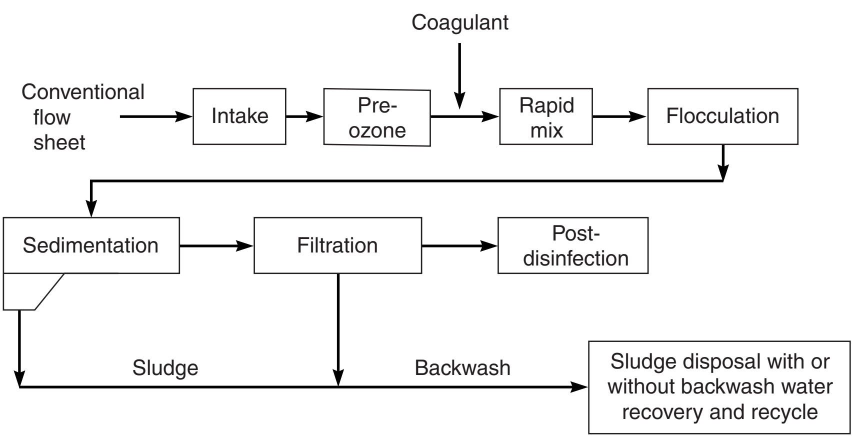


FIGURE 14.3 Schematic of conventional water treatment train with preozonation.

- *Conventional technology.* The removal efficiency for the conventional treatment plant using post-chlorine disinfection was assumed as 2 logs (i.e., the concentration of oocysts in the water is reduced by a factor of 10^2, or 100).
- *Ozone pretreatment alternative.* Ozone pretreatment in addition to the baseline technology was assumed to reduce oocyst concentrations by between 0.5 and 1.5 logs (a factor of 3.16 to 31.6). The precise removal efficiency was assumed to follow a triangular distribution with a lower bound of 0.5, an upper bound of 1.5, and a mode of 1.0.
- *In-home filter alternative.* The in-home reverse osmosis filter systems were assumed to remove all protozoa such as *Cryptosporidium.*

14.4.2 DBP Risk Assumptions

- DBP-induced health effects are assumed to include cancer, developmental toxicity, and reproductive toxicity.
- The individual DBP cancer risks are assumed to be a linear function of the average daily dose.
- A response addition model was used to estimate the mixtures risk for this endpoint.
- It is assumed that DBP-induced cancer is manifest as bladder, colon, and rectal cancer.
- Developmental effects are assumed to be permanent and to result in severe lifetime dependency and a decreased life expectancy.
- Unidentified total organic halides (TOX) values were used to estimate the amount that could be associated with producing a developmental, reproductive, or carcinogenic health risk using Quantitative Structure Activity Relationship (QSAR) predictions (Rice et al., 2001b).

14.4.3 Microbial Risk Assumptions

- The assumed concentration of *Cryptosporidium* in source waters for treatment plants was derived from measurements taken at the intake based on a study of the treatment plant in Trenton, N.J. (LeChevallier et al., 1998).
- Risks of contracting diarrhea and related symptoms, such as severe illness and death, were estimated from the predicted exposures.
- It is assumed that age-specific total tap water consumption rates for the AIDS subpopulation are the same as the corresponding rates for the general population. It is assumed that the age-specific unheated tap water consumption rates for the AIDS subpopulation are 70 percent of the corresponding rates for the general population.
- Probabilistic techniques were used to estimate infection.

14.4.4 Cost Assumptions

- Health effect costs are assumed to depend on three factors: the tap water consumption rate, the incremental probability of an adverse health effect associated with each liter of water consumed, and the cost (measured in lost QALYs) associated with each health event.
- Financial costs considered by the case study are limited to the direct costs of implementing the technologies evaluated consisting of the capital costs necessary for installing the technology and the ongoing operational costs.

14.4.5 Response Addition Model

For this case study, response addition was assumed across chemicals to estimate cancer, developmental, and reproductive human health risks from exposure to a mixture of DBPs found in the distributed drinking water. Response addition has been used for the risk assessment of mixtures of carcinogens (Gaylor et al., 1997; U.S. EPA, 1989). The following equation describes the response addition model used in the case study:

$$R_m = \left(1 + \frac{C_u}{C_k}\right) \cdot Y\left(\frac{\text{L}}{\text{kg·d}}\right) \cdot 0.001\left(\frac{\text{mg}}{\mu\text{g}}\right) \cdot \sum_{i=1}^{n} S_i\left(\frac{1}{\text{mg/kg·d)}}\right) \cdot C_i\left(\frac{\mu\text{g}}{\text{L}}\right) \tag{14.1}$$

where n = number of DBPs in mixture known to cause a specific effect
R_m = total mixtures risk from both n known DBPs and unidentified TOX in mixture
C_i = concentration of the ith DBP in mixture, μg/L
C_u = total concentration of measured but unidentified TOX in mixture associated with causing a specific effect, μg/L
C_k = sum of concentrations of n DBPs in mixture known to cause a specific effect, μg/L
Y = daily human tap water consumption per body weight, L/kg·d
S_i = slope factor for humans for ith DBP, (mg/kg·d)$^{-}$

The final risk estimate R_m for the entire mixture is the sum of the risks from the DBPs known to cause a specific effect and from the unidentified TOX associated with that effect. The total risk is reflected by the first term on the left side of Eq. (14.1), $(1 + C_u/C_k)$. This term is multiplied by the total risk for the known DBPs, which is the product of the other terms on the right side of Eq. (14.1). This calculation effectively sums the risk from the known DBPs and a scaled value of that known risk equal to the ratio of the concentration of unidentified TOX to the total concentration of known DBPs (C_u/C_k). To characterize the distribution of plausible values for R_m, distributions were developed for each of the parameters C_i, C_u, S_i, and Y.

14.4.6 Application of QALY Concept

The following list briefly summarizes the basic approach used to assign values to each of the health endpoints considered in the case study:

- *DBP-induced cancer illness.* Treatment costs per year are the difference between a typical year of life and the value of a year of life undergoing cancer treatment (0.5). The treatment is assumed to last 2 years.
- *DBP-induced cancer death.* The QALY cost of this event is the value of the life years lost.
- *DBP-induced reproductive toxicity.* The cost of a single year of infertility is assumed to be 0.12 QALY. This cost is assumed to affect both males and females. This cost is assumed not to apply to the AIDS subpopulation.
- *DBP-induced developmental toxicity.* Only the cost to the child to be born is considered. It is assumed that this individual will suffer a decreased quality of life due to severe dependence stemming from the developmental toxicity. The cost of this dependency is the difference between the value of a typical year of life and the value of a year of life in a state of severe dependence (0.55 QALY). It is also assumed that this individual suffers a decreased length of life with a life expectancy of 50 years. It is assumed that members of the AIDS subpopulation do not have offspring; hence, this cost is assumed not to apply in their case.
- *Mild GI illness.* For the general population, the cost of a year of GI illness is assumed to be the difference between the value of a typical year of life and the value of a year of life in severe pain (0.64 QALY). A mild illness is assumed to last 1 day and, hence, has a cost equal the cost of a year of illness divided by 365 days. For members of the AIDS subpopulation, this cost is assumed to equal one-half the average cost for members of the general population.
- *Moderate to severe illness.* This condition is assumed to last 14 days; its cost is therefore assumed to be 14 times the cost of a mild GI illness.
- *Death due to microbial infection.* The QALY cost of this event is the value of the life years lost.

Tables 14.2 and 14.3 summarize the QALY costs for these health endpoints for the general population. Specifically, the QALY costs in Tables 14.2 and 14.3 represent the net present value of the lost QALYs due to the health effect at the time the effect occurred. Tables 14.2 and 14.3 also list the corresponding values for the AIDS subpopulation. Note that although the QALY cost is the same for death resulting from cancer and death resulting from microbial infection, the latter is, in practice, of greater

TABLE 14.2 Results Summary: 3 Percent Discount Rate—General Population of 460,000 Individuals and AIDS Subpopulation of 429 Individuals—20-Year Analytical Time Frame

		NPV of total incremental cost, $	NPV of total incremental QALYs	Expected CE ratio, $ per QALY
Ozone	General population*	10,800,000	5,372	4,134‡
	AIDS population	10,000	7,684	2.27
	Total population	9,780,000†	10,420	1,532
Home filters	AIDS subpopulation	1,130,000	11,636	152

*General population results represent the average (i.e., expected value) for the average tap water consumer.

†The total population cost is rounded and hence does not equal the sum of general population and AIDS subpopulation costs. This rounding is consistent with the imprecision in the estimated size of these two groups.

‡The average CE ratio does not equal the incremental technology cost divided by the average QALY gain.

TABLE 14.3 Results Summary: 5 Percent Discount Rate—General Population of 460,000 Individuals and AIDS Subpopulation of 429 Individuals—20-Year Analytical Time Frame

		NPV of total incremental cost, $	NPV of total incremental QALYs	Expected CE ratio, $ per QALY
Ozone	General population*	9,780,000	4,519	4,235‡
	AIDS population	9,120	5,901	2.71
	Total population	9,780,000†	10,420	1,719
Home filters	AIDS subpopulation	997,000	8,746	178

*General population results represent the average (i.e., expected value) for the average tap water consumer.

†The total population cost is rounded and hence does not equal the sum of general population and AIDS subpopulation costs. This rounding is consistent with the imprecision in the estimated size of these two groups.

‡The average CE ratio does not equal the incremental technology cost divided by the average QALY gain.

consequence because there is effectively no latency period in the case of microbial illness. Cancer, on the other hand, is generally manifest many years after an exposure; its cost must therefore be appropriately discounted to reflect this delay. Clearly from Tables 14.2 and 14.3 the use of an intervention (ozone) that increases cancer risk slightly but substantially reduces microbial risk is a cost-effective strategy.

The uncertainty associated with these estimates has not been quantified, even though it is clear that these values are not known precisely. To assess the potential importance of this uncertainty, it is assumed that the plausible range of QALY cost values for each health endpoint is log-uniformly distributed between half the point estimate derived in this section and twice this point estimate. It is clear that in this example the application of ozone for protection against microbial infection clearly outweighs any negative effect associated with the formation of DBPs.

14.5 SUMMARY AND CONCLUSIONS

A major objective of drinking water treatment is to provide safe drinking water. Conventional drinking water treatment combined with disinfection has proven to be effective in achieving this objective and is often considered to be one of the major public health advances in modern times. In the United States, chlorine which is the disinfectant most often added to treated water for microbiological protection poses the dilemma of a risk tradeoff. Chemical disinfection reduces risk of infectious disease, but the interaction between chemical disinfectants and precursor materials in source water results in the formation of DBPs. Frequently, environmental exposures pose these types of problems.

Often, the objective of the risk assessor working in collaboration with the risk manager is to make choices between alternatives or to find the "optimal" decision point that minimizes the overall risk to exposed populations.

The EPA's National Center for Environmental Assessment (NCEA) and National Risk Management Research Laboratory (NRMRL) have initiated research into a concept called the Comparative Risk Framework Methodology (CRFM). It provides a systematic way to assess interdependent environmental health risks and compares the impact of alternative interventions designed to minimize these risks. Interventions are defined as protective measures either in place or proposed that interdict between a potentially harmful agent and a susceptible population. In this context water treatment could be considered an environmental health intervention designed to protect human populations against potentially pathogenic microorganisms in drinking water while minimizing their exposure to potentially harmful disinfection by-products.

To illustrate this application the CRFM has been applied, on a case study basis, to the comparison of three water treatment strategies for protecting against microbial contamination while minimizing exposure to DBPs. One of the major advantages of the CRFM is that it requires that each assumption that underlies the analysis is explicitly identified. The NCEA and NRMRL are currently extending this concept to exposures associated with drinking water distribution systems.

14.6 ACKNOWLEDGMENTS

The authors would like to acknowledge the assistance of Ms. Bette Zwayer and Mr. Merrel Robinson of the EPA's National Center for Environmental

Assessment and Mr. Steven Waltrip of the EPA's Water Supply and Water Resources Division.

14.7 REFERENCES

Bellar, T. A., Lichtenberg, J. J., and Kroner, R. C. (1974). "The occurrence of organohalides in chlorinated drinking water." *JAWWA,* 66(12):703–706.

Gaylor, D. W., Axelrad, J. A., Brown, R. P., et al. (1997). "Health risk assessment practices in the U.S. Food and Drug Administration." *Reg. Toxicol. Pharmacol.,* 26:307–321.

Gold, M. R., Russell, L. B., Siegel, J. E., and Weinstein, M. C. (1996). *Cost-Effectiveness in Health and Medicine.* New York: Oxford University Press.

Haddix, A. C., Teutsch, S. M., Shaffer, P. A., and Dunet, D. O. (1996). *Prevention Effectiveness: A Guide to Decision Analysis and Economic Evaluation.* New York, Oxford: Oxford University Press.

LeChevallier, M. W., Norton, W. D., Abbaszadegan, M., and Atherholt, T. B. (1998). *Short-term Variability of Giardia Cyst and Cryptosporidium Oocyst Concentrations in a Surface Water Source Used for Potable Water.* New Jersey Department of Environmental Protection Division of Science and Research, April.

Lykins, B. W., Clark, R. M., and Goodrich, J. A. (1992). *Point-of-Use/Point-of-Entry for Drinking Water Treatment.* Chelsea, Mich.: Lewis Publishers, Inc.

Miltner, R. J., Rice, E. W., and Stevens, A. A. (1990). "Pilot-scale investigation of the formation and control of disinfection byproducts," *1990 AWWA Annual Conference Proceedings,* Cincinnati, Ohio, 2:1787–1802.

Owens, J. H., Miltner, R. J., Rice, E. W., et al. (2000). "Roto pilot scale ozone of *cryptosporidium* and other microorganisms in natural water." *Ozone Sci. Eng.* 22:501–517.

Putnam, S. W., and Graham, J. D. (1993). "Chemicals versus microbials in drinking water. A decision sciences perspective." *JAWWA,* 85(3):57–61.

NAS (National Academy of Sciences). (1986). *Committee on Drinking Water, Drinking Water and Health.* Washington, D.C.: National Academy Press, 6:457.

Rice, E., Murphy, P. A., Teuschler, L. K., et al. (2001a). "U.S. EPA's comparative risk framework" in *Microbial Pathogens and Disinfection By-products in Drinking Water: Health Effects and Management of Risks.* Craun, G. F., Hauchman, F. S., and Robinson, D. E. (eds.), Washington, D.C.: ILSI Press, pp. 371–394.

Rice, E., Teuschler, L. K., Moudgal, C., et al. (2001b). "Risk assessment of complex mixtures of disinfection by-products (DBPs): methods for considering unidentified DBPs" in *Microbial Pathogens and Disinfection By-products in Drinking Water: Health Effects and Management of Risks.* Craun, G. F., Hauchman, F. S., and Robinson, D. E. (eds.), Washington, D.C.: ILSI Press, pp. 571–580.

Rook, R. J. (1974). "Formation of haloforms during chlorination of natural waters." *Water Treat. Exam.* 23(2):234–243.

U.S. EPA. (1989). *Risk Assessment Guidance for Superfund,* Vol. 1, Part A. EPA/540/1-89/002.

U.S. OMB (Office of Management and Budget) (1996a). *Economic Analysis for Federal Regulations under Executive Order 12,* January 11, 1996, p. 866.

U.S. OMB (Office of Management and Budget) (1996b). *Guidelines and Discount Rates for Benefit-Cost Analysis of Federal Programs,* Circular No. A-94, Revised. February 6, 1996.

P • A • R • T • 4

CASE STUDIES

CHAPTER 15
ADAPTIVE MANAGEMENT FOR WATER SUPPLY PLANNING: SUSTAINING MEXICO CITY'S WATER SUPPLY

Richard L. Skaggs
Lance W. Vail
Steve Shankle
Pacific Northwest National Laboratory
Richland, Washington

15.1 INTRODUCTION

Modern water resource managers are constantly required to balance multiple, conflicting, incommensurate objectives in an environment characterized by high levels of uncertainty, varying data quality and availability, and competing models and approaches. The reliability of water resources management policy and decisions depends on the ability of measurements, response models, process models, and policy models to interact with each other across the variety of temporal and spatial scales each represents. It also requires a cautious, probing, adaptive approach founded on fundamental economic principles, the success of which depends upon improved understanding, predictive accuracy, and iterative performance assessment.

Over the past 2 decades, adaptive management has emerged as the advocated approach to resolve natural resource conflicts in the face of significant uncertainty. Adaptive management is a systematic and rigorous, scientifically defensible program of learning from the outcomes of management actions, accommodating change, and improving management (Walters and Holling, 1990). While advocated, adaptive management has proven difficult to implement due to (1) its direct and indirect costs (e.g., the costs of improved understanding through research and the political risks of potentially having clearly identified failures are two of the barriers to its use), and (2) the lack of readily available, widely implemented tools for resolving critical natural resource management issues (scientific methods for analyzing, understanding, and managing problems within complex and anthropogenically altered environmental systems).

Central to the natural resource management challenge is ensuring the sustainability of water resources in the face of expanding domestic, industrial, and recreational demands. It was noted by Mays (1996) and the National Research Council (NRC, 1999) that the continued viability of historical approaches to development, industrialization and resource use are in question. Problems today are couched in terms of preserving and/or restoring the physical, chemical, and biological integrity of our natural resources. And the complexity of the planning process is compounded by rapidly changing demographics and urban development, uncertain impacts of global and regional climate change, and overallocated water resources. Further there is the growing expectation on the part of decision makers and the public that alternative water resource management strategies integrate multiple dimensions including, among others, economic, ecological, social, as well as engineering requirements.

The purpose of this chapter is to discuss the applicability of the adaptive management process to water supply planning. The historical development of adaptive management and its key components are summarized. The barriers to its successful implementation and the commensurate benefits are outlined. Finally, a case study is presented wherein the necessity for and the value-added of an adaptive management approach is used to evaluate alternative solutions for ensuring the long-term viability of the principal water supply for Mexico City.

15.2 ADAPTIVE MANAGEMENT

One of the original articulations of the adaptive management concept was that of C. S. Holling (1978). It was presented as an alternative approach to environmental impact assessment and management with an essential notion being the integration of environmental with economic and social understanding throughout the design process and beyond implementation. One of the problems it was intended to address was that the fundamental properties of a development or policy are firmly established early in the design process such that significant change can only be introduced in the process through conflict and extraordinary pressure from interest groups. Adaptive management, then, was conceived as a systems approach building on ecology, physical sciences, and systems sciences. It is experimental in nature, utilizing trial and error and allowing for both the known and unknown. Its goal is to embrace uncertainty, accepting partial understanding of processes, and produce policies and designs that are less sensitive to the unexpected. These ideas are captured well in a more recent definition of adaptive management contained in Sit and Taylor (1998): "Adaptive management is a systematic process for continually improving management policies and practices by learning from the outcomes of operational programs. Its most effective form—'active' adaptive management—employs management programs that are designed to experimentally compare selected policies or practices, by evaluating alternative hypotheses about the system being managed." It is further stated that "active adaptive management can involve deliberate 'probing' of the system to identify thresholds in response and to clarify the shape of the functional relationship between actions and response variables."

Adaptive management is generally described as a recursive process consisting of four main steps: (1) plan, (2) act, (3) monitor, and (4) evaluate as shown in Fig. 15.1 (Rauscher, 1999; British Columbia Forest Service, 2000). Each of these steps, as they relate to water supply planning, is discussed below.

15.2.1 Plan

Planning most often consists of the steps conventionally included in water supply planning such as problem identification, alternative development, and alternative assessment. Planning efforts begin with a means or process to identify or highlight a set of problems, needs, or issues to be addressed. Once needs are identified, available data and information are assembled and analyzed leading to development of a range of preliminary alternative solutions. The adaptive management approach suggests that these alternatives should be designed to explicitly state and

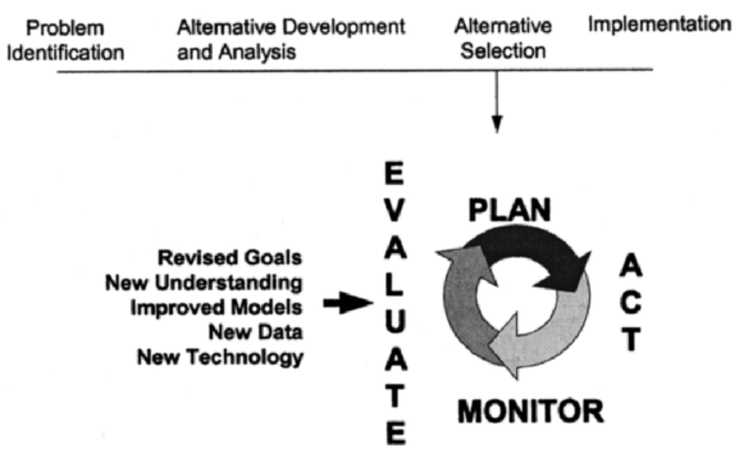


FIGURE 15.1 Adaptive management process. (From Rauscher 1999 and BC FS, 2000)

subsequently test experimental hypotheses related to test assumptions and projections about the performance of management actions. Inherent in developing this list is the need for decision makers to articulate the overall objectives of the planning effort. Identifying the objectives is essential to establishing the metrics for success of any actions ultimately taken. Selected metrics must be relevant to the objectives and responsive to potential management actions over a range of time scales and relevant spatial scales (e.g., site, watershed, region, or basin). Conceptual- and process-based simulation models can then be used to investigate cause-effect relationships between the candidate actions and metrics.

With project objectives in mind, it is at this early phase in the process that the philosophy of adaptive management must be embraced. An outcome from the analysis of available information should be a measure (qualitative or quantitative) of uncertainty and associated risks for the various alternatives. Uncertainty may be the result of our inability to forecast variables such as climate, population, or industrial growth. The uncertainty can derive from our inability to quantify key physical and biological processes affected by the alternative measures. For example, despite the ever increasing spatial resolution of remote sensing imagery of land cover, our ability to predict the effects of land-use change on adjacent instream water quality (e.g., temperature, sediment load, flow velocity, and depth) tends to be comparative and imprecise. We can speculate with some confidence that clear-cutting tends to increase runoff and erosion, but it is difficult to determine the amount of runoff and sediment attributable to a given land-use change scenario. Or, sensitivity studies may show that results are relatively insensitive to assumptions about a given parameter. Thus, in the spirit of adaptive management, the identification of uncertainty and the "unknown" at the planning phase should not necessarily lead to massive data collection efforts in lieu of action. But instead, the planning effort should evaluate incremental, potentially robust alternatives, that through their implementation and subsequent monitoring, would provide feedback on the impacts of change on processes, fill information gaps, and demonstrate the relative effectiveness of selected alternatives.

15.2.2 Act

The BCFS (1999) discusses the difference between "passive" and "active" adaptive management. Frequently the default approach is one of *passive* adaptive management such that an action plan is developed around the "best" forecast outcome, hypothesizing that the analysis results are true. *Active* adaptive management refers to the development of action plans that are explicitly designed as management experiments with a significant objective being hypothesis testing and learning to fill in knowledge gaps. The power in adaptive management is investment in management experiments for compounded benefits in the form of reliably improved future decisions. The active approach suggests consideration of multiple alternatives. Primary considerations in alternative selection should be based on comprehensive tradeoff analyses related to the various long-term objectives (i.e., economic, social, political cost/benefits) as well as risks of negative outcomes with a bias toward small-scale and robustness. Another important consideration is the potential for filling key knowledge gaps that might lead to improved solutions at multiple locations in the future. To this end, where possible, multiple alternatives should be selected, perhaps at pilot scales with replications and controls to establish statistical independence, minimize bias, etc. Concepts and theories underlying this approach are described in detail by Sit and Taylor (1998).

15.2.3 Monitor

Monitoring programs must be designed to support testing of hypotheses associated with the selected alternative(s). Monitoring results and their subsequent evaluation are the basis for management improvement. Important considerations are the frequency, timing and duration, and spatial scales for monitoring, both pre- (as baseline) and postimplementation. It should also be determined what level or degree of response is important to subsequent management assessments of the given alternatives to ensure that the monitoring program has sufficient detection sensitivity and resolution. Sit and Taylor (1998) discuss three types of monitoring: (1) implementation monitoring to support operations and determine if all management guidelines and regulations have been followed, (2) effectiveness monitoring to provide the basis for evaluating whether management actions are producing the desired results, and (3) validation monitoring to evaluate the correctness of basic underlying assumptions.

15.2.4 Evaluate

As shown in Fig. 15.1, the evaluation stage closes the loop on each iteration of the adaptive management process. It is during the evaluation phase that actual results from the implemented management actions are compared to forecasts made during detailed planning for incorporation into future decisions. An important aspect of this stage is deconvolving cause and effect between results that are understood and attributable to the management actions and results that are not understood and outside management control. This suggests going beyond simply detecting trends to the point of discerning the degrees of influence in cause-effect relationships (e.g., R^2, coefficient of determination). To a large extent success in doing this is dependent upon the statistical design of the implementation and the monitoring systems. It is also important to be careful in selecting appropriate indicators, exploiting all available quantitative and qualitative information, and to utilize statistical approaches that are robust with respect to sample size. It is also emphasized that in the long term, negative as well as positive (i.e., predicted) outcomes are valuable.

15.3 ESSENTIAL ELEMENTS FOR APPLICATION OF ADAPTIVE MANAGEMENT TO WATER SUPPLY PLANNING

Water supply systems are inherently complex. Along with their complexity are limitations in our ability to characterize and/or predict state variables and parameters associated with key processes, uncertain driving functions, etc. With respect to these uncertainties and associated risks, an important aspect of the power behind the adaptive management process is its basic faculty for acknowledging and embracing uncertainty. Other elements associated with adaptive management that need to be emphasized in discussing its application to planning and development of water supplies are that decision goals and objectives be economics based, with a bias toward robust solutions. When uncertainties are extreme, preference should be given to incremental advancement toward long-term goals and objectives (i.e., try something and see what happens) to avoid the default "do nothing" alternative. Linked to the idea of moving forward, even incrementally, in the face of uncertainty is the concept of *value of information* (VOI). Each of these elements is elaborated upon below.

15.3.1 Economics Based

An imperative to successfully implementing adaptive management is establishing concrete program goals and objectives. A subelement to articulating goals and objectives is identifying the criteria against which alternatives will be evaluated for selection. Given the diversity of interests and stakeholders involved in planning and implementing modern water system projects, there will undoubtedly be multiple and diverse criteria under consideration. As a reminder (at the risk of stating the obvious), the suggestion here is that the criteria selected should be solidly founded in economics-based reasoning. Katz and Koss (2000) make a similar argument for addressing salmon recovery in the Pacific Northwest and suggest several decision rules for designing salmon recovery measures in the face of uncertainty. Economic-based rules equally applicable to water supply planning include:

- Socioeconomic benefits must be realistically assessed, while costs must accurately reflect the true societal value of dedicated, scarce resources. Consequently, alternatives with benefits exceeding costs should be viewed as candidates for further consideration while those with costs exceeding benefits should be discarded.
- Measures should be ranked according to their cost-effectiveness such that if two or more measures produce equivalent outcomes, the least-cost measure should have higher priority consideration. Costs in this context include whole life costs (i.e., life cycle capital and operating expenses) as well as accurate costs of scarce resources (land, water, water quality, etc.).

- Opportunity costs must be explicitly considered. These should include not only alternative uses of funding required for a particular water resources project, but the alternative uses of the scarce resources as well.
- Inherent in embracing uncertainty is the acceptance of failure for the sake of learning. Given failure then, one must be prepared to invoke the notion of sunk costs. *Sunk costs* are past expenditures that are irretrievable. If an alternative proves to be ineffective (i.e., socioeconomic costs are greater that benefits) from the standpoint of future costs, it should be abandoned.

15.3.2 Robustness

Closely related to the economics concepts discussed above, Katz and Koss (2000) also discuss ideas suggesting the value in a bias toward "robustness" when evaluating alternatives. The Encarta web-based dictionary gives a definition for the adjective *robust* as "capable of recovery: able to recover from unexpected conditions during operation." Clearly, striving for this aspect of robustness in spite of uncertainties and unknowns is a goal of the adaptive management process. More specifically, these ideas are

- Future options have value and should be preserved to the extent possible. If feasible, selections should be avoided that preclude alternatives that currently do not measure up, but could do so with additional information or potentially emerging new technology.
- One such option is that of reversibility. Alternatives having the potential for destroying a nonrenewable resource or asset should be viewed as inferior to those that avoid such irreversible costs assuming the total cost picture is acceptable.
- In conducting cost-benefit analyses, consideration should be given to the cost of inflexibility associated with large-scale projects. Given projects of equivalent cost-benefit ratios, it should be recognized that smaller measures developed incrementally offer inherently more flexibility and opportunity to make future decisions with less uncertainty while exploiting the benefits of technology creep. Conversely, large projects reduce planning flexibility and are more difficult to treat as sunk costs if found to be ineffective.

15.3.3 Value of Information

An important concept inherent in adaptive management is that of value of information. An engineering view of *value of information* is simply the ability of additional information to reduce uncertainty in design variables, thereby accruing cost savings from a less conservative (i.e., lower cost) design. The value of additional (new) information can be exploited in a number of ways. At one extreme is the use of new information merely as data in a purely statistical context (i.e., least squares analysis). This is the minimal use of new information wherein each new piece of data simply increases the degrees of freedom of the statistical model by one, with only marginal impact on standard errors of estimation and a minimal increase in understanding. At the other end of the spectrum, information can be interpreted in the context of formulated models that fully utilize a priori information in representing process knowledge. In this context, new information can contribute to reducing uncertainty of predictions on multiple fronts such as model formulation improvement, enhancing model calibration, and/or model validation.

A powerful element of the adaptive management approach is its reliance on value of information in making investments for new data. The concept enables a decision maker to fully utilize available information to analyze alternative solutions with full consideration given to the level and sources of uncertainty in the analysis. If those uncertainties impose excessive conservatism and therefore cost burden on the preferred alternative(s), then a clear tradeoff can be established on the costs (direct and indirect) of collecting new information, and the resultant benefits. Costs can include the actual costs of data collection, management, and analysis as well as indirect costs such as the impacts of delaying action until new data are obtained.

Benefits can also be viewed as direct or indirect. Direct benefits include avoided costs associated with project failures plus the savings derived from reduced uncertainty and design factors of safety. Indirect benefits are achieved principally from the increased process knowledge gained from new information and extension of that knowledge to future management decisions in the same water system or to similar systems.

To illustrate this further, consider an important activity associated with adaptive management, which is postaction monitoring. There are two important aspects to the value of such monitoring. The first is increased understanding of the cause-effect relationships (i.e., what effect does a given management action have on the system in question?). The second is monitoring the degree of success or failure of the management action. Both aspects have considerable value in guiding future management actions to minimize their cost and maximize their effectiveness. Focusing on the latter as a way of illustrating the value (and justifiable cost) of monitoring, consider two identical projects, one implemented with monitoring, one without. Using a decision tree analysis, the justifiable cost for monitoring can be roughly estimated to be the probability of project failure times the sum of the project operating cost plus the estimated damage associated with a failed project (see Holling, 1978). Extending this concept to projects with similar objectives within a region, it can be seen that monitoring the success or failure of a given project provides information useful to the design and operation of future projects. For example, if postproject monitoring detects project failure or significantly less than expected success, several benefits accrue:

- The failing or less than successful project can be aborted, avoiding continued operating costs as well as minimizing continued damage associated with the failure.

- An alternative project can be implemented with an increased probability of success.
- All future projects with similar objectives using potentially similar technologies will have a lower probability of failure.

From this highly simplified example it can be seen that the costs and benefits of monitoring for a given project can be estimated. Value is quantifiable in terms of detecting failures and minimizing unanticipated results and/or impacts. When implemented on a regional basis, the value of additional information is multiplied many times over. And finally, the value of information (when collected in a scientifically sound manner) is compounded further by advancing the state of understanding about cause-effect relationships between natural and anthropogenic changes and project performance, enabling true adaptive management and improved future management decision making.

15.4 CASE STUDY

To illustrate the broad versatility and applicability of adaptive management concepts to water supply planning, an analysis focused on the sustainability of Mexico City's water supply is presented. The Mexico City problem confronts considerable uncertainty in multiple dimensions (e.g., prediction of state variables and lack of knowledge about key processes). And, the nature and scale of the problems result in a large number of potential alternative solutions.

15.4.1 Mexico City Water Challenge

The federal district of Mexico City has recognized that existing water management practices in the federal district are insufficient to adequately provide for a long-term, safe supply of water. Current water supply estimates are that 51 m^3/s are extracted from the Valley of Mexico Aquifer and another 23 m^3/s are imported via surface conveyance from outside the city. Since early in the twentieth century, there has been a deficit in the balance between extraction and recharge to the aquifer. While precise measurements of the deficit are unavailable, field measurements show that for decades the deficit has produced an average decline in the water table of approximately 1 m per year (NRC, 1995). Currently the deficit is estimated to be on the order of 17 m^3/s.

While "mining" of the aquifer poses no immediate threat to the reliability of the city's water supply (i.e., there is a vast quantity of water remaining in the aquifer), future sustainability of the water supply is undermined by long-term risks associated with aquifer water quality degradation and/or losing access to imported water. Mass balance studies reported by the NRC (1995) suggest that at current depletion rates, the aquifer storage volume is equivalent to 200 to 300 times the annual draft.

However, the NRC further notes several factors associated with overpumping that could significantly degrade the reliability of the aquifer as Mexico City's principal water supply. As the clay layer overlying the aquifer is dewatered and subsequently consolidates and fractures, the potential for contamination of the aquifer from surface and near-surface activities (e.g., leakage from the sewer collection system) increases. It is also noted that deep well tests indicate potential geology-derived water quality problems at deeper levels within the aquifer. Equally important are the severe city-wide impacts from subsidence associated with continued drawdown of the aquifer.

There are many options that the federal district of Mexico City could consider for strengthening water management practices to help assure the long-term supply of water at an acceptable price (e.g., aquifer enhancement, water system performance improvement, and residential and industrial demand reduction). A collaborative study was conducted for the federal district by Battelle Memorial Institute, Instituto Mexicano del Petróleo, and the Universidad Autónoma Metropolitana to prepare an integrated strategy that identifies a comprehensive suite of feasible options and lays out a program for their implementation (Battelle, 2001). Given the scale and potential acuteness of the problems to be addressed and considerable uncertainties in understanding of present and future stresses and demands on the water system, an adaptive management approach was adopted for the planning effort. The study was based almost entirely on existing data. Even though the aquifer and the water management system have been the focus of a number of past studies, there do remain gaps in understanding of the aquifer, both in how it behaves as a natural system and in how human activity has affected it. However, the situation with the water supply is reaching what some have called a crisis level. Therefore, an adaptive approach will be necessary to move ahead with demonstration projects and engineered solutions as quickly as is reasonably possible before all the desired data are available.

15.4.2 The Problem

The challenge of providing high-quality water at a reasonable price to all the residents of Mexico City will only grow in size and complexity in the coming years. While many have noticed the warning signs—the increase in subsidence-related problems and increasing costs of maintaining the infrastructure for water—the real impacts of the water crisis have yet to be felt. Accordingly, most in the public have not placed the coming water crisis in a sufficiently central position to assure prompt and effective action. Solving the water crisis will not be easy or inexpensive. However, nothing significant can be done unless the public and decision makers come to understand the significance of the crisis and the potential for its solution. Many different water management actions (e.g., leak reduction, demand reduction, and enhanced natural recharge) have been advanced as the potential solution to the whole problem. However, the facts do not support the efficacy of simple solutions. What is needed is an integrated and aggressive water management program in order to assure the long-term availability of water.

The first question to be addressed is, "What is the magnitude of the water crisis?" Most people in the city have access to sufficient water of adequate quality today. However, there are many indicators that the current water system is already under stress. In addition to the subsidence problems brought on by excessive drawdown of the aquifer, there are increasing problems with contamination of wells in some parts of the valley. There is also increased regional competition for the same water resources, concern on the part of industry that necessary sources of quality water will not be available in the long term, the potential that some of the federal subsidies that currently make water affordable to local citizens will be withdrawn, and the stubborn fact that some residents are not currently served by the water system.

The second question is, "Is this situation sustainable?" The answer is clearly no. A highly simplified conceptual model of the city's water system is shown in Fig. 15.2. Currently about 23 m^3/s are imported from outside the city, and about 51 m^3/s are withdrawn from the aquifer (of which 42 m^3/s are delivered to the water distribution system). Of the 65 m^3/s input to the water distribution system, approximately 40 percent is unaccounted for (i.e., lost to leaks, unauthorized use). Ultimately, on the order of 40 m^3/s wastewater is "exported" outside the city to the north and used primarily for crop irrigation. Focusing on the aquifer itself, Fig. 15.3 presents the water balance for the aquifer based on current estimates of withdrawals and recharge components. The major components consist of 25 m^3/s of natural recharge, approximately 9 m^3/s of recharge derived from irrigation

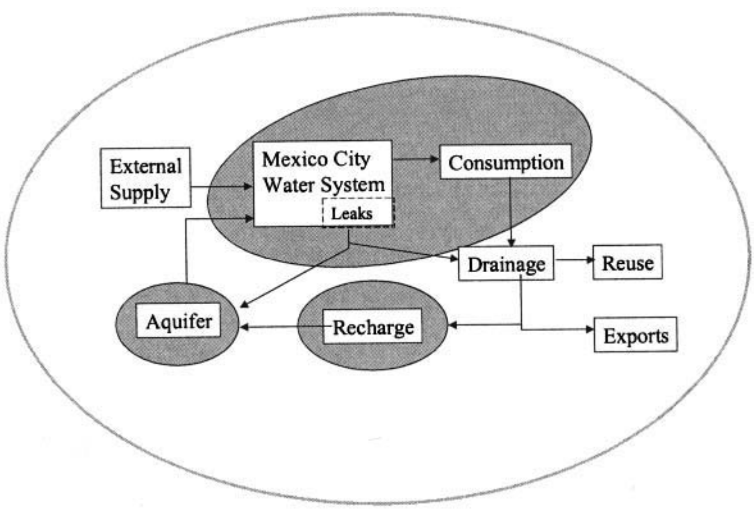


FIGURE 15.2 Simplified conceptual model of Mexico City's water system.

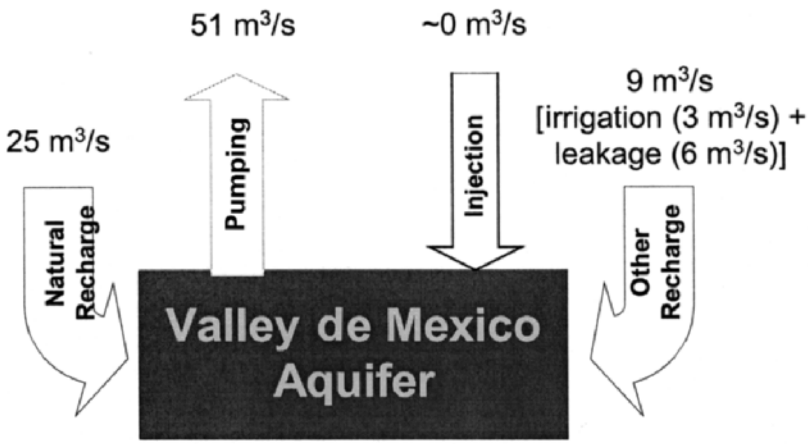


FIGURE 15.3 Water balance for the Valley of Mexico Aquifer (based on 1999 estimates). (Shankle, et al., 2002)

within the city and leakage from the water distribution system, and finally pumping from the aquifer at an average rate of 51 m^3/s.

While large reserves of water remain in the aquifer, the current level of exploitation will almost certainly result in growing subsidence problems in Mexico City and potentially a decrease in water quality as well. More importantly, unless aggressive action is taken, the stress on the aquifer is likely to accelerate in magnitude and effects.

15.4.3 Problem Assessment

Several water system development scenarios were analyzed as the basis for evaluating the magnitude of Mexico City's long-term water management challenges. The scenarios were conceived to systematically consider the sustainability of the city's water supply in terms of overall water balance and water system performance. There were several important findings from these scenarios. First, it was determined that the projected stress on the aquifer is significant and will become unsustainable in the near future. While most people currently have access to sufficient water of adequate quality, with even modest growth rates over the next 20 years the overdraft of the aquifer would continue to increase. For example, taking a very conservative view of the future wherein Mexico City experiences virtually no population growth, and only a 2 percent annual economic growth rate over the next 20 years, the current rate of overexploitation of the aquifer can be expected to grow by 2.5 times during this period. As illustrated by Fig. 15.4, the current deficit of 17 m^3/s can be expected to grow to over 30 m^3/s in 2010 and almost 46 m^3/s in 2020. While all

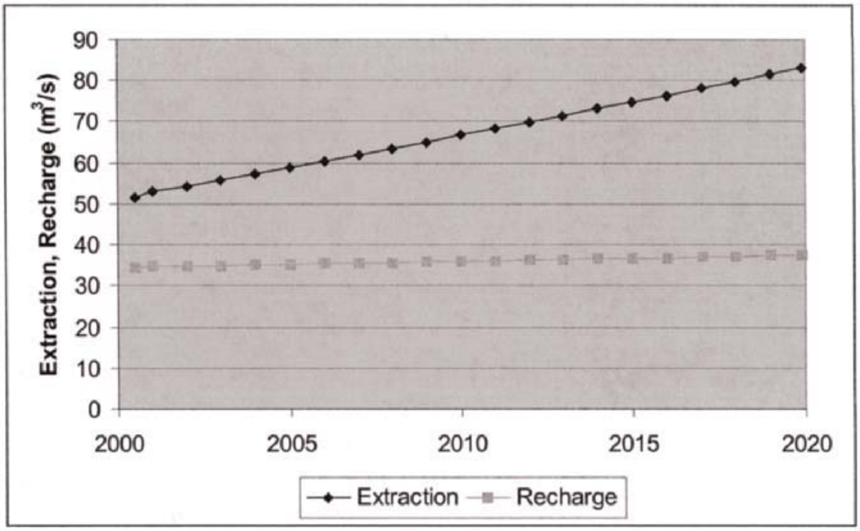


FIGURE 15.4 Aquifer extraction and recharge: 2 percent economic growth.

the consequences of this level of annual drawdown are not precisely known, it can be anticipated that significant subsidence and water quality problems will result, and that water shortages will be likely.

Second, it was determined that only an integrated, multifaceted approach to addressing the water problem was feasible. No single intervention (e.g., demand reduction, leak reduction, enhanced natural recharge, or artificial recharge) could by itself balance the aquifer over the long term. Nor is it apparently feasible to meet the city's growing needs with enhanced imports from other watersheds. However, each of the interventions considered holds the potential for contributing significantly to the overall solution of the water problem. For example, an analysis was conducted with the same conservative growth assumptions as above and the following set of mitigative actions over a 20-year period:

- Increase in the amount of natural recharge of the aquifer using engineered solutions for bracketing purposes
- Reduction in the use of potable water by industry by 50 percent (substituting treated wastewater)
- Reduction in potable water use by households of 3 percent
- Reduction in the amount of leakage in the potable water system (from 37 percent of volume to 20 percent of volume).

Results show that even with these actions, the projected overdraft of the aquifer would continue, achieving a level of approximately 23 m^3/s by 2020. Extending the analysis to include the addition of artificial recharge on the aquifer water balance, it was determined that a system to inject treated wastewater was a critical component of a potential long-term solution. Figure 15.5 shows the projected aquifer water balance over 20 years with the additional assumption that treated wastewater will be injected into the aquifer at a rate of 10 m^3/s within 10 years. The increased recharge provided by the injection activity reduces the deficit to less than 2 m^3/s by about 2010 and reduces the projected deficit in 2020 to just over 10 m^3/s.

Mexico City's water challenges have been developing for a long time. They are driven by factors largely outside of any one agency's control: geography, geology, internal migration, and economic growth patterns are just a few of the factors that have contributed to the current set of problems facing the city's water supply. Just as no one factor is responsible for the current state of affairs, it is unlikely that a single answer will resolve the problems facing the system.

A number of potential solutions to Mexico City's water situation were examined. Solutions considered include

- Aquifer enhancement
  - Injection of treated wastewater into the aquifer
  - Enhanced natural recharge
- Improved system performance
  - Leak reduction
  - Demand reduction

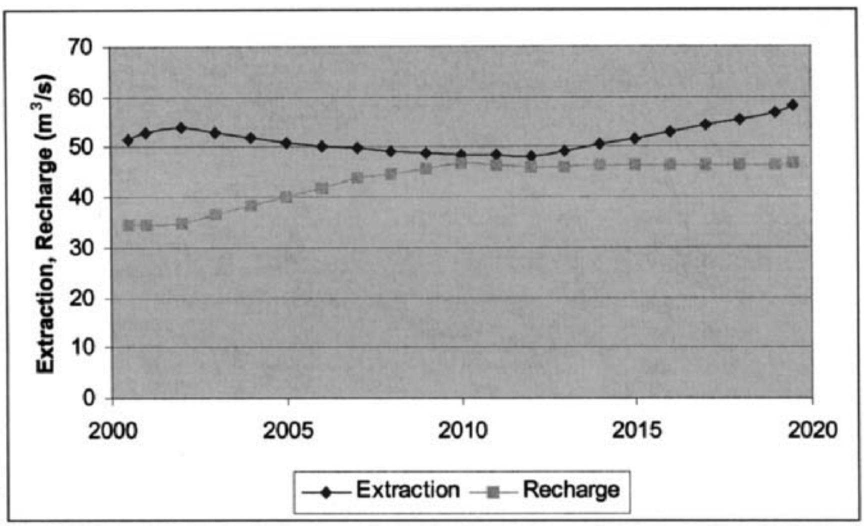


FIGURE 15.5 Mexico City aquifer extraction and recharge: 2 percent economic growth with 10 m^3/s injection (i.e., artificial recharge).

- Residential demand reduction
- Industrial demand reduction
- Nonpotable reuse

It is acknowledged that the ultimate solution will integrate aspects of one or all of these actions; however, discussing planning concepts for each of these actions is beyond the scope of this chapter. For the purposes of illustrating the application of the adaptive management approach, a more detailed description of the planned program for wastewater treatment and injection is presented.

15.5 AQUIFER ENHANCEMENT THROUGH WASTEWATER TREATMENT AND INJECTION

In spite of what has been learned about the system, implementation must be initiated in the face of considerable uncertainties that include technical and economic feasibility, long-term efficacy, and social acceptability. Adaptive management principles demand a disciplined, long-term, integrated systems approach that requires continuous refinement of the tools, models, and collective understanding of the problem, conducting carefully designed monitoring of the results of each management action, and adjusting the overall solution set in response to continuous learning. Adaptive management requires a blend of short-term and long-term actions and a set of performance measures that are monitored to provide feedback on performance to improve future decisions. This posture of adaptive management depends upon embracing authentic stakeholder involvement and education, applying conventional technologies in creative ways that fit the particular complexities of the Mexico City challenge, and willingly conducting "responsible experiments" to enlarge the set of future options available to achieve sustainability.

The centerpiece of the long-term plan developed for Mexico City is initiation of an aggressive demonstration program for aquifer injection of treated wastewater. The preliminary schedule for implementing the program is shown in Fig. 15.6 which stipulates initiation of injection by the end of program year 2 with a target of up to 10 m^3/s injection by the end of year 8.

The program begins with a two-stage demonstration injection program as the foundation. The goals of the demonstration program's first stage are (1) to begin recharge operations at the first site within 2 years of the program initiation and (2) bring a second recharge site on line by the end of year 3. The two sites will have an approximate combined recharge capacity of 750 L/s. Stage II of the demonstration project will consist of upscaling the stage I results to facilities having recharge capacities on the order of 1000 to 2000 L/s. Subsequent to stage II, the program will proceed to full implementation of the water reclamation injection system which could include recharge capacity targets of approximately 5 m^3/s by year 6 and 10 m^3/s by year 8.

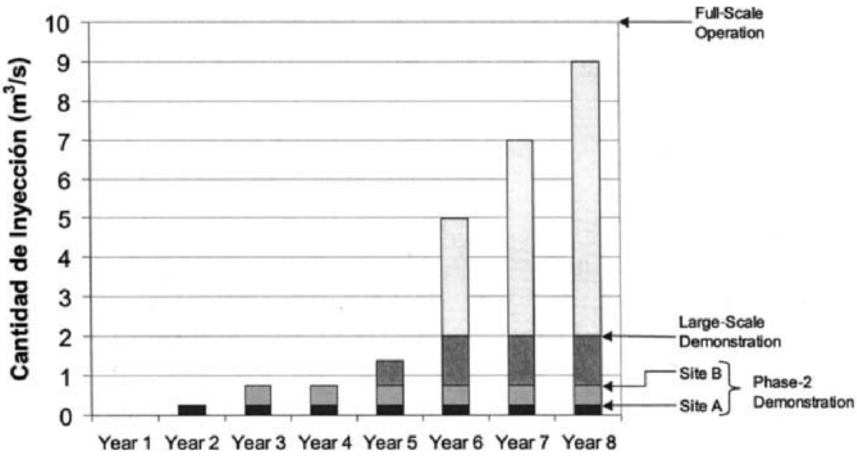


FIGURE 15.6 Phased demonstration program schedule.

The demonstration approach proposed here will take full advantage of existing facilities with suitable wastewater streams, and, as necessary, additional treatment chains will be implemented. The sites will be designed such that they will become an integral part of the operational recharge system once the demonstration phase has been completed.

The proposed demonstration phase has been designed and will be implemented according to our adaptive management approach. The initial demonstration sites will utilize conservatively designed conventional technologies for the treatment and injection systems. Throughout the implementation at each site (e.g., characterization, design, construction, and commissioning), methods will be severely scrutinized as to their cost, effectiveness, reliability, and consistency. In addition, whenever possible without unduly impacting the overall project schedule, new methods and techniques will be tested experimentally, striving to maximize learning from each demonstration site. To this end, the objectives of the demonstration phase are to

- Verify the design of injection wells and the ability to control local chemical and biological fouling at the point of injection.
- Design the program for extraction, including the process for empirically verifying aquifer behavior and the necessary residence time in the aquifer to ensure public safety, and of the siting and design of extraction wells.
- Establish required treatment levels and ensure that the necessary wastewater treatment systems are installed to achieve full compliance with all regulations.
- Implement appropriate monitoring systems to achieve full regulatory compliance and to optimize treatment system design and operations to maximize efficiency, reliability, and economy.

An important element of the demonstration program and successful implementation of the adaptive management approach is the development and deployment of monitoring systems for each of the recharge sites. Within the overall adaptive management process, monitoring is the basis for capturing the lessons learned from each of the steps in implementing the demonstration program and other initiatives, and successfully advancing to the subsequent stages of the Mexico City project.

With regard to long-term groundwater extraction and injection operations, the groundwater monitoring program provides immediate feedback on the quality of water used in the short term and provides data that can be used to predict the quality of groundwater extracted from the aquifer in the future. A comprehensive, consolidated groundwater monitoring program must consider all uses of the aquifer and must specify a number of specific parameters, such as horizontal and vertical monitoring locations as determined from a knowledge of groundwater flow patterns in the aquifer, both natural and as impacted by anthropogenic activities.

The monitoring systems will consist of three components:

1. *Operational monitoring.* Monitoring and control systems to maintain optimal and safe operation of aboveground (wastewater treatment and transportation) and below-ground (injection wells) systems. For example, particulates, organics, dissolved minerals, and physical parameters (e.g., temperature and pressure drops) must be closely monitored to ensure effective and safe wastewater treatment. Monitoring in the subsurface will provide immediate feedback on the quality of injected water as well as the change in groundwater quality over time as the injected water mixes with the ambient groundwater.
2. *Compliance monitoring.* Aboveground and below-ground monitoring focused on wastewater treatment and groundwater quality parameters as dictated by regulatory and operational criteria.
3. *Biomonitoring program.* This program will be developed to include appropriate bioassay techniques for detecting chemicals of toxicological concern that may not be detected using conventional monitoring and exhaustive chemical analyses.

15.6 KEY RESULTS

Aquifer injection is a direct approach to the problem of aquifer overexploitation. Under an aquifer injection program, water that would otherwise be lost to the system as wastewater is reclaimed and used to replenish the aquifer itself. This study evaluated the feasibility of injecting as much as 10 m^3/s into the aquifer. The initial analysis confirmed both that this level of injection is technically feasible and that this much water is potentially available for injection. In future years, as more of the demand reduction activities come into effect, it may be necessary to enhance the collection system in order to provide this level of water to the injection system. The following are a summary of other aspects of the results that explicitly reflect the adaptive management approach.

15.6.1 Economics Based

At the current stage of the analysis, a strict benefit-cost analysis was not conducted. Instead, an attempt was made, based on available data, to minimize the cost of achieving a given level of recharge. The analysis of alternative systems for recharging the Valley of Mexico Aquifer considered several interdependent factors:

- The desired recharge rate
- Source, quantity, and quality of wastewater available for recharge
- The performance of potential recharge sites including recharge capacity, compatibility of recharge water with ambient groundwater, proximity of pumping wells, and associated residence time of recharged water before it is extracted
- Capital and operating costs of water treatment, transport from treatment facilities to recharge sites, and injection systems

While artificial recharge has been found to be technically feasible in other cities, it is also potentially an expensive option. To provide a framework for evaluating each of these factors and their interdependence, one goal in selecting a recharge option is to optimize an objective function simply stated as "minimize C" where C, the total capital and operating costs, is expressed as

$$C = f\,(\text{treatment costs, transport costs, injection costs})$$

subject to the following overarching constraints:

- The total injection rate equals or exceeds the prescribed recharge rate.
- The residence time at each injection site equals or exceeds the minimum required residence time (most likely, 12 months).

Injected water at each injection site is treated to a sufficient level to meet all regulatory requirements, satisfy public and environmental health concerns, and minimize the risk of impacts from geochemical or microbiological processes (such as aquifer or well screen plugging).

The first constraint requires that, for a given development scenario for the Mexico City water system, the prescribed demand for recharging the aquifer must be satisfied. For example, if the development scenario indicates that to achieve the system objectives recharge to the aquifer must equal or exceed 5 m^3/s, then the recharge system would be conceived to inject, at one or more recharge sites, a cumulative total of 5 m^3/s. In turn, recharge at individual sites would be constrained by the site-specific injection capacities based on local hydrogeologic characteristics as well as constraints 2 and 3.

The second constraint relates to an important aspect of recharge system reliability (that is, maximizing treatment potential and mixing reclaimed water with natural groundwater within the aquifer). D'Angelo et al. (1998), in summarizing draft regulations proposed by the state of California, suggest a minimum residence time of 12 months and a minimum horizontal separation between injection and pumping wells of 600 m.

The final overarching constraint refers to the level of treatment for the reclaimed wastewater before it is injected. The minimum requirement is that reclaimed water be treated to comply with applicable regulatory standards (for example, drinking water standards). For this analysis, additional treatment requirements are determined based on geochemical and microbiological assessments of the compatibility between the injected water and the ambient groundwater at each injection site. Before the full-scale recharge system is implemented, reclaimed water treatment levels need to be thoroughly evaluated through demonstration studies to make certain that proposed criteria ensure a microbiologically and chemically safe groundwater recharge system.

15.6.2 Robustness

Given the enormous capital and operating costs associated with resolving Mexico City's problems, a premium was placed on preserving future flexibility and minimizing reversibility in selected alternatives. The demonstration program is designed to begin at a relatively small scale and "scale up" based on greater understanding of problems to be encountered and optimal solutions. It also enables the city to take full advantage of "technology creep" recognizing that considerable work is under way throughout the world in exploring treatment requirements for direct and indirect wastewater reuse. In addition, the plants associated with the demonstration program, while having a strong experimental aspect to their initial operation, will ultimately be incorporated directly into the city's wastewater treatment network and the aquifer injection system.

15.6.3 Experiment

The planned initial actions for the Mexico City water system are fundamentally experimental. A two-staged demonstration recharge program is recommended as an essential foundation for developing Mexico City's water reclamation injection system with a targeted capacity of approximately 10 m^3/s. The goal of the demonstration program's first stage is to have recharge systems operational at two sites with a combined recharge capacity of 750 L/s within 3 years of initiating the program. The demonstration plants will be constructed and operated in a phased approach to establish optimum treatment processes, injection water composition, well configurations, recharge rates, and monitoring and control system designs.

Because little is known about the biogeochemistry of the Valley of Mexico Aquifer, a carefully designed and executed program of characterization and experimentation is needed before the aquifer is artificially recharged on a significant scale. The following approach is designed to ensure injection well capacities can be sustained in the long term:

- Simultaneously characterize (horizontally and vertically) the sediment chemistry, bulk groundwater geochemistry, and microbiology in cored samples at candidate injection sites. This work should include basic sediment properties (for example, lithology, organic carbon concentration, and porosity), and an assessment of microbial community diversity.
- Relate the characterized vertical groundwater heterogeneity directly to the microbial community. This characterization may provide information on the microbial nutrient limitations by showing the correlation between microbial groups and chemical characteristics in the aquifer.
- Utilize the proposed demonstration site injection facilities to experimentally test hypotheses regarding the control of microbial communities by nutrient limitation. Candidate injectant chemistries can be tested against microbial community response by monitoring geochemical response to nutrient addition in nearby monitoring wells. Candidate elements to be analyzed are dissolved oxygen, nitrogen, and reduced species such as ferrous iron and sulfide.
- Recore the impacted sediment and characterize potential changes in the microbial community, organic carbon chemistry of the cored sediment, and porosity. This activity will provide engineering assurance that recharge activities and chemistries are compatible with the aquifer biogeochemistry.

Demonstration-scale operations of the membrane water treatment process are also necessary to gain enough knowledge to establish relationships between the feed water quality, product water quality and water quantity with membrane choices, process characteristics, pretreatment options, and operating conditions. Once the feed and product water streams are quantified with the demonstration scale membrane process, effective integration of the product stream with other water sources and existing geochemistry in the injection wells may be achieved.

15.6.4 Value of Information

An example of the concept of value of information in the implementation of adaptive management to the Mexico City study relates to the limited understanding of key processes important to the design and operation of an aquifer recharge program. As noted above, a carefully designed and executed program of characterization and experimentation is recommended before a full-scale aquifer recharge program is implemented. The program should include construction of demonstration-scale injection facilities to experimentally test hypotheses regarding the control of microbial communities by nutrient limitation. In the demonstration facilities, scientific characterization would set operating parameters for injections of treated recharge water. The plants would be operated (treated water would be injected into the aquifer), and the reactions to the injected water would be monitored to test the hypothetical response of the aquifer system. During an extended period of operations, the injection composition, recharge mechanism, well configuration, and recharge rate would be optimized through iterative scientific measurements and engineered adjustments. Such a program will provide engineering assurance that recharge activities and chemistries are compatible with the aquifer biogeochemistry. Further, the program is essential to confirming key performance criteria and optimization for the proposed design, including the following:

- Reclaimed water from the proposed treatment processes meets identified quality standards.
- Recovered water (that is, blended ambient groundwater and injected water) does not pose unacceptable risks.
- Unit capital and operating costs are optimized.
- Long-term operations of the treatment and injection systems are sustainable at design capacities.

15.7 SUMMARY AND CONCLUSIONS

Adaptive management was conceived as an approach to environmental impact assessment and management that encourages the integration of environmental with economic and social understanding, and the notion of experimentation and response. As a systematic and rigorous scientifically defensible program of learning from the outcomes of management actions, accommodating change, and improving management, it can also be effectively applied to water supply planning. Adaptive management was conceived as a systems approach building on ecology, physical sciences, and systems sciences. It is experimental in nature, utilizing trial and error and allowing for both known and unknown information. Its goal is to embrace uncertainty, accepting partial understanding of processes, and producing policies and designs that are less sensitive to the unexpected. One of the problems it was intended to address was that fundamental properties of a development or policy are firmly established early in the design process such that significant change can only be introduced in the process through conflict and extraordinary pressure from interest groups.

These issues are equally important to the long-term planning and development of water supply systems given the potential for extraordinary change associated with demographics, climate, regulation, technology, etc.

The application of adaptive management to water supply system planning was demonstrated in a case study focused on an assessment of alternatives for ensuring the long-term sustainability of Mexico City's water supply. Actions taken to address the long-term sustainability of Mexico City's supply system must necessarily be undertaken in the face of significant uncertainties. These uncertainties include technical and economic feasibility, long-term efficacy, and social acceptability. By taking an integrated systems approach, however, that combines the aggressive, continuous development of information, models, and decision tools, and includes a set of activities directed at artificial recharge, leak reduction, and demand reduction, the city can establish an adaptive, durable framework for achieving the goal of sustainability. This posture of adaptive management depends upon applying conventional technologies in innovative ways and willingly conducting "responsible experiments" to enlarge the set of options available to achieve sustainability. Adaptive management implies continuous refinement of the tools, models, and collective understanding of the problem, conducting carefully designed monitoring of the results of each management action, and adjusting the overall solution set in response to continuous learning.

15.8 REFERENCES

British Columbia Forest Service (BCFS). 1999. *An Introductory Guide to Adaptive Management.* Victoria, B.C., Canada. 24 pp.

Holling, C. S. 1984. *Adaptive Environmental Assessment and Management.* John Wiley & Sons, New York, 337 pp.

Katz, M., and P. Koss. Addressing Fish Uncertainty: The Quest for Rational Decisionmaking. In *Proceedings of the PSU Salmon Symposium,* Portland State University, July 2000.

Mays, L. W. (ed.). 1996. *Water Resources Handbook.* McGraw-Hill, New York.

National Research Council. 1999. *New Strategies for America's Watersheds.* National Academy Press, Washington, D.C., p. 301.

National Research Council. 1995. *Mexico City's Water Supply, Improving the Outlook for Sustainability.* National Academy Press, Washington, D.C.

Rauscher, H. M. 1999. Ecosystem management decision support for federal forests in the United States: A review. *Forest Ecology and Management* 114:173–197.

Shankle, S. A., R. Skaggs, J. L. Olson, and S. W. Gajewski, 2002, Mexico City Water Challenge: Blueprint for Sustainability, Report PNWD-3152, Battelle, Pacific Northwest Laboratory, Richland, WA.

Sit, V., and B. Taylor (eds.). 1998. Statistical Methods for Adaptive Management Studies, Res. Br., B.C. Min. For., Victoria, BC, Land Manage. Handb. No. 42.

Walters, C. J., and C. S. Holling, 1990. Large-scale management experiments and learning by doing. *Ecology* 71(6):2060–2068.

CHAPTER 16

WATER INFRASTRUCTURE MANAGEMENT: AN OVERVIEW OF EUROPEAN MODELS AND DATABASES

Patrick Eisenbeis
Cemagref Bordeaux
Cestas, France

Pascal Le Gauffre
INSA Lyon
URGC—Hydrologie Urbaine
Villeurbanne, France

Sveinung Saegrov
SINTEF Chemistry
Trondheim, Norway

In Europe, rehabilitation policies for drinking water pipes are based more on budgetary restrictions than on an authentic technical and economic consideration. In this reactive approach, only pipes that have failed or have been repaired at more than some subjective rate are rehabilitated. This approach leads to a very low rehabilitation rate and does not allow for the assessment and management of risks.

This chapter presents different preventive and proactive models and methods that should be developed for use by water utilities. These models were created in different European research centers and universities. They can be classified into two major types according to their objectives: (1) models assessing the structural state of the pipes based on statistical methods and (2) models assessing and comparing rehabilitation strategies, based on technical and economic assumptions.

The first model type uses short- or long-term data maintenance records. It is made up of two stages. First an analysis of the influencing factors is produced. These factors are either specific factors, such as material and diameter, or environmental factors, such as soil, traffic, or pipe location. Secondly, the state of the pipe is assessed or forecast using specific models, such as Poisson or Weibull distributions. The model concludes by classifying pipes either in groups or individually. The second model type requires a description of the network (age, length, material, and failure rate) and integrates economic data.

Previous European case studies are used to classify data requirements. For each model, we define *required data, highly significant data,* and *useful data.* Previous results are given as illustration.

At the end of this chapter, we present case studies about available data in European services. They show that available databases remain incomplete regarding failure analysis and modeling and rehabilitation forecasting.

16.1 INTRODUCTION: PRACTICES AND EXPECTATIONS IN EUROPEAN WATER UTILITIES

Theoretically, the rehabilitation and maintenance policies should combine two components: a *reactive* approach and a preventive or *proactive* approach (Zwingelstein, 1996; Smit, 1993; Saegrov et al., 1999). In European water utilities two types of maintenance management may be observed: policies based solely on a reactive approach and policies progressively integrating preventive and proactive components in order to reduce the current risks and the future problems that can be evaluated and forecast.

In the first approach, a financial envelope intended for renewal or renovation is fixed according to mainly financial criteria and is independent of the state of the asset stocks. The rehabilitated mains are selected according to emergency criteria, such as the number of breaks, but also, generally, according to the rehabilitation program of roadworks. In this case, the water service rarely takes into account the pipe or environmental characteristics. This reactive approach does not allow for risk management. In addition, the foreseeable acceleration of deterioration in average network conditions is likely to create financial problems for water services, involving large and increasing investment in the next few years.

The second approach is proactive. The water service determines where to invest money and labor after taking into account the state of the pipes and forecasting their degradation. This approach then dominates in the choice of the global quantity to be rehabilitated, as well as in the choice of the individual pipes. However, this policy requires a good knowledge of the network environment and its failures. This means the installation of a computerized database, preferably in the form of a geographic information system (GIS). Implementation of this approach may allow:

1. *Forecasting of investment planning.* Elaboration and simulation of a long-term strategy
2. *Integration of risk management and cost management in rehabilitation programs.* Assessment and reduction of risks such as severe damages and/or disruptions due to bursts, critical water interruptions to sensitive customers, hydraulic deficiencies due to repairs, etc.

However, in the framework of a proactive policy, water services will always have to implement a reactive component of the rehabilitation management: coordination with roadworks, reduction of water quality deficiencies, etc. The reactive approach cannot be avoided, but a proactive one has to complement it in order to create an efficient, long-term rehabilitation policy.

We present in Table 16.1 the relative importance of main rehabilitation objectives according to 12 European water utilities. Two main objectives are preeminent in most of these services: the improvement of water quality and the reduction of water mains failures and their consequences (damages on the urban environment, traffic disruptions, interruptions, lack of pressure, etc.). The first objective is achieved mostly by using a reactive approach. Monitoring is an indispensable process. The evaluation of performance indicators (Alegre et al., 2000) allows for the definition of zones where water quality deficiencies are unacceptable (noncompliance with European standards and/or numerous complaints) and hydraulic modeling is a complementary process to define pipes that have to be rehabilitated. The second preeminent objective is more suited to being managed with a preventive approach. Failure rates can be analyzed and forecast, and, consequently, risks of damages and other detrimental events may be assessed and managed.

Services need to preserve data in databases but also use technical and economic tools to aid decision making with regard to the rehabilitation of pipes. This chapter presents some of the results of the European Working Group A "Diagnosis of Water Infrastructure" of COST UCE Action C3 "Diagnosis of Urban Infrastructure" and additional results provided by the ongoing research project Computer Aided Rehabilitation of Water Networks (CARE-W).

In Sec. 16.2, we present models linked with rehabilitation of drinking water networks and developed in European research centers and universities. Then in Sec. 16.3, we present availability of data, useful for these models, in several large European water services. Finally, we define the research needs in this field.

16.2 MODELS FOR FAILURE ANALYSIS AND FAILURE AND/OR REHABILITATION FORECASTING

Several failure analysis and forecasting models have been developed in European universities and research centers. The main objective of the development of these

TABLE 16.1 Objectives of Rehabilitation Projects, According to 12 European Water Utilities*

Objectives	1: Overriding importance	2: Important	3: Minor importance	4: Not important	Data or tools missing (to take into account this objective)
Improve hydraulic performance	1	6	3	1	2
Improve water quality	5	5	2	0	3
Reduce operation and maintenance costs	3	7	2	0	6
Reduce water losses	2	7	3	0	3
Reduce the number of mains failures and their consequences	6	5	1	0	4
Reduce the age of water at the customer tap	0	3	5	4	5
Maintain or improve average condition of network	3	6	1	1	6

*Bristol, United Kingdom; Brno, Czech Republic; Codigoro, Italy; Dresden, Germany; Ferrara, Italy; Greater Lyon, France; Lausanne, Switzerland; Oslo, Norway; Reggio Emilia, Italy; Roubaix-Tourcoing, France; Stuttgart, Germany; Trondheim, Norway.

models is to provide water utilities with several tools that will allow for an assessment of the state of the pipes and the consequence for network function and an estimation of the required investment. Each model was applied in several large urban and rural water services, but their utilization is still experimental.

16.2.1 Principles

In Table 16.2, different European models are presented, classified according to their objectives and their principles. We discuss each model here.

AssetMaP: Asset Maintenance Procedure (INSA Lyon, France). This procedure had been initially proposed by Malandain (1999) and Malandain et al. (1999) when they were studying the water infrastructure of the urban community of Lyon. Two complementary components are briefly presented:

- *AssetMaP_1.* In the first step, homogeneous pipe categories are defined by a statistical analysis (Poisson regression, see AssetMap_2) or *a priori*; then fail-

TABLE 16.2 Some Approaches and Models Developed in European Research Centers

Models	Objectives and principles		
AssetMaP (INSA Lyon, France)	Statistical analysis of failure rates with Poisson regression	→ Categories & rate ratios →	Forecast of failure rates with Markov models applied to pipe categories ↓ Simulation of rehab. Strategies: Forecast of the overall failure rate according to rehab. hypotheses
Failnet_Stat (Cemagref, France)	Statistical analysis of failure occurrences with survival analysis and Weibull model	→ Influencing factors, failure risk function →	Forecast of failure number with Monte Carlo simulation applied to pipes, ranking pipes according to this number
Failnet_Reliab (Cemagref, France)	Hydraulic modeling considering pipe failures	→ Pressures, available consumptions →	Assessment of hydraulic reliability indices (global, per pipe or per node), according to available consumption modeling, failure probability, node importance
Clustering analysis (WRc and Imperial College, U.K.)	Spatial and clustering analysis of failures with GIS	→ Causes of failures →	Define failure risk zones (Waterfowl)
Winroc (NTNU, Norway)	Statistical analysis of failure occurrences with nonhomogeneous Poisson process	→ Influencing explanatory variables, failure intensity →	Forecast of number of failure, ranking pipes according to this number
Aquarel (SINTEF, Norway)	Hydraulic modeling considering pipe failures	→ Pressures, pressure indices →	Hydraulic reliability assessment according to failure probabilities, at pipe or node level
KANEW (Germany)	Analysis of existing stock infrastructure, cohort survival models	→ Evolution of pipe categories →	Comparison of rehabilitation strategies
UtilNets (North West Water, U.K.; Computer Technology Institute, Greece; SINTEF, Norway; Tecnic, Italy; UBIS, Germany	Physical assessment of failure rates considering corrosion rate and external load	→ Structural safety factors, loads and stresses on pipes →	Forecast failure rate by pipe

ure rates (without rehabilitation) are forecast with conceptual models (Markov processes) and calibrated with historical data (see Fig. 16.1). So several pipe categories can be studied, and introducing a hypothesis on rehabilitation rates, rehabilitation strategies, it is thus possible to forecast failure rates for future years. This procedure gives the opportunity to define models for pipe categories using incomplete failure records.

- *AssetMaP_2.* Statistical analysis of failure rates by Poisson regression is elaborated. The influence of failure factors can be characterized with rate ratios, and thus pipe categories can be defined. These rate ratios can be used in AssetMap_1 in order to calibrate simultaneously a set of aging functions associated with the categories.

Failnet: Analysis and Forecasting of Water Network Failures—Hydraulic Reliability Model (Cemagref, France). The method of analyzing and forecasting water pipe failures has been applied with data from several urban and rural water services, with short (10 years) or long (50 years) maintenance records (Bordeaux, suburb of Paris, Alsace, Charente-Maritime) (Eisenbeis, 1994; Eisenbeis et al., 1999; Le Gat and Eisenbeis, 2000).

The analysis and forecasting of water network failures is composed of three steps:

1. *Analysis of failures.* A proportional hazard model, by maximizing likelihood, allows identification of the significant factors and assessment of their influence.

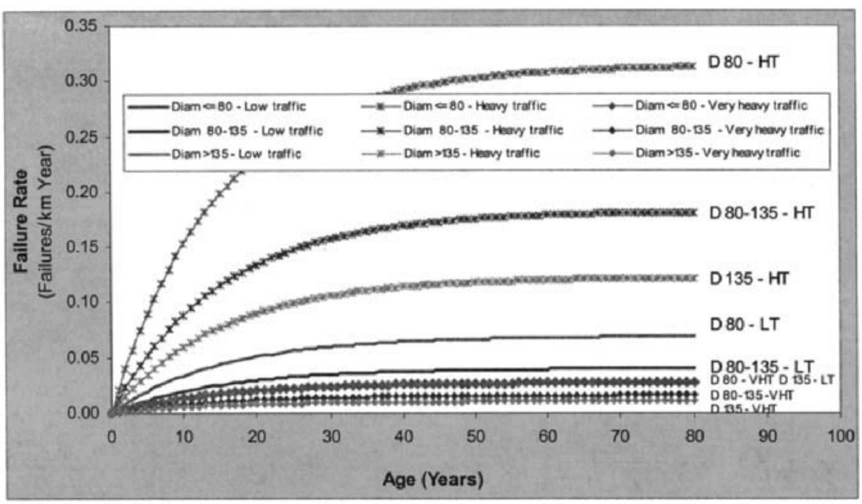


FIGURE 16.1 Failure rate functions for nine pipe categories (AssetMap).

2. *Definition of survival functions.* Based on a Weibull model, these functions are defined according to the material and the number of previous failures, taking into account the influencing factors selected in step 1.
3. *Forecasting the number of failures for a defined period.* Using a Monte-Carlo method, the number of failures is forecast from the survival functions. It is then possible to classify the pipes or a group of pipes. This forecast can be used in combination with a hydraulic reliability model, in an economic model, or alone as one of the rehabilitation criteria.

This hydraulic reliability model is based on a specific hydraulic model, calculating available consumption depending on pressure (Piller et al., 2001). One model is computed for each pipe failure. The impact of the failure is assessed regarding the quotient between available consumption and original demand. Indices then give pipe reliability, node vulnerability, and global network reliability (see Fig. 16.2) (Brémond and Berthin, 2001).

GIS and Cluster Analysis Techniques (WRc, United Kingdom). Cluster analysis techniques (Poulton, 2001) examine areas where bursts are concentrated in order to identify the underlying cause of failure for that location. Sophisticated

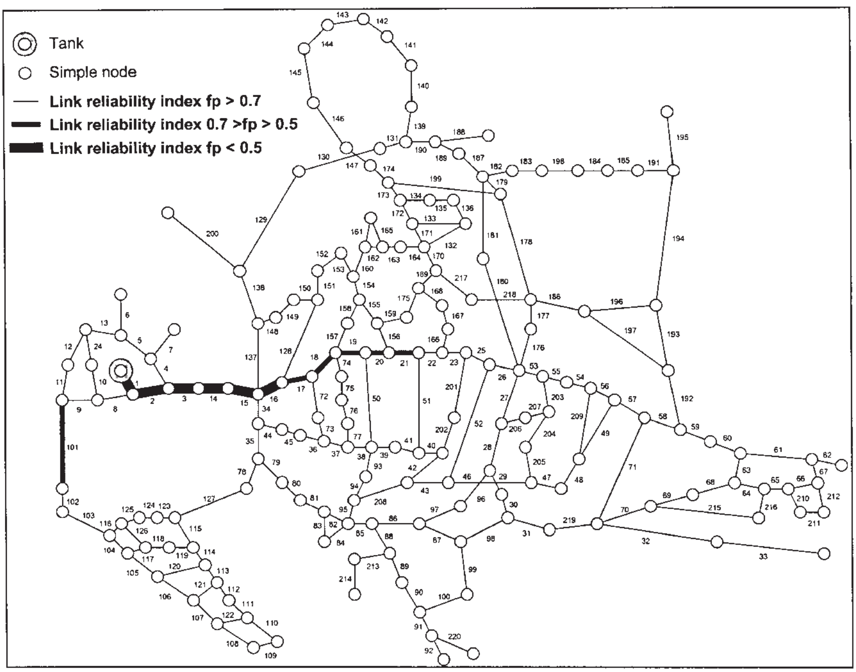


FIGURE 16.2 Link hydraulic reliability (fp) assessed with Failnet-Reliab.

GISs are used for the analysis. Local environmental factors such as soil type, cover type, and traffic loading may all affect the burst rate. Seasonal failure trends also provide valuable information about reasons for bursts. Once climate effects have been removed from annual burst numbers by statistical analysis, an underlying deterioration rate for the system may be determined. This can be compared with the deterioration rate for the burst clusters.

The technique highlights areas where rehabilitation should be targeted. A whole-life costing software package such as WRc's Waterfowl can then be used to determine the most economic rehabilitation technique in the long term.

Winroc/Aquarel: Vulnerability of Water Networks (NTNU/SINTEF, Norway). Statistical models based on modified nonhomogeneous Poisson processes (Winroc) and Weibull functions calculate the probability of failures of single pipes (see Fig. 16.3). The probability of failure is modeled for groups of pipes characterized by material, construction year, water quality, surrounding soil, and diameter. For those pipes with a record of very few failures (i.e., large-diameter pipes), the failure probability is estimated from professional judgment (Røstum and Schilling, 1999).

The vulnerability of water supply for any customer caused by any network failure is calculated with a combined hydraulic and reliability model. The hydraulic software EPANET is being extended by a routine (Aquarel) for estimating the probability of network failures and the consequence of this on water supply reliability (Saegrov, König et al., 2001).

These models are being tested with data from Norwegian municipalities.

KANEW: Exploring Rehabilitation Strategies (Dresden University of Technology, Germany). KANEW is a cohort survival model for infrastructure that has been developed at Karlsruhe University to determine future needs of infrastructure

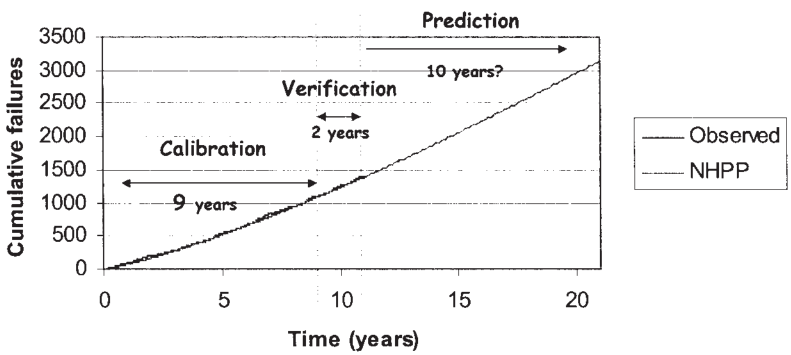


FIGURE 16.3 Construction and calibration of failure rate function for Winroc.

renewal. It has been produced as user-friendly software at Dresden University of Technology in a research project sponsored by the Research Foundation of the American Water Works Association (AWWA-RF, REP 265, released in 1999). It is now available in Access 2.0 and '97 versions to AWWA-RF subscribers. Basically, KANEW projects the number of pipes that reach the end of their service lives from the existing stock of infrastructure, differentiated by periods of installation and by types of elements with distinctive life spans. Service life is assumed to be a random variable, starting after some time of resistance and being characterized by a median age and a standard deviation or age that would be reached by a certain percentage of most durable elements. The user can choose these parameters of the Herz distribution. Projections are generated for pessimistic and optimistic assumptions of service lives that are derived from failure and rehabilitation statistics of typical elements of infrastructure in the past. For new materials, intelligent estimates of future behavior are made. The cohort survival model of KANEW is a tool for exploring network rehabilitation strategies (Herz, 1996, 1998).

UtilNets. UtilNets (König and Saegrov, 2001) calculates the remaining service life of single pipes and water networks from a deterministic model. The model considers internal and external corrosion of cast-iron and ductile iron pipes. The corrosion will vary to a large extent along the length of the pipeline. This is modeled in a probabilistic distribution function of corrosion.

UtilNets also includes the most important external load conditions. These are soil weight, traffic load, uneven foundation (pipe acting as a "beam"), pipe wall temperature, and frost action. As for corrosion, the load variation along the pipeline is modeled by a probabilistic function, and the extreme values are determined from worst-case estimates, for example, no foundation over one pipe length (beam length = 6 m). The system comprises a GIS-based user interface, and results are presented as thematic maps (and tables).

16.2.2 Requirements on Databases

Overview. Table 16.3 presents the data required by or useful to these models. These data were determined from studies applying the models to different services. For example, Failnet was tested on four services (two urban services, Bordeaux and a suburb of Paris, and two rural services situated in the Charente-Maritime and Alsace areas). For this model, the data depended on their availability and on their influence on failure rate.

For all these models, four variables must be known: (1) ages of the pipe segments, (2) lengths of the pipes, (3) pipe material, and (4) recorded pipe breaks except for the model KANEW. These are the basic variables necessary for establishing a database concerning pipes and their failures. The pipe diameters, although not completely necessary for KANEW, can be added to these data.

TABLE 16.3 Description of the Required Data for Each Model

Maintenance data and description of the asset stock	AssetMaP	Failnet	KANEW	Cluster analysis	Winroc/ Aquarel	UtilNets
Recorded breaks	*	*	‡	*	*	†
Ages of pipes	*	*	*	‡	*	*
Lengths of pipes	*	*	*		*	*
Material	*	*	*	*	*	*
Diameter	†	†	‡	*	*	†
Soil	†	†		*	*	*
Traffic	†	†		‡	‡	*
Location of pipes	†	†			‡	‡
Water pressure	†	‡			‡	†
Number of previous breaks		†			†	†
Type of joints		‡			‡	
Burst type				‡	‡	‡
Pipe condition				‡	‡	
Tree locations				‡		
Contours				‡		

*Required data.
†Highly significant data (according to previous studies).
‡Useful data, but not studied.

Among these models, only Kanew does not require additional data. The other variables used by the models are environmental data, concerning soil, traffic, or the location of the pipe in the street. The WRc model can be distinguished as it also needs water pressure and burst type. We can note that new data could be required if, with experience of new services, they are defined as highly significant for these sites.

More Details about the Variable "Material." Table 16.4 shows that the material highly influences the occurrence of failures. The values in this table are the failure rate or hazard compared to the rate or hazard of gray cast-iron (GCI) pipes. In Reggio Emilia (Italy), the failure rate is larger for GCI and asbestos cement (AC) pipes. In this case the average age of the pipe is not taken into account. In fact, the pipe age is 1 to 3 years for polyethylene (PE), 4 to 15 years for polyvinyl chloride (PVC), 16 to 35 years for AC, and 36 to 115 years for GCI (Di Federico et al., 1997). It can also be seen that, for the GCI pipes, the failure risk decreased from 1994 to 1996, thanks to a policy of decreasing water pressure in the service.

TABLE 16.4 Some Relative Risks RR Calculated from Previous Studies

	Reggio Emilia			Failnet	Winroc		
					Trondheim	Oslo	Bergen
	1994	1995	1996	Bordeaux	1988–1996	1976–1998	1978–1999
PE	0.01	0.11	0.25		0.06	0.22	0.06
PVC	0.21	0.25	0.30		0.01	0.33	0.12
Asbestos cement (AC)	0.34	0.64	0.68		1.92		1.44
Steel	0.08	0.11	0.15				
Gray cast iron (GCI)	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Ductile iron (no corr. prot.)				0.81	1.75		
Ductile Iron (corr. protect)				0.81	0.22		0.12

RR = (failure rate of the concerned material/failure rate of GCI) for Reggio-Emilia and Winroc and RR = *h*(ductile cast iron)/*h*(GCI), where *h,* the hazard function, is calculated with the Failnet model. If the value is more than 1, the material will break more than the GCI.

For Bordeaux, ductile cast-iron and GCI pipes are compared, even after eliminating the influence of age. It shows that GCI pipes break more often than ductile cast-iron pipes. We can note that, in Norway, asbestos cement and nonprotected ductile cast-iron pipes are more vulnerable than GCI pipes.

More Details about the Variable "Soil." Table 16.5 shows that as for diameter, the surrounding soil acts in the same way in the three services applying Failnet as for Assetmap and Winroc. The differences in value for this relative risk may arise from different definitions of soils.

In Trondheim, using the model Winroc, a rough classification has been applied to represent the soil:

Very aggressive. Tidal zone, groundwater level, natural soil with resistivity under 750 ohm · cm, pH less than 5, polluted by chemicals, stray current, alun slate

Moderate aggressive. Clay, wetland, inhomogeneous

Not aggressive. Natural soil resistivity over 2500 ohm · cm, dry conditions, sand, moraine

In Lyon, types of soil have been defined by scanning geologic maps (scale 1/50,000) with few significant results from statistical analysis (see example in Table 16.4). In addition to that, areas with geotechnical risks (soil movements)

TABLE 16.5 Some Relative Risk Calculated from Precedent Studies (Soil)

Failnet			AssetMap	Winroc
$RR = \dfrac{h(\text{corrosive soil})}{h(\text{noncorrosive soil})}$			$RR = \dfrac{BR(\text{other})}{BR(\text{alluvium})}$	$RR = \dfrac{h(\text{clay})}{h(\text{nonclay})}$
Bordeaux (GCI, 1st fail.)	Charente M. (GCI, 1st fail.)	Suburb of Paris (GCI, 1st fail.)	Lyon (GCI)	Trondheim
1.75	3.64	1.33	1.4	3.09

Note: BR is the average failure rate of each group of pipes.

have been defined by a previous study (Vinet, 1991). This variable appears highly significant in explaining problems with joints and leak frequency.

Estimation by Poisson regression gives a relative failure risk (RR) of 2.22 between a zone with geotechnical risks and another zone, with a 95 percent confidence interval of [1.21; 4.08].

In Bordeaux, the type of soil has been defined from a specific study, measuring the resistivity of the soil on half of the zone. In Paris and in Charente-Maritime, it was defined using staff knowledge of soil corrosivity, which may lead to a more subjective definition.

More Details about the Variables "Traffic" and "Location." For Failnet, traffic is considered to be a qualitative variable (0 or 1) according to the number of vehicles per hour or the type of road (local or national road). Some results are presented in Table 16.6.

For AssetMap and the case of Lyon, six situation types have been used (see Table 16.7). These results display low differences between light and heavy or very heavy traffic and can be considered an underestimation of relative risks. Actually it has been shown (Malandain et al., 1999), that the location of pipes must be considered as uncertain data. In studying this uncertainty on a sample (a 211-km network in Villeurbanne) with a Bayesian approach, the point estimate of the relative risk increased from 1.6 to 4 (see Table 16.8). These results demonstrate that improving accuracy and reliability of databases appears to be highly profitable for failure modeling.

16.3 AVAILABILITY AND MANAGEMENT OF DATA: SOME CASE STUDIES

16.3.1 Some Case Studies Concerning European Cities

In Sec. 16.3, the data used and available in previous or current studies are presented. It concerns 12 services in Europe (Lyon and Roubaix-Tourcoing in France;

TABLE 16.6 Some Relative Risks Calculated from Precedent Studies (Traffic Location) for Failnet

Failnet		
$RR = \dfrac{h(\text{high traffic})}{h(\text{low traffic})}$		
Bordeaux (GCI, 1st fail.)	Charente M. (GCI, 1st fail.)	Suburb of Paris (GCI, 1st fail.)
2.30	3.00	1.77

TABLE 16.7 Some Relative Risks Calculated from Precedent Studies (Traffic Location) for AssetMap

(Traffic) × (pipe location) Case of Lyon (Malandain, 1999)	Point estimate of RR = BR (Ei) / BR (E0) and 95% confidence interval
E0: pipe under pavement	1
E1: under roadway with light traffic (<25 trucks/day)	1.13; [0.98; 1.30]
E2: under roadway with heavy traffic (25 < < 300 trucks/day)	1.35; [1.18; 1.56]
E4: under secondary road	1.35; [1.18; 1.56]
E3: under roadway with very heavy traffic (>300 trucks/day)	1.80; [1.37; 2.35]
E5: under main (national) road	1.80; [1.37; 2.35]

TABLE 16.8 Some Relative Risks Calculated from Precedent Studies (Traffic Location) for AssetMap

Pipe location (Malandain et al., 1998)	Point estimate of RR and 95% confidence interval, *without* considering uncertainty of data	Point estimate of RR and 95% confidence interval, *considering uncertainty* of data (Bayesian approach)
L0: pipe under pavement	1	1
L1: pipe under roadway	1.64; [1.06; 2.54]	4; [2.68; 5.96]

Lausanne in Switzerland; Reggio Emilia, Ferrara, and Codigoro in Italy; Bristol in the United Kingdom; Oslo and Trondheim in Norway; Dresden and Stuttgart in Germany; and Brno in the Czech Republic). As they were chosen as experimental sites, they are not completely representative of an average service. But the objective is to present these data, the difficulty in collecting them, and the need to manage incomplete, imprecise, or uncertain data. Table 16.9 shows the different data available in these services.

16.3.2 Improving Databases: A Need to Improve Failure Analysis and Forecasting

Improving databases means dealing with incomplete, imprecise, and uncertain data. This should be taken into account when considering stakes and costs. In presenting previous results obtained with European models (Sec. 16.2), we have defined three types of data: required, highly significant, and useful. This can be considered as a first step of advice for municipalities concerned about implementing or improving a water infrastructure database for maintenance management.

As we can see in Fig. 16.4, available databases remain incomplete regarding failure analysis and modeling, and rehabilitation forecasting. The *ages* of the pipes and the type of *material* at a least level (both required data) remain partly unknown, and this should lead to investigation of archives and/or searches for efficient heuristics in order to estimate these two variables. It seems obvious that imprecise data concerning the ages of pipes should be greatly preferable to incomplete databases, regarding failure or rehabilitation forecasting.

Highly significant data, such as *soil, traffic,* and *pipe location,* are poorly represented in databases. The WRc and NTNU/SINTEF approaches have shown that soil characterization could even be considered as required data for cities confronted with aggressive and very aggressive soils. As we have seen in "More Details about the Variables 'Traffic' and 'Location'" in Sec. 16.2.2, *traffic* and *pipe location* should also be considered as efficient risk indicators. An evaluation or estimation of *traffic* intensity should be possible in most cases. One remaining problem is to deal with the uncertainty of the *location* variable (under pavement versus under roadway). Probabilistic tools should be tested and used in order to take into account this uncertainty in failure analysis and forecasting (Malandain et al., 1998).

Finally, data on *water pressure* are confirmed to be more and more available (due to increasing implementation of hydraulic modeling of water networks). In considering studies by NTNU/SINTEF (on Norwegian networks), INSA (on a 250-km network in Lyon) (Malandain, 1999) and Cemagref (on the Lausanne network) (Le Gat et al., 2000), and previous studies in the United States (Andreou et al., 1987), which have shown the significance of this variable, results from hydraulic modeling appear to be important data to include in pipe databases.

TABLE 16.9 Available Data in the Previously Studied Services

Description of the asset stock	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7	Case 8	Case 9	Case 10	Case 11	Case 12
Total length of mains, km	3,000	490	700	890 (su. + di.)[a]		400	1,100	7,694	1,600	750	1,800	1,326
No. of pipe segments in DB	50,000	~3,000	7,000	15,886				76,161	37,000	7,000		16,531
GIS (yes/no/in progress)	Yes	Yes	Yes	Yes	No	In progress	Yes	Yes	Yes	Yes	In progress	Yes
Reliably recorded failures[b]	1993 →	1985 →	1993 →	1994 →	1995 →	None	1973 →	1994 →	1981 →	1988 →	1997 →	1978 →
Age of pipes (in databases)[b]	S: 62%; C: 21%[c]	99%	95%	75%	0%	0%	100%	100%[d]	97%	98%	50%	99%
Length of pipe segments[b]	100%	100%	100%	100%	100%	100%	93%	100%	100%	100%	100%	100%
Material[b]	S: 98%; C: 50%[c]	97%	100%	Su: 85%; Dis: 95%[a]	100%	100%	92%	80%	97%	95%	90%	99%
Diameter[e]	100%	100%	100%	100%	100%	100%	100%	90%	100%	100%	90%	100%
Type of joint		0%	95%		0%	0%	0%	30%	0%	100%	60%	100%
Soil[e]	100%	0%	0%	In progress	0%	0%	0%	100%[f]	Thematic maps	Thematic maps	Soil 60%; bedding 20%	31%

TABLE 16.9 Available Data in the Previously Studied Services (*Continued*)

Description of the asset stock	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7	Case 8	Case 9	Case 10	Case 11	Case 12
Traffic[e]	100%	100%	0%	In progress	0%	0%	0%	50%	0%	0%	0%	0%
Location of pipes (road-pavement)[e]	98%	0%	0%	100%[g]	0%	0%	100%	100%	98%	100%	50%	100%
Water pressure	20% in progress	0%	100%	[h]	100%	0%	100%	70%	100%	100%	100%	100%

[a]Su: supply, di: distribution.
[b]:Required data.
[c]:C: city, S: suburbs.
[d]:50% known, others estimated from pipe material.
[e]:Highly significant data (according to previous studies).
[f]:Corrosivity and fracture potential class.
[g]:Average of total length, one orthogonal measurement from the properties each 120 m of length.
[h]:Pressure spot measurements, permanent district flow metering.

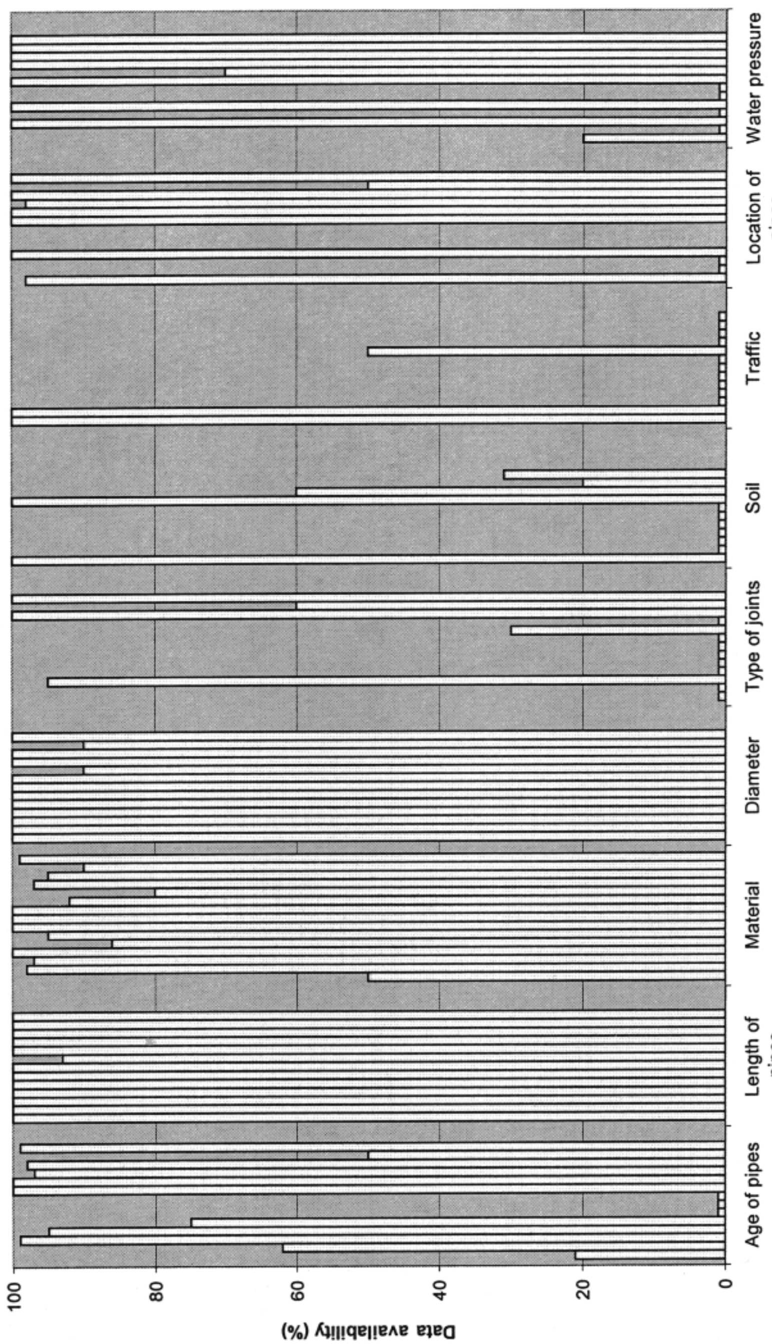


FIGURE 16.4 Availability of the data (*each bar represents percentage of available data in one service*).

16.4 CONCLUSION: ONGOING RESEARCH AND REQUIREMENTS ON URBAN DATABASES

This chapter presents some of the results of the Working Group A "Diagnosis of Water Infrastructure" of COST UCE Action C3 "Diagnosis of Urban Infrastructure." Previous studies, involving some large European cities and research centers and universities, made it possible to develop and test several models, dealing with failure modeling or rehabilitation forecasting. These studies provide or make a contribution toward efficient tools for assessing and predicting the condition of water asset stocks and rehabilitation needs, but it appears that further work is necessary in order to extend knowledge and practical use.

First, in order to be validated and adapted to each water service, these models need to be applied and compared on more services, from various countries, involving other pipes and environmental conditions. Second, testing and using these models implies, for water services, the set up and maintenance of accurate databases. This first level of a survey, presented here, provides information about ongoing projects and efforts to be made concerning the implementation of maintenance and asset databases. Developing methods to deal with incomplete databases, and imprecise or uncertain data, appears also to be a research need. Finally, we have to notice that the development of practical use of maintenance models and the setting up of efficient maintenance procedures in European water services implies the development of new tools, such as a guide for best management practices or software.

Further research, involving both research laboratories and water services, on a European scale, appears to be a necessary way to reach these different goals. This is also the background for establishing the project CARE-W (EU project, contract no. EVK1-CT-2000-00053). CARE-W (Saegrov, Røstum et al. 2001) deals with water supply networks and their problems of aging, such as structural failures, hydraulic insufficiencies, leakage, deteriorating water quality, and increasing maintenance costs that impact on the urban environment. CARE-W aims to establish a rational framework for water network rehabilitation decision making. In other words, rehabilitate the right pipelines at the right time by using the right rehabilitation technique at a minimum total cost, before failure occurs.

The project is managed in five main tasks:

1. *Construction of control panel for performance indicators for rehabilitation.* To define appropriate performance indicators for water network rehabilitation and determine realistic values for these indicators
2. *Description and validation of technical tools (failure forecasting models and hydraulic reliability models).* To test and improve them on several networks and to establish a procedure for the selection of appropriate models according to available data in the services

3. *Elaboration of a decision support for annual rehabilitation program.* To establish a procedure helping the definition of annual rehabilitation programs, with a multicriteria decision based on technical and economic points of view
4. *Elaboration of long-term strategic planning and investment.* To develop a long-term rehabilitation scenario manager and design a procedure for rehabilitation strategy evaluation
5. *Elaboration and validation of CARE-W prototype.* Specifying data input/output and storage requirements for the integrated package, producing a specification for software integrating optimally the use of the tools

The ultimate goal of the project is to develop a suite of tools, which provide the most cost-effective system of maintenance and repair of water distribution networks, with the aim of ensuring reliable water supply that meets social, health, economic, and environmental requirements. More information is given at *www.unife.it/care-w.*

16.5 ACKNOWLEDGMENTS

The authors acknowledge the participants of the European Working Group "Diagnosis of Water Infrastructure" of COST UCE Action C3 "Diagnosis of Urban Infrastructure," especially the persons who contribute to data collecting: Paul Conroy, WRc, United Kingdom; Raimund Herz, from University of Technology of Dresden, Germany; Marcello Schiatti from AGAC, Reggio Emilia, Italy; J. E. Oddevald from Oslo municipality, Norway. They also thank all water services, which provided some data.

16.6 REFERENCES

Alegre H., Hirner W., Baptista J. M., Parena R. Performance indicators for water supply services. *Manual of Best Practices Series,* IWA Publishing, London, ISBN 1 900222 18 3, 2000, 150 pp.

Andreou S. A., Marks D. H., Clark R. M. A new methodology for modeling break failure patterns in deteriorating water distribution systems: Applications. *J. Advances in Water Resources,* vol. 10, March 1987, pp. 11–20.

AWWARF, Quantifying Future Rehabilitation and Replacement Needs of Water Mains, Final Report of Research Project 265, Denver, 1998.

Brémond B., Berthin S. Reliability of a drinking water supply system. *IWA International Conference in Brno-Czech Republic,* May 16–18, 2001, pp. 167–174.

COST Action C3, Diagnosis of Urban Water Supply and Wastewater Infrastructure, Proceedings from COST C3 and of action workshop, Brussels, May 18–19, 2000, 195 pp.

Di Federico, V., Mazzacane S., Schiatti M. Active control of water distribution systems: The Emilia Romagna experience. *Fifth Conference on Pipeline Construction, Hamburg,* October 1997.

Eisenbeis P. Modélisation statistique de la prévision des défaillances sur les conduits d'eau potable. Doctoral thesis, University Louis Pasteur of Strasbourg, Collection Etudes Cemagref no. 17, 190 pp. + appendix, 1999.

Eisenbeis P., Røstum J., Le Gat Y. Statistical models for assessing the technical state of water networks-some European experiences. *Annual Conference of AWWA,* Chicago, June 20–24, 1999.

Herz R. Aging processes and rehabilitation needs of drinking water distribution networks. *J. Water SRT-Aqua,* vol. 45, 1996, pp. 221–231.

Herz R. Exploring rehabilitation needs and strategies for water distribution networks. *J. Water SRT-Aqua,* vol. 47, 1998, pp. 1–9.

König A., Saegrov S. Evaluation of UtilNets—a decision support system for water mains rehabilitation. *IWA International Conference in Brno, Czech Republic,* May 16–18, 2001, pp. 182–188.

Le Gat Y., Eisenbeis P. Using maintenance records to forecast failures in water networks. *Urban Water,* vol. 2, no. 3, 2000, pp. 173–181.

Malandain J. Modélisation de l'état de santé des réseaux de distribution d'eau pour l'organisation de la maintenance. Doctoral thesis, INSA Lyon, May 1999, 206 pp.

Malandain J., Le Gauffre P., Miramond M. Organizing a decision support system for infrastructure maintenance: Application to water supply systems. *Journal of Decision,* vol. 8, no. 2, 1999, pp. 203–222.

Piller O., Brémond B., Eisenbeis P. Water supply network modeling including head-dependent demand and elevated nodes. *IWA International Conference in Brno, Czech Republic,* May 16–18, 2001, pp. 72–80.

Poulton M. Elements of an integrated approach for studying leakage and bursts in water distribution of pipelines. Ph.D. thesis, Imperial College of Science, Technology and Medicine, London, September 2001.

Røstum J., Schilling W. Predictive service-life models used for water network management. *14th EJSW,* Dresden, September 8–12, 1999.

Saegrov S., König A., Milina J., Rohrer P., Røstum J., Schilling W., Selseth I. Forecasting water network rehabilitation needs—experiences Experiences from Norway. *IWA International conference in Brno, Czech Republic,* May 16–18, 2001, pp. 160–166.

Saegrov S., Melo Baptista J. F., Conroy P., Herz R., Le Gauffre P., Moss G., Oddevald J. E., Rajani B., Schiatti M. Rehabilitation of water networks: survey of research needs and on-going efforts. *Urban Water,* vol. 1, no. 1, 1999, pp. 15–22.

Saegrov S., Røstum J., Tuhovcak L., Eisenbeis P., Herz R., Le Gauffre P., Baptista J., Conroy P., Di Federico V., Mazzacane S., Schiatti M. Computer aided rehabilitation of water networks. *ISTT International Conference of No-dig Technology, Prague,* September 2001.

Smit A. M. *Reliability Centered Maintenance,* McGraw-Hill, New York, 1993.

Vinet, L., Typologie et Prévention des risques et accidents geotechniques en site urbain : les collines de la ville de Lyon entre 1977 et 1990, Doctoral Thesis INSA Lyon, 1991. 335 pp.

Zwingelstein G. La maintenance basée sur la fiabilité, Hermes, Paris, 666 pp.

CHAPTER 17
ISRAEL: URBAN WATER INFRASTRUCTURE IN THE DESERT

Hendrik J. Bruins
Department of Man in the Desert
Jacob Blaustein Institute for Desert Research
Ben-Gurion University of the Negev
Sede Boker Campus, Israel

Now the rest of the acts of Hezekiah and all his might, and how he made the pool and the conduit and brought water into the city, are they not written in the Book of the Chronicles of the Kings of Judah?
2 Kings 20:20

17.1 INTRODUCTION

A remarkable engineering project for urban water supply was carried out some 2700 years ago in Jerusalem, as mentioned in the above Hebrew text. Jerusalem has a semiarid climate and is situated near the edge of the Judean desert. Rainfall occurs mainly in winter, while the long hot summer is completely dry for some five months from May to October.

Deserts and drylands have been a challenge for humanity since time immemorial. Settlement and development of arid regions necessitate first and foremost a secure water supply. Potable water is fundamental to human life. A person needs a few liters (L) of fluids per day as a physiological requirement. In hot desert climates the required daily amount may reach more than 5 to 10 L of water, because much moisture is lost through sweating.

Imagine a scenario where urban water supply in a dry region would come to a halt, for whatever reason, and there would be no alternative water supply in the short run. How many days can a healthy person survive without moisture intake? The answer varies with the temperature of the ambient environment (Hayward and Oguntoyinbo, 1987). The survival period without drinking water is about 18 days in a moderate temperature range of 16 to 21°C (61 to 70°F) for a person in rest and in the shade (Table 17.1). However, in a very hot environment of 39 to 49°C (102 to 120°F) the survival period without water intake is only 2 days (Table 17.1). These figures are not intended as small print in a water supply contract with a utility company. Yet, such information is very relevant in crisis and disaster situations (Bruins, 2000a).

The impending danger of an urban water crisis suddenly became a cause of great concern in the time of King Hezekiah during the period of the First Temple, as Jerusalem was under military threat from Assyria (2 Kings 20:20; Isaiah 22:11; 2 Chronicles 32:2–4, 30). A siege could cut the water supply to the city and force the people to surrender or die of dehydration.

The Gihon spring was the main water source of ancient Jerusalem. The spring is situated inside a cave in the Kidron Valley (Fig. 17.1). It does not maintain a constant flow, but is a siphon-type karst spring, fed by groundwater through cracks in the cave floor (Cahill and Tarler, 1994). The water flows intermittently for a duration of about 40 min every 6 to 8 h. Its flow characteristics may in fact be the origin of its name, as Gihon is derived from a Hebrew verb meaning "to erupt" or "burst forth" (Shiloh, 1994). The outflow frequency of the intermittent spring varies with the season of the year and the annual amount of precipitation.

As the Gihon spring was located just outside the city walls, urgent strategic planning was required to safeguard the water supply to Jerusalem, in view of the military threat from Assyria in the late eighth century B.C. Thus King Hezekiah and his water engineers developed an ingenious water supply system. A tunnel was dug in the limestone rock to channel the water underground into the city proper. The length of the tunnel is 533 m, and it descends gradually by only 2 m (Issar, 1990). The water flows through the outlet into a reservoir known as the Pool of Siloam (Fig. 17.1).

TABLE 17.1 Days of Expected Survival without Water

Daily minimum temperature		Daily maximum temperature		Days of survival without drinking water
°F	°C	°F	°C	
61	16	70	21	18
70	21	81	27	14
81	27	90	32	7
84	29	100	38	4
95	35	109	43	2½
102	39	120	49	2

Note: The days of survival, at different ambient temperatures, are average figures for a healthy person, who is at rest and in the shade.

Source: Hayward and Oguntoyinbo (1987), adapted after Bruins (2000a).

Two crews of miners dug through the solid limestone rock from both ends of the tunnel. They gradually worked their way onward, and the two crews could hear the sound of their digging tools as they approached each other. Finally they met exactly at the same spot and the subterranean water diversion channel was completed. How do we know? This was recorded in an ancient Hebrew inscription (Fig. 17.2) engraved in the rock wall near the outlet of the aqueduct tunnel. The inscription was accidentally discovered in 1880 when the country was part of the Ottoman Empire. It was removed by the Turkish authorities and is now exhibited in the Museum of the Ancient Orient at Istanbul (Vilnay, 1971).

The latter part of the text is translated as follows (Vilnay, 1971), the length of the *cubit* being about 45 cm (a foot and a half):

> …and while there were still three cubits to be cut through, (there was heard) the voice of a man calling to his fellow, for there was an overlap in the rock on the right (and on the left). And when the tunnel was driven through, the quarrymen hewed (the rock), each man toward his fellow, axe against axe; and the water flowed from the spring toward the reservoir for 1,200 cubits, and the height of the rock above the head(s) of the quarrymen was 100 cubits.

The distance between the Gihon spring and the outlet of the underground tunnel, as the crow flies, is 320 m, much shorter than the total length of the tunnel, which is 533 m (Fig. 17.1). The reasons for the large detour were an enigma for the archaeologists. Issar (1990) describes how hydrogeologic investigations solved this problem. The tunnel seems to follow a natural karstic channel in the Turonian limestone. This explains, according to Issar, how the two teams of quarrymen, digging from both ends of the tunnel, could meet at the same spot, while they did not dig in a straight line. It also explains the varying and often unneces-

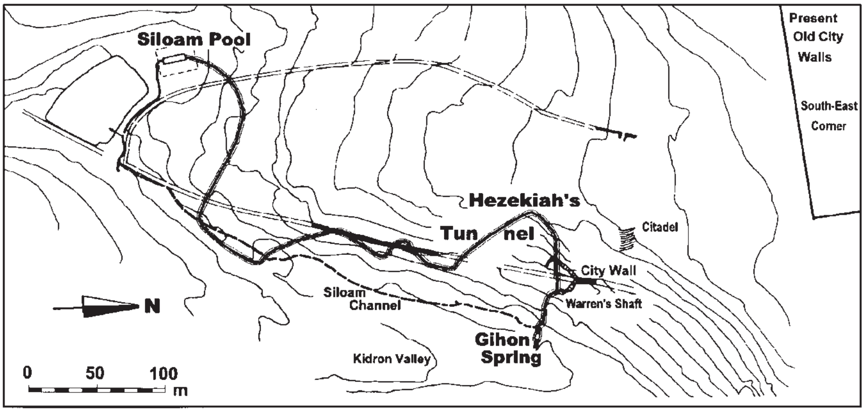


FIGURE 17.1 Hezekiah's tunnel and other iron-age water systems in Jerusalem in the City of David. (Cahill and Tarler, 1994, p. 43)

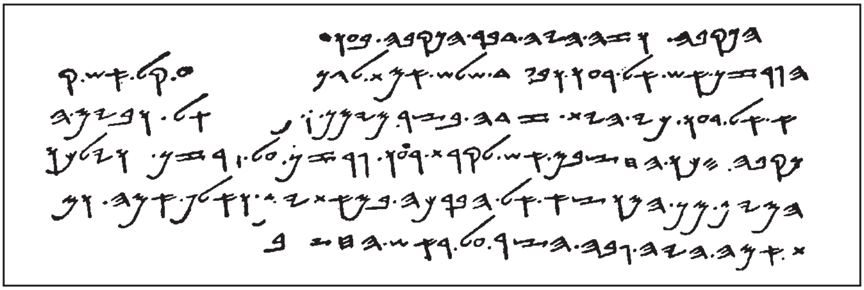


FIGURE 17.2 Inscription in ancient Hebrew inside Hezekiah's tunnel.

sary height of the tunnel, ranging from 1.1 to 3.4 m. Finally it helps us understand how the two crews of quarrymen could hear each other (Issar, 1990).

After completion of the tunnel, water from the Gihon spring flowed intermittently into the Pool of Siloam. The latter reservoir, situated at a safe place within the city, functioned as a buffer storage. This ingenious system secured the urban water supply to the ancient inhabitants of Jerusalem.

17.2 GEOGRAPHY AND CLIMATE OF ISRAEL

The state of Israel is quite small, being about the same size as New Jersey. Its length is about 425 km (264 mi), and its maximum width, in the northern Negev, is just 115 km (71 mi). The climate of Israel is comparable to some extent with that of California, ranging from more humid areas in the north to deserts in the south and east. Four different seas form an important part of the geography of Israel. The Mediterranean Sea in the west is of great significance for the climate, as it supplies moisture to the air. The Sea of Galilee, also known as Lake Kinneret, is situated in the northern part of the Rift Valley. It constitutes the principal resource of potable surface water in the country. Further south lies the Dead Sea, a terminal lake that is unique in two respects: (1) it occupies the lowest spot on the terrestrial surface of the Earth, and (2) it is the most saline sea in the world, containing 322 g of salts per liter, being nine times more saline than ocean water. The Red Sea in the southernmost part of Israel, also situated in the Rift Valley, forms part of the Indian Ocean. The coast near Eilat exhibits attractive coral reefs in a spectacular landscape setting.

The rapid landscape and climate changes within a small area give the country of Israel a unique flavor and beauty. Deserts make up more than half of the country. The Judean desert is situated east of the central mountain ridge (Fig. 17.8), as the rain-shadow effect aggravates the aridity of the area. The average annual rainfall decreases in the Jerusalem area rapidly from 600 mm (25 in) in the western part of the city to only 100 mm (4 in) near the Dead Sea over a distance of just 25 km (16 mi) eastward. The Negev Desert makes up the entire southern half of Israel. It forms part of the largest planetary desert belt on Earth, which stretches over 8000 km (5000 mi) from the Atlantic coast of Morocco eastward to the Thar Desert in India.

Israel lies near the southern limit of cyclonic rains, which bring precipitation to the country during late autumn, winter, and early spring. The number and trajectories of these low-pressure systems often determine the amount of rainfall. The country is usually devoid of any rainfall in the 6-month period from May to October. Large interannual variations occur in the amount of precipitation, which result in the regular occurrence of meteorological drought. A study about rainfall variability in Israel by Amiran (1994) showed that precipitation in drought years is about 30 to 40 percent less than the long-term annual average.

17.3 ARIDITY, WATER RESOURCES, AND THE RAINWATER-HARVESTING CIVILIZATIONS IN THE NEGEV DESERT IN ANTIQUITY

How dry is dry? *Bioclimatic aridity* is the principal constraint for habitation and development of drylands. Climatic variability adds a factor of uncertainty on top of the above constraint, because rainfall variations in drylands are large and unpredictable from year to year. Drought is a natural hazard, related to climatic variability, which can develop in a disaster (Wilhite, 2000; Bruins, 2000b). Drylands can be classified on the basis of bioclimatic aridity. UNESCO (1979) produced a map of the world distribution of arid regions, in the framework of the *Program on Man and the Biosphere* (MAB). This map is based on the *P/ETP moisture index*, where P is the average annual precipitation and ETP the average annual potential evapotranspiration. ETP was calculated according to the Penman (1948) method, which requires data on solar radiation, atmospheric humidity, and wind.

Four *aridity classes* are distinguished (Table 17.2): *hyperarid*, *arid*, *semiarid*, and *subhumid zones* (UNESCO, 1979; Bruins and Lithwick, 1998; Bruins and Berliner, 1998). *Rainfed agriculture*, based on direct precipitation, is generally possible in the subhumid and semiarid zones, albeit with risks due to drought. The arid and hyperarid zones are too dry for rainfed agriculture, but may be suitable for raising livestock. Cattle require pasture in the wetter part of the arid or semiarid zone, while goats and camels are also suited for the hyperarid zone.

The *World Atlas for Desertification* (Middleton and Thomas, 1997) presents a map of the arid zones, also based on the P/ETP moisture index, but potential evapotranspiration (ETP) is calculated by the Thornthwaite (1948) method. The UNESCO (1979) map was not established on a systematic global set of time series of meteorological data. A map for the period 1951–1980 (Middleton and Thomas, 1997) had to be based on Thornthwaite ETP calculations, because not enough meteorological data are available worldwide to use the Penman formula. The

TABLE 17.2 The Arid Zones and Their Characteristics

Climatic zone	P/ETP ratio (Penman method)	P/ETP ratio (Thornthwaite method)	Interannual rainfall variability, %	Possible rainfed land-use
Hyperarid	<0.03	<0.05	~100	Limited livestock grazing, mainly by camels and goats
Arid	0.03–0.20	0.03–0.20	50–100	Classical livestock grazing zone (less suited for cattle); runoff agriculture
Semiarid	0.20–0.50	0.20–0.50	25–50	Rainfed agriculture (more viable with added runoff or irrigation); livestock grazing (all species)
Subhumid	0.50–0.75	0.50–0.65 Dry subhumid	<25	Rainfed agriculture (more viable with added runoff or irrigation); livestock grazing (all species)

Source: UNESCO (1979); Middleton and Thomas (1997); Bruins and Berliner (1998); Bruins and Lithwick (1998).

Thornthwaite method underestimates ETP values in very dry environments, while overestimating ETP values in subhumid regions. Hence, the lower boundary for the hyperarid zone and the upper boundary for the dry subhumid zone were adjusted. The two classifications are presented in Table 17.2.

Having a P/ETP index of about 0.05, the central Negev is situated close to the hyperarid zone. The area is much too dry for normal rainfed agriculture. Natural springs in the region are scarce, while groundwater is usually very deep. Urban water supply, therefore, could not be based on wells or springs. However, the landscape geomorphology and the presence of silt-loam loessial soils (Yaalon and Dan, 1974; Bruins, 1986) enabled the development of a special niche in terms of human ecology, based on rainwater harvesting. Cisterns were built to catch drinking water (Fig. 17.3) and rows of check-dams were built in suitable streambeds to make terraced fields for food production (Kedar, 1957; Evenari et al., 1982; Hillel, 1982; Bruins, 1986).

Thus a rather unique development of human-environmental interactions took place in the central Negev Desert. The desert landscape was changed by extensive human intervention (Bruins, 1986, 1990). Cisterns in the region have been described by Evenari et al. (1971, 1982).

Six urban centers gradually developed in the central Negev hills during Nabatean–Byzantime times (Fig. 17.4): Avdat, Mamshit, Nizzana, Shivta, Rehovot, and Haluza. The Nabateans lived as pastoralists in southern Jordan and penetrated the Negev, as they developed the spice trade from the Arabian peninsula via Petra, their capital (Fig. 17.4), to Gaza at the Mediterranean coast (Negev, 1986).

FIGURE 17.3 An oval-shaped cistern (Hemed), probably dating to the iron age, still collects runoff water today in the central Negev Desert. (Photograph H. J. Bruins)

They established the above places during the third to first centuries B.C., which developed into towns during the Byzantine period. Population estimates range from 71,000 to 25,000 people (Broshi, 1980).

Urban water supply was largely based on *rainwater harvesting*. Numerous cisterns were dug into the rock, while conduits were designed to collect the runoff water from roofs, pavements, or natural catchments (Figs. 17.5 to 17.7). A sophisticated rainwater-harvesting civilization developed in the heart of the Negev Desert, based on an average annual rainfall amount of about 100 mm (4 in).

17.4 WATER RESOURCES DEVELOPMENTS IN ISRAEL IN THE TWENTIETH CENTURY

Britain conquered the land of Israel from the Ottoman Turkish empire in 1917 during World War I. The country remained occupied by the British until the end

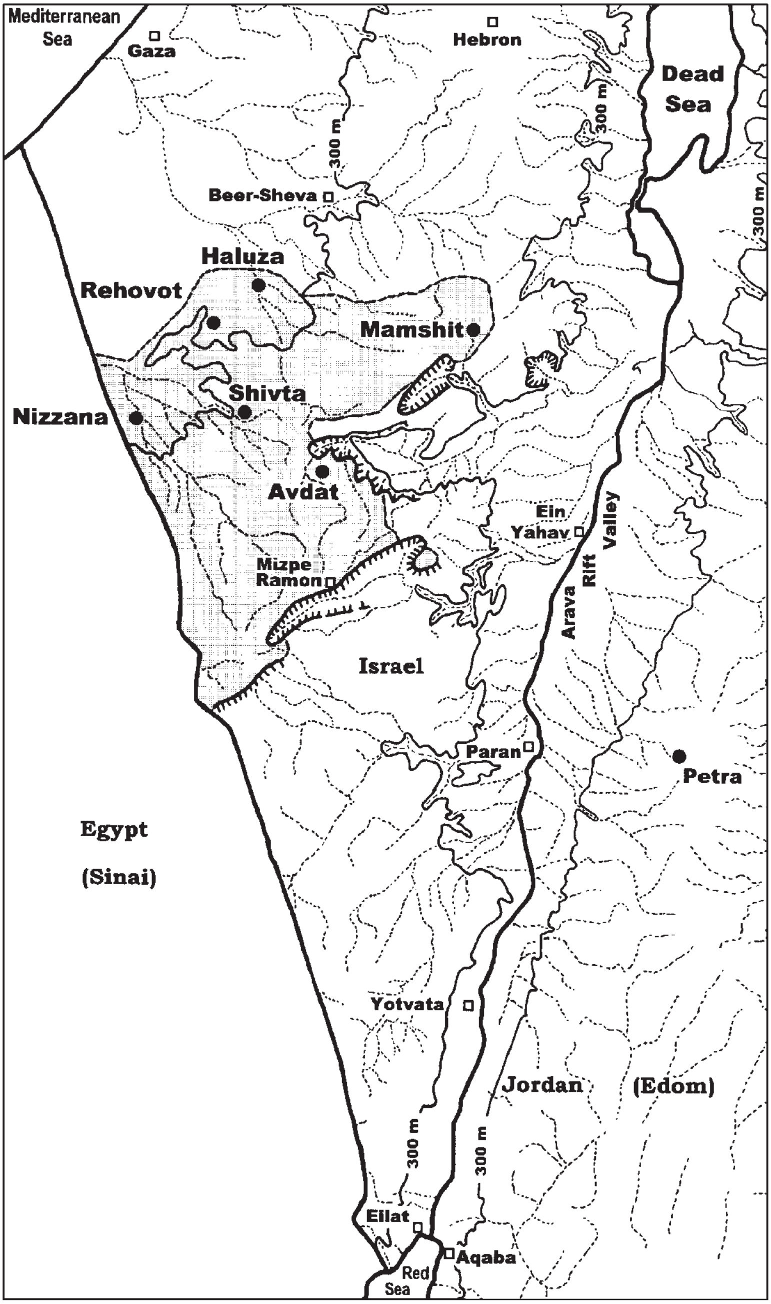


FIGURE 17.4 The location of six ancient towns in the Negev Desert, within an area characterized by rainwater harvesting for urban water supply and agriculture in antiquity. (After Bruins, 1986, p. 8)

FIGURE 17.5 The acropolis of the ancient Nabatean–Byzantine city of Avdat is situated on a flat hilltop in the central Negev Desert. A conduit enters the acropolis from outside the city walls. It used to channel runoff water from a natural catchment into the main cistern on the acropolis. (Photograph H. J. Bruins)

of the mandatory period in 1948, when Israel proclaimed independence. The British authorities did not make any drastic changes in the Ottoman code of law, which regulated water use in the country (Tamir, 1973). This legal heritage was found to be unsuited for modern development. It was, for example, forbidden to carry water from one region to another. Hence, a new and entirely different set of laws had to be created to regulate water management in the country (Arlosoroff, 1974). The water resources in Israel are centrally governed by the state, based on the Water Law enacted in August 1959. The Water Metering Law of 1955 ensures rigorous information about all water production, water supply, and water consumption in the country (Arlosoroff, 1974; Bruins 1999).

American scientists and politicians played an important role in the proposition of regional water projects. The idea of comprehensive water development was already suggested and described in the 1930s by Lowdermilk (1946). The American engineer James B. Hays worked out the idea, which he modeled on the Tennessee Valley Authority. The main components of his visionary project were summarized by Orni and Efrat (1971, pp. 447–449):

> It included points such as maximum use of groundwater sources and interception of ephemeral storm water, drainage of the Hula Lake and swamps, and irrigation of the

FIGURE 17.6 Opening of an underground cistern outside the city walls of Avdat. (Photograph H. J. Bruins)

FIGURE 17.7 In the foreground a cistern, situated inside an ancient Byzantine church. (Photograph H. J. Bruins).

> Lower Jordan Valley. All headstreams of the Jordan were to be collected by a countrywide carrier and diverted close to their sources so that the water could flow down freely without need of much pumping; water from the Litani River was to be led down from Lebanese territory into the Hula Valley. The cheap electricity thereby produced was to be at Lebanon's disposal, and the water at Palestine's. The Yarmuk River was to be diverted to the Sea of Galilee to replace the Jordan waters. Most important, a seawater canal was to be dug from the Haifa Bay to the Dead Sea to replace the Jordan's flow of fresh water and to use the level difference of 395 m (over 1300 ft) between the Mediterranean and the Dead Sea for electricity production on a vast scale.

Political developments after 1948 caused a gradual reduction of this visionary regional scheme. In 1953, President Eisenhower appointed Ambassador Eric Johnston to work out a water infrastructure plan acceptable to all the parties involved: the Jordan Valley Development Plan (Wolf, 1996). No mutual Arab–Israeli agreement was reached. However, the East Ghor Canal, largely built with American aid, was a positive spin-off for Jordan of the Johnston proposals. Water from the Yarmuk River was diverted and piped to the eastern part of the Jordan Valley. At about the same time in the 1950s Israel began to build the first phases of a large canal to divert water from Lake Kinneret.

The National Water Carrier System in Israel was designed by TAHAL, a public corporate body established in 1952 to plan water development. The Mekorot Water Company, founded in 1937, has built the water infrastructure in the country and handles its operations. The main part of the National Water Carrier System was completed in 1964. Lake Kinneret forms the pivotal part of the system, which also incorporates the Coastal Aquifer and Western Mountain Aquifer. These are the three most important water resources in the country.

The National Water Carrier System begins in the northwestern part of Lake Kinneret (Fig. 17.8). Here water is pump-lifted through pressure pipes from the lake surface at about −210 m to an altitude of 152 m: a rise of 362 m! Then the water runs in an open canal for 16 km (10 mi) to a reservoir with an area of 20 hectare and a volume of 0.8 million cubic meters (mcm). The water passes deep gorges by siphons, is pump-lifted again to an altitude of 147 m, passes through underground tunnels, enters another reservoir of 1.5 mcm, passes again through a tunnel, and then flows through prestressed concrete pressure pipes with a diameter of 270 cm (108 in). The main conduit reaches the northwestern Negev Desert and has a total length of 142.5 km (88.5 mi) (Orni and Efrat, 1971).

About 80 percent of all the water used in Israel is derived from three principal sources: Lake Kinneret, the Coastal Aquifer, and the Western Mountain Aquifer (Grinwald and Bibas, 1989). The annual contribution of groundwater is more important than surface water (Table 17.3).

The National Water Carrier System, designed as a pressure pipe system, can supply more than 450 mcm annually. Its peak delivery capacity is 20 m^3/s. The secondary piping network to urban communities and villages are placed about 1 m underground, designed to withstand at least 6 atmospheres (atm) of working pressure. Gate valves and air-release valves are installed above the ground to simplify maintenance and operation.

Lake Kinneret is the main water source of the National Water Carrier System. An average amount of 450 mcm was pumped from the lake annually, which was considered a safe yield. However, in recent years the lake has reached record low levels, due to drought and excessive pumping (Bruins, 1999, 2000a). Each meter of water in the lake represents about 170 mcm. The maximum water level is −208.9 m, above which flooding would occur in the town of Tiberias and other villages around the lake. A sluice gate at the southern end of the lake regulates the outflow of water. The minimum acceptable lake level, the so-called red line, used to be −213 m (Grinwald and Bibas, 1989; Schwarz, 1990). However, the red line was artificially lowered by another meter, due to the drought. There is a risk that water quality in the lake, particularly salinity, will deteriorate to unacceptable levels if pumping continues below the red line.

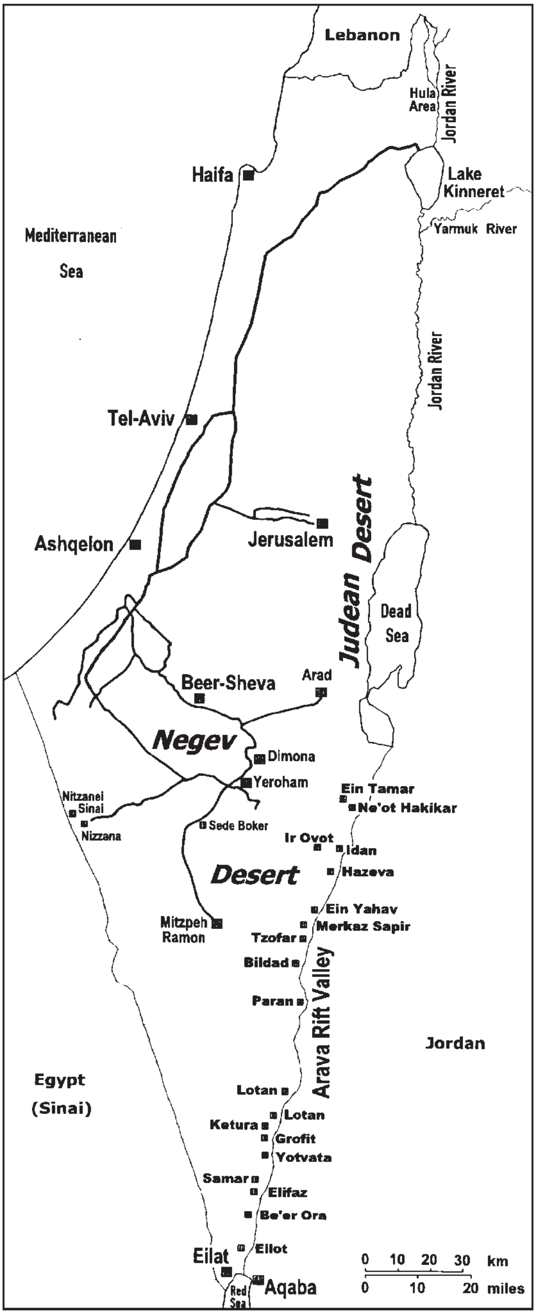


FIGURE 17.8 Urban water supply in the deserts of Israel is based on three approaches: (1) The National Water Carrier System brings water from Lake Kinneret south until the town of Mitzpeh Ramon. (2) Local brackish groundwater in the Arava Rift Valley is desalinated in each village for drinking water. (3) Eilat receives desalinated seawater.

TABLE 17.3 Short-term Water Potential in Israel in Million Cubic Meters (mcm) per Year, as Compared to Actual Supply in 1984/85 and Planned Supply for the Year 2000.

Water source	Fresh water	Saline water	Total	1984/85	2000
Groundwater: boreholes	768	132	900		
Groundwater: springs	82	100	182		
Total groundwater	850	232	1,082	1,340	1,115
Surface water: Hula Valley	122		122		
Surface water: Lake Kinneret	490		490		
Surface water: saline water		20	20		
Surface water: outflow Kinneret	−20		20		
Total Kinneret Basin	592	20	612	620	660
Floodwater	160		160	40	80
Recycled wastewater	241		241	110	275
Losses				−60	−40
Total water supply	1,843	252	2,095	2,050	2,090

Source: State Comptroller (1990). Report on the management of water in Israel; Jerusalem. Schwarz, J. (1990). Management of the water resources of Israel. *Israel Journal of Earth Sciences* 39:57–65.

17.5 URBAN WATER INFRASTRUCTURE IN THE NEGEV DESERT

The National Water Carrier System supplies water to the population in the northern and central Negev Desert. The town of Mitzpeh Ramon (Fig. 17.8) is the southernmost urban community that receives water from the above system. Though the government operates the water supply network through the Mekorot Company, maintenance of the system inside villages and towns is the latter's responsibility. The municipalities have to plan their secondary distribution network from the point of the Mekorot outlet. The consumers receive water on the basis of annual, seasonal, and monthly allocations.

Beer-Sheva is a rapidly expanding city situated in the northern Negev Desert. The climate is arid with hot, dry summers and mild winters; the average annual rainfall is 200 mm (8 in). The name of Beer-Sheva is first mentioned in the Book of Genesis in relation to the patriarchs: Abraham, Isaac, and Jacob. The modern town of Beer-Sheva was founded in 1900 during the last phase of the Ottoman period. The population increased from a mere 8300 in 1950 to more than 180,000 in 1999. Beer-Sheva is the fourth largest city in Israel, after Tel-Aviv, Jerusalem, and Haifa.

The western part of the Mountain Aquifer extends south to Beer-Sheva. Its main reserve is situated in the Yarkon-Tanninim basin east of the Coastal Plain.

The aquifer is of a karstic nature with high conductivity and swift flows (Grinwald and Bibas, 1989). Use of the Yarkon-Tanninim aquifer increased rapidly since the 1950s, reaching annual withdrawal levels of more than 400 mcm. These figures are higher than the long-term safe yield estimated at 310 mcm/year (Schwarz, 1990). Preservation of its water quality is of great importance, also for Beer-Sheva. Pumping from groundwater in the area for urban use in the city causes a northward flow of brackish groundwater from the central Negev.

During 1998 Beer-Sheva received some 18 mcm of water from the Mekorot Water Company. Urban consumption of water in Israel is based on an allocation of 100 to 180 m^3 of water per household per year. The average person in Israel uses some 135 L of water per day, considerably less than in the United States.

Treated wastewater can be used in the urban sector to replace freshwater for the irrigation of municipal parks and street vegetation. The city of Beer-Sheva is planning a new neighborhood with 8000 dwellings, in which treated wastewater, recycled from its own use of urban water, will be used in combination with modern drip irrigation techniques. A completely separate piping system will be built to convey the purified wastewater from the city treatment plant to the new neighborhood (Bruins, 2000c).

Urban water supply in the Arava Rift Valley (Fig. 17.8), situated in the eastern part of the Negev Desert, is completely different. This area is not connected to the National Water Carrier System. Here desalination of local brackish groundwater constitutes the basis of the water supply system. The town of Eilat on the shore of the Red Sea receives most of its water from the desalination of seawater.

The Arava Rift Valley has a hyperarid climate (Table 17.2); the average annual rainfall is very low, being 30 mm (1 in) in Eilat, slightly increasing to about 50 mm (2 in) in the northern part. The mountains in Jordan to the east of the Arava Rift Valley receive more rainfall, up to 500 mm (20 in) in the north and up to 200 mm (8 in) in the south (Salameh and Bannayan, 1993). The potential evaporation in the Arava Valley is extreme and reaches up to 5000 mm/year, due to the high temperatures, low humidity, and prevailing northern winds. Precipitation and runoff vary greatly from year to year and are highly unpredictable. The annual groundwater recharge is estimated at 44 mcm/year. The safe yield is about 30 mcm/year, an amount more or less used by Jordan and Israel.

The groundwater reservoir in the Arava Rift Valley is shared by Israel and Jordan. The geologic structure is quite intricate, also due to the north-south strike-slip movement of the two continental plates along the length of the valley. In terms of surface hydrology the region is divided into two parts, the northern Arava draining toward the Dead Sea and the southern Arava draining toward the Red Sea. The three major aquifers in the area are the Shallow Aquifer Hydraulic Complex, the Cretaceous Hydraulic Complex, and the Deep Sandstone Aquifer Hydraulic Complex. Recharge of the groundwater comes from various sources, including water of different ages and qualities (Issar, 1985; Adar et al., 1992; Salameh and Bannayan, 1993).

The upper aquifer complex comprises Quaternary and Late Tertiary sedimentary and alluvial deposits, which may reach a depth of 550 m. Annual recharge from occasional rainstorms and wadi flows is important. Five dams have been built at various places in the Arava Valley to arrest the floodwaters in surface basins. Laronne (1990) investigates the functioning and sedimentation rates in such basins in Israel. The upper aquifer complex may also receive recharge from lower aquifers along faults and fractures. The majority of the wells in the area tap this aquifer system. Adar et al. (1992) made a quantitative assessment of recharge in the Israeli part of the southern Arava Valley. Total recharge is estimated at 12.5 mcm/year, most of which is derived from the deep Nubian sandstone aquifer. Only some 20 percent of the recharge is contributed by rainfall and wadi flows, mainly originating from the Jordanian mountains to the east. The latter recharge, however, is the most important source of freshwater.

The Cretaceous Hydraulic Complex consists of an upper and lower group, each bearing different characteristics. The aquifer has generally been developed for irrigation and water supply schemes. Recharge comes from ephemeral flows in wadis, particularly from the high Edom Mountains in Jordan that receive more rainfall than the lower Negev hills. There may also be recharge flows from the deep Nubian sandstone aquifer to the Cretaceous aquifer.

The Deep Sandstone Aquifer Hydraulic Complex receives minor recharge from local rainfall. Some recharge originates in northwestern Saudi Arabia and in Sinai, but a significant part of the water appears to be fossil (Issar, 1985; Salameh and Bannayan, 1993). Development of this aquifer in Jordan at Disi-Mudawwara for water supply to Amman is being considered by the Jordanian government. The total abstraction at present, including by Saudi Arabia, exceeds the safe yield of the aquifer. The deepest drilling for groundwater in Israel to a depth of 1500 m was made near Paran in the central Arava Valley (Fig. 17.9). The water has a temperature of about 60°C and is derived from the deep Nubian sandstone aquifer complex. Water quality is quite good, containing about 250 to 300 mg chlorine per liter, as well as some iron and sulfur (Yossi Schmaya, 2001, personal communication, Mekorot).

The quality of the groundwater in the Arava Rift Valley varies from almost fresh to saline brines. The upper aquifer usually has the best water quality. Nevertheless, desalination is required for urban and domestic water supply. In 1996 there were some 40 reverse osmosis units operational in the Arava Rift Valley. Each village has its own desalination unit, which is only for drinking water (Fig. 17.10). The desalination of brackish groundwater in each reverse osmosis unit is about 5 m^3/h with a maximum output of 100 m^3/day. A dual piping system brings both desalinated water and regular groundwater to each home. The desalinated water is used for drinking and cooking. The other piping system provides slightly brackish groundwater for all other household requirements, including showers, washing, toilet flushing, cleaning, and gardening. The Mekorot Water Company supplies all the water and is responsible for the pumping, the piping network, and all water treatments (Yossi Schmaya, 2001, personal communication, Mekorot).

FIGURE 17.9 Pumping of (warm) groundwater in the Arava Rift Valley from a depth of 1500 m. (Photograph H. J. Bruins)

The cost difference between desalination of relatively low salinity brackish water and seawater is in the range of 1 to 4. Resources of brackish water are limited, however, so seawater desalination is also required for water supply to the town of Eilat at the Red Sea coast. The year 1962 marks the onset for seawater desalination in Israel, as it was decided to construct a Multi-Stage Flash distillation plant near Eilat with a capacity of 3800 m^3/day. The actual supply of desalinated seawater to the town began in 1965. The first reverse osmosis plant to desalinate brackish water, having a capacity of 200 m^3/day, was constructed in 1968 at Yotvata, 40 km north of Eilat. Another Multi-Stage Flash distillation plant for seawater desalination in Eilat was constructed in 1969 to supply more water to the growing population (Levite, 1973; Arad et al., 1973; Glueckstern, 1996).

A more sophisticated reverse osmosis seawater desalination plant has begun operating in Eilat in 1997, producing at present 10,000 m^3/day. An additional module with the same capacity is being planned to double the output in the near future. The net daily production of desalinated water in the southern Arava Valley is at present 46,000 m^3, of which 36,000 m^3 is derived from brackish groundwater and the remainder from the Red Sea. The town of Eilat uses about 14 mcm of desalinated water per year and produces some 7 mcm of treated wastewater, which is used in agriculture. Farmers in the Arava Valley mix different qualities of groundwater to reach the appropriate level of salinity for a certain crop. Whereas the domestic water supply to the villages in the Arava Rift Valley is based on a dual piping system—one for desalinated water and the other for slightly brackish groundwater—the town of Eilat has only one urban piping system containing desalinated water.

FIGURE 17.10 Desalination of brackish groundwater for drinking water in the village of Ein Yahav in the Arava Rift Valley. (Photograph H. J. Bruins)

17.6 SUMMARIZING COMMENTS AND CONCLUSIONS

Israel has a long history of urban water supply systems in dryland and desert areas. Ancient Hebrew texts in the Bible and rock inscriptions describe the planning and building, some 2700 years ago, of an intricate underground tunnel with a length of 533 m. This subterranean canal diverts water from an intermittent karst spring, situated outside the city walls of ancient Jerusalem, to a reservoir inside the town.

A unique urban water supply system existed in the Negev Desert some 1500 years ago, based on rainwater harvesting and the construction of cisterns and conduits. The settlements were established by the Nabateans as caravan stations for camels and later developed into towns. The peak of development of this rainwater-harvesting civilization occurred during the Byzantine period.

Technological advances in modern times enabled the construction of a National Water Carrier System that transports water from Lake Kinneret in the north of Israel to part of the Negev Desert in the south. Water is initially pump-lifted upward by about 362 m. The precious fluid passes subsequently through a complex system of open canals, siphons, and tunnels. The main conduit is built of prestressed concrete pressure pipes with a diameter of 270 cm (108 in). The length of the main channel is about 140 km. The entire system, including secondary networks, is much longer, reaching the desert town of Mitzpeh Ramon in the central Negev.

The surface of Lake Kinneret is about 210 m below ocean level, and the latter town is situated at about 900 m above sea level. Such altitude differences of 1110 m underline the considerable energy costs involved to operate the National Water Carrier System. Annual water quantities of 450 mcm have often been pumped and delivered through the network. Sustained overutilization of the system, coupled with drought, have led to a water crisis at present. Lake Kinneret and the main aquifers have reached their lowest levels in living memory.

The Arava Rift Valley in the eastern Negev constitutes the hyperarid desert region of Israel. The National Water Carrier System does not reach this area, and the urban water infrastructure is completely different. It is based on the desalination of brackish groundwater and seawater. There are three aquifer systems in the Arava Rift Valley associated with different geologic strata of Quaternary–Late Tertiary, Cretaceous, and older Nubian sandstone layers above the Precambrian basement. The water quality varies considerably from saline brines to almost fresh water.

Each village in the Arava Rift Valley has its own reverse osmosis plant and a dual water distribution network. The output capacity of the desalination unit is modest and reaches 100 m^3/day. Therefore, the desalinated water is only used for drinking water and cooking, supplied in a separate network. Another piping system supplies slightly brackish groundwater for all other household uses, such as washing, cleaning, toilet flushing, and garden irrigation.

The town of Eilat at the Red Sea shore is the first and so far the only town in Israel that receives desalinated seawater, since 1965. Also brackish groundwater is desalinated, but Eilat has a single urban water supply network. Unlike the other Arava Valley villages, desalinated water in Eilat is used for all household purposes. A more sophisticated reverse osmosis seawater desalination plant has begun operating in Eilat in 1997, producing 10,000 m^3/day. The town of Eilat uses about 14 mcm of desalinated water per year and produces some 7 mcm of treated wastewater, which is used for irrigated agriculture in the region.

17.7 ACKNOWLEDGMENTS

My thanks to Mr. Ami Shacham, Director of the Arava Drainage Authority, for useful discussions in his office, together with Mr. Yossi Schmaya, water engineer of Mekorot. I appreciate the advice from my colleagues Prof. Jonathan B. Laronne

(chairman, Israel Association of Water Resources and member of the Department of Geography and Environmental Development, Ben-Gurion University of the Negev) and Dr. Eilon Adar (Department of Environmental Hydrology and Microbiology, Jacob Blaustein Institute for Desert Research, Ben-Gurion University of the Negev).

17.8 REFERENCES

Adar, E., Rosenthal, E., Issar, A. S., and Batelaan, O. (1992). Quantitative assessment of the flow pattern in the southern Arava Valley (Israel) by environmental tracers in a mixing cell model. *Journal of Hydrology* 136:333–354.

Amiran, D. (1994). *Rainfall and Water Policy in Israel.* The Jerusalem Institute for Israel Studies, No. 55 (in Hebrew).

Arad, N., Glueckstern, P., and Kantor, Y. (1973). Desalting plants in operation and under construction in Israel. In J. Bonné, S. Grossman-Pines, and Z. Grinwald (eds.), *Water in Israel. Part A, Selected Articles,* Tel-Aviv: Ministry of Agriculture, Water Commission, Water Allocation Department, pp. 143–151.

Arlosoroff, S. (1974). Legal, administrative and economical means for the preservation and efficient use of water in Israel. In *Water in Israel* (Reprints), Ministry of Agriculture, Water Commission.

Broshi, M. (1980). The population of western Palestine in the Roman-Byzantine period. *Bulletin of the American Schools of Oriental Research (BASOR)* 236:1–10.

Bruins, H. J. (1986). *Desert Environment and Agriculture in the Central Negev and Kadesh-Barnea during Historical Times.* Published Ph.D. dissertation, The University of Wageningen, The Netherlands; Stichting Midbar Foundation, Nijkerk.

Bruins, H. J. (1990). The impact of man and climate on the Central Negev and northeastern Sinai deserts during the Late Holocene. In S. Bottema, G. Entjes-Nieborg, and W. van Zeist (eds.), *Man's Role in the Shaping of the Eastern Mediterranean Landscape.* Rotterdam: Balkema, pp. 87–99.

Bruins, H. J. (1999). Drought management and water supply systems in Israel. In E. Cabrera and J. García-Serra (eds.), *Drought Management Planning in Water Supply Systems.* Dordrecht / Boston / London: Kluwer Academic Publishers, pp. 299–321.

Bruins, H. J. (2000a). Proactive contingency planning vis-à-vis declining water security in the 21st century. *Journal of Contingencies and Crisis Management* 8:63–72.

Bruins, H. J. (2000b). Drought hazards in Israel and Jordan: policy recommendations for disaster mitigation. In D. A. Wilhite (ed.), *Drought: A Global Assessment,* Vol. II, *Hazards and Disasters Series.* London: Routledge, chap. 42, pp. 178–193.

Bruins, H. J. (2000c). La eficiencia en la gestion del agua en Israel y la ciudad de Beer-Sheva. In A. Estevan and V. Viñuales (eds.), *La Eficiencia del Agua en las Ciudades.* Zaragonza: Bakeaz, Bilbao & Fundación Ecología y Desarrollo, pp. 205–234.

Bruins, H. J., and Lithwick, H. (1998). Proactive planning and interactive management in arid frontier development. In H. J. Bruins and H. Lithwick (eds.), *The Arid Frontier—Interactive Management of Environment and Development.* Dordrecht / Boston / London: Kluwer Academic Publishers, chap. 1, pp. 3–29.

Bruins, H. J., and Berliner, P. R. (1998). Bioclimatic aridity, climatic variability, drought and desertification. In H. J. Bruins and H. Lithwick (eds.), *The Arid Frontier—Interactive*

Management of Environment and Development. Dordrecht / Boston / London: Kluwer Academic Publishers, chap. 5, pp. 97–116.

Cahill, J. M., and Tarler, D. (1994). Excavations directed by Yigal Shiloh at the City of David, 1978–1985. In H. Geva (ed.), *Ancient Jerusalem Revealed.* Jerusalem: Israel Exploration Society, pp. 31–45.

Evenari, M., L. Shanan, and N. Tadmor (1982). *The Negev: The Challenge of a Desert* (2d expanded edition). Cambridge, Mass.: Harvard University Press.

Glueckstern, P. (1996). Short and long term desalination options for enhancement and quality improvement of water supply. *International Conference of Water Resource Management Strategies in the Middle East.* Association of Engineers and Architects in Israel, Society of Water Engineers, Herzlia, Book of Abstracts.

Grinwald, Z., and M. Bibas (1989). *Water in Israel.* Tel Aviv: Ministry of Agriculture, Water Commission, Water Allocation Department.

Hayward, D. F., and Oguntoyinba, J. S. (1987). *Climatology of West Africa,* London: Hutchinson.

Hillel, D. (1982). *Negev—Land, Water, and Life in a Desert Environment.* New York: Praeger.

Issar, A. S. (1985). Fossil water under the Sinai Negev peninsula. *Scientific American* 253:104–111.

Issar, A. S. (1990). Water shall flow from the Rock: Hydrogeology and climate in the lands of the Bible. Berlin: Springer-Verlag.

Kedar, Y. (1957). Water and soil from the desert: some ancient agricultural achievements in the central Negev. *The Geographical Journal* 123:179–187.

Laronne, J. B. (1990). Probability distribution of event sediment yields in the northern Negev, Israel. In J. Boardman, I. D. L. Foster, and J. A. Dearing (eds.), *Soil Erosion on Agricultural Land.* New York: John Wiley.

Levite, G. A. (1973). Towards the era of desalination. In J. Bonné, S. Grossman-Pines, and Z. Grinwald (eds.), *Water in Israel. Part A, Selected Articles.* Tel-Aviv: Ministry of Agriculture, Water Commission, Water Allocation Department, pp. 135–142.

Lowdermilk, W. C. (1946). *Palestine, Land of Promise.* London: Golancz.

Middleton, N., and Thomas, D. (1997). *World Atlas of Desertification* (2d ed). London: Arnold.

Negev, A. (1986). *Nabataean Archaeology Today.* New York: University Press.

Orni, E. and Efrat, E. (1971). *Geography of Israel.* Jerusalem: Israel Universities Press.

Penman, H. L. (1948). Natural evaporation from open water, bare soil and grass. *Proceedings of the Royal Society, Section A,* 193:120–145.

Salameh, E., and Bannayan, H. (1993). *Water Resources of Jordan: Present Status and Future Potentials.* Amman: Friedrich Ebert Stiftung.

Schwarz, J. (1990). Management of the water resources of Israel. *Israel Journal of Earth Sciences* 39:57–65.

Shiloh, Y. (1994). Excavations directed by Yigal Shiloh at the City of David, 1978–1985. In H. Geva (ed.), *Ancient Jerusalem Revealed.* Jerusalem: Israel Exploration Society, pp. 46–54.

State Comptroller (1990). Report on the Management of Water Resources in Israel. State Comptroller, Jerusalem (in Hebrew).

Tamir, O. (1973). Administrative and legal aspects of water use in Israel. In J. Bonné, S. Grossman-Pines, and Z. Grinwald (eds.), *Water in Israel. Part A, Selected Articles,* Tel-Aviv: Ministry of Agriculture, Water Commission, Water Allocation Department, pp. 171–180.

Thornthwaite, C. W. (1948). An approach towards a rational classification of climate. *Geographical Review* 38:55–94.

UNESCO (1979). Map of the world distribution of arid regions. Explanatory note. Paris: Unesco, Man and the Biosphere (MAB). Technical Notes 7.

Vilnay, Z. (1971). *The Guide to Israel.* (4th ed.). Jerusalem: Hamakor Press.

Wilhite, D. A. (ed.) (2000). *Drought: A Global Assessment,* Vol. I & II, *Hazards and Disasters Series,* London: Routledge.

Wolf, A. T. (1996). Middle East Water Conflicts and Directions for Conflict Resolution, *Food, Agriculture, and the Environment Discussion Paper 12,* International Food Policy Research Institute, Washington, D.C.

Yaalon, D. H., and Dan, J. (1974). Accumulation and distribution of loess-derived deposits in the semi-arid and desert fringe areas of Israel. *Zeitschrift für Geomorphologie Suppl. Bd.* 20:91–105

CHAPTER 18

USING GIS AND HYDRAULIC MODELING TO EVALUATE SUSCEPTIBILITY OF WATER DISTRIBUTION SYSTEMS TO INTRUSIONS: A CASE STUDY

Trevor R. Lindley
Malcolm Pirnie, Inc.
Phoenix, Arizona

Steven G. Buchberger
University of Cincinnati
Cincinnati, Ohio

18.1 INTRODUCTION

From 1971 to 1998, 113 of 619 documented waterborne illnesses were attributed to distribution system deficiencies (Craun and Calderon, 2001). These deficiencies include backflow at connections (both controlled and uncontrolled), contamination during infrastructure repair, pathogen entry along buried infrastructure during periods of low pressure, dissolution of pipe material (e.g., copper or lead), and contamination of storage tanks. The World Health Organization (WHO) reports contamination of distribution pipelines may arise from intermittent supply, low pressure in the distribution network, leaking pipes, and inadequate wastewater collection systems (WHO, 2000); these conditions make the distribution system susceptible to contaminant intrusions. A *distribution system intrusion* is defined here as the unintentional introduction of an external contaminant into the system at some location after treatment. Intrusion susceptibility, even in a well-operated system, provides the basis for maintaining a distribution system chlorine residual in some countries (Trussel, 1999; Haas, 1999; LeChavallier, 1999).

To overcome these deficiencies and minimize the risk of distribution system contamination, water utilities implement *best management practices* (BMPs). BMPs minimize intrusion susceptibility by eliminating deficiencies that lead to a possible contamination event. Table 18.1 shows common management practices employed to minimize contamination of drinking water distribution systems. At any location in the distribution system, the failure of one or more of these practices heightens the possibility of an intrusion event; the failure of multiple practices at the same location will likely lead to contamination. Multiple failures of common operating practices led to one noteworthy and tragic distribution system waterborne disease outbreak in Cabool, Missouri, in 1989 and 1990 (Geldreich, 1996; Clark, 2000).

Eliminating intrusion susceptibility can be challenging; the task is complicated by the age of the systems, the buried nature of the facilities, ongoing funding issues, and the system's spatial expansiveness. Two tools, namely geographic information systems (GISs) and hydraulic network modeling are being used by utilities to overcome these (and other) challenges.

18.2 GIS AND HYDRAULIC MODELING

Historically, data associated with a utility's GIS and data associated with the utility's hydraulic modeling system have been maintained separately. Key components typically maintained in the GIS realm (i.e., main locations and attributes, junction elevations, land-use for demands) are often used to build the hydraulic model; however, use of other data components is rare. In addition, hydraulic model results are infrequently passed back to the GIS to support further spatial analysis in concert with other layers. The one-way flow of information does not utilize the abundance of information available in many existing GISs. Even utilities with newly established GIS and modeling capabilities can correlate seemingly unrelated data layers to increase the applicability of the hydraulic model and effectively apply the modeled results. Barcellos (2000) used a GIS analysis to investigate health outcomes related to urban sanitation conditions in Brazil. Each information source was a layer with a purpose and characteristic that allowed for spatial operations and population-at-risk calculations in a GIS environment. Others have spatially correlated bacteriological excursions to local hydraulic conditions observed via hydraulic modeling (McMath and Casey, 2000). Water quality concerns researched through hydraulic and water quality modeling have been linked to the served population in a GIS environment or GIS-type analysis (Maslia et al., 2000; Clark, 2000).

TABLE 18.1 Common Management Practices for Minimizing Contaminant Intrusions*

Management practice	Typical value or practice	Reference for management practice	Potential result if practice is not employed
Maintenance of minimum pressure	20 psi (140 kPa)	Ten-state standard (GLUMRB, 1997)	Low distribution system pressure increases backflow risk at connections; extremely low pressure may result in intrusions along buried infrastructure
Backflow prevention device installed on connections deemed hazardous Inspection and maintenance of backflow preventer	Air gap, reduced pressure zone backflow preventer, double check valves-select control device based on perceived risk	AWWA standards C510 and C511 (AWWA, 1997a, b) AWWA Manual M14 (AWWA, 1990) *Manual of Cross Connection Control* (FCCHR, 1993) *EPA Cross Connection Control Manual* (USEPA, 1989)	Hazardous processes (industry, hospitals, mortuaries, etc.) connected to the potable supply may operate at pressures higher than potable water mains.
Separation distance between water mains and pathogen source	10 ft (3.05 m) for water mains parallel to sewer lines or sewer force mains 18 in (46 cm) for vertical separation for water main crossing a sewer	Ten-state standard (GLUMRB, 1997	Sewer systems (storm, sanitary, combined) typically exhibit some leakage (magnitude depending on construction practices and age) at joints or crack. Pathogens may exist exterior to these systems (AWWARF, 2001). The pathogens could be very near the exterior of a distribution system if placed too close together

TABLE 18.1 Common Management Practices for Minimizing Contaminant Intrusions* (*Continued*)

Management practice	Typical value or practice	Reference for management practice	Potential result if practice is not employed
Control main leakage	Target < 10–15% unaccounted for water	Varies by utility	Mains with high leak rates (manifest by high break rates or pitometer surveys) are mains subject to intrusion during a low-pressure situation within the main
Covered storage facilities	All finished water reservoirs should be covered	*Guidance Manual for Covered Storage* (AWWARF, 1999) AWWA standard D310 (AWWA, 1996) Ten-state standard (GLUMRB, 1997)	Uncovered reservoirs or improperly maintained hatches on reservoirs provide an easy entry point to contamination
Disinfection of new or repaired facilities	25 mg/L chlorine for 24–48 h 10 mg/L residual after 24 h 50–100 mg/L for 3 h	AWWA standard C651 (AWWA, 1992a) AWWA standard C652 (AWWA, 1992b)	Inadequate chlorination may result in system contamination after main is placed on-line. Contamination may come from surrounding environment

*The AWWARF report *Pathogen Intrusion into the Distribution System* (2000) provides a comprehensive review of distribution system management practices.

What follows is a case study of a drinking water distribution system intrusion susceptibility analysis. The objective of the study was to identify locations in the drinking water distribution system where intrusion susceptibility exists or where the BMPs noted in Table 18.1 are not being met. The actual analysis is a series of spatial data queries and hydraulic modeling that are designed to yield the location(s) of intrusion susceptibility due to the presence of distribution system deficiencies. Once susceptible locations were identified, an intrusion analysis was performed to investigate the hydraulic connectivity between the locations susceptible to intrusion events and nearby sensitive populations. The intrusion modeling serves as a prioritization mechanism for addressing the susceptible locations posing the greatest risk to the consumers.

The information reported here focuses on potential contamination due to intrusions at locations along buried infrastructure (i.e., water mains) and intrusions due to backflow at a service connection. The case study involved the following steps:

1. Data needs assessment
2. Spatial refinement of hydraulic model
3. Main break and structural assessment
4. Pressure modeling and pressure susceptibility
5. Contaminant source analysis
6. Colocation of susceptibility conditions
7. Sensitive population analysis

18.3 DATA NEEDS ASSESSMENT

Review of Table 18.1 suggests susceptibility is based on three critical susceptibility conditions (Lindley and Buchberger, 2002):

- Adverse pressure gradient
- Contaminant intrusion pathway (i.e., hydraulic connection)
- Contaminant source

Identifying the *locations* where one or more of these conditions exist is paramount and the crux of the susceptibility analysis. Table 18.2 summarizes these three conditions, system deficiencies that generate them, and some spatial data layers that may assist in objectively identifying where these three conditions exist. In the case study, one or more spatial layer(s) represent each of these conditions. As more susceptibility conditions—represented as data layer(s)—are manifest at a given location, the intrusion susceptibility increases as depicted in Fig. 18.1.

TABLE 18.2 Critical Conditions Leading to an Intrusion Event and Available Data to Identify Locations with Critical Conditions

Critical condition	Deficiencies contributing to critical condition	Data source related to deficiencies (GIS layer)
Adverse pressure gradient (water main pressure lower than external fluid pressure)	• Deficient system design (routine low pressure) • Main breaks (random low pressure) • Fire flow (random low pressure) • High consumer demands • Extreme consumer process • Transients due to valve shutting, pump cycling, change in flow direction	• Calibrated hydraulic network model to provide pressure information • Main break data to assess break patterns that may influence systemwide pressure
Contaminant intrusion pathway (presence of a pathway between treated potable water in the main and external contaminant source)	• Leaking water main • Faulty appurtenance • Open storage reservoir • Failed backflow preventer* • Improper disinfection during main repair/replacement	• Main break data may suggest mains that are not structurally sound and subject to intrusion during low pressure • Locations of all connections*
Contaminant source (existence of a reservoir of water/fluid or substance that would degrade water quality if allowed into a water main)	• Water mains buried near sewer lines† • Sewer overflow • Septic systems, leach fields • Hazardous process fluids in user facilities (assuming these processes are connected to the • Submerged water mains (high groundwater or other waterways)	• Locations of sewers and septic systems • Locations of streams, lakes, high groundwater, or other water bodies • Locations of hazardous connections

**All* distribution system service connections including residential, commercial, and industrial connections create an intrusion pathway. Certain types of connections are typically backflow protected due to perceived risk of internal process liquids. Residential connections typically do not have backflow prevention.

†Refers to storm, sanitary, or combined sewers.

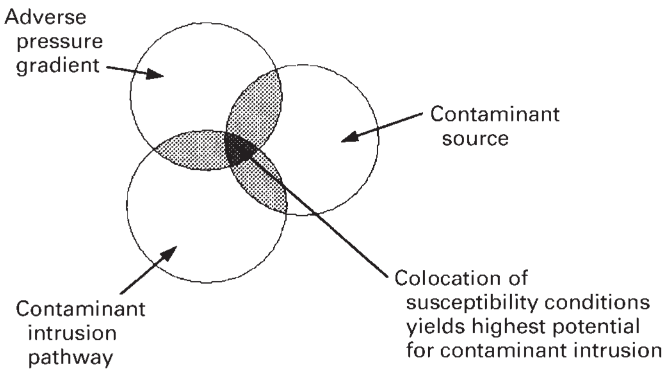


FIGURE 18.1 Susceptibility conditions contributing to distribution system intrusions.

At the conclusion of the data needs assessment, a utility having many of the required data layers was identified. A portion of the utility's service area was selected for application of the susceptibility analysis. The study area is typical of urban and suburban service areas in North America.The study area encompasses approximately 38 mi^2 (98 km^2), includes three pressure zones, with 282 mi (454 km) of water mains (down to service connections), 18,900 connections, and an average demand of 8.3 million gallons per day (mgd) (31,400 m^3 day). The hydraulic model used for this analysis is a skeleton; it contains 100 percent of the mains greater than 12 in (30.5-cm) in diameter, 87 percent of the 12-in (30.5-cm) mains, 64 percent of the 8-in (20.3-cm) mains, and 12 percent of the 6-in (15.2-cm) mains. The total length of mains represented in the hydraulic model is 150 mi. The hydraulic model has seen recent use by the utility for planning purposes.

The demand data including the patterns used in the extended period simulations (EPS) were also provided by the utility. All hydraulic analysis was done using the public domain EPANET software (Rossman, 2000). The GIS layers used in the susceptibility analysis were acquired from the utility and a regional GIS where the water utility is located. All spatial analysis was done using ArcView GIS.

18.4 SPATIAL REFINEMENT OF THE HYDRAULIC MODEL

Many utility distribution system hydraulic models are skeletons of the actual system; the case study model was no exception. A hydraulic model's anticipated use (e.g., capital planning) usually drives the level of detail built into the model at the outset. Traditional hydraulic skeletons only have spatial accuracy at the nodes; the nodes are connected by straight-line links without true spacial representation. This approach is adequate from a hydraulic standpoint but inadequate to utilize the abundant data layers in a GIS where location is the fundamental link between layers.

The existing skeleton model was refined to more accurately represent the actual system mains. This served to increase the number of model nodes (and model pipe links) in the system, which in turn allowed for a more refined pressure surface. The *x, y* coordinates for the new nodes were found by digitizing within the model against a backdrop of water mains maintained for utility management in the regional GIS. The elevations for the new nodes were found by matching the new nodes to an existing GIS layer of elevation point data. All nodes now inherit the main attributes where they are located (i.e., each node has attributes of *x, y* coordinates, elevation, diameter, vintage, material, etc.). The new nodes are zero demand nodes that are essentially placeholders for hydraulic modeling purposes. Adding the new nodes and model links where they actually occur in space moved the hydraulic model from a traditional skeleton, where only a handful of nodes are modeled with straight-line link connections, to a skeleton whose modeled pipes actually have spatial representation.

Table 18.3 summarizes the changes brought about by the spatial refinement. The level of effort for digitizing the new links and nodes and assigning the elevations to the nodes was approximately one full-time equivalent for 1 week.

18.5 MAIN BREAK AND STRUCTURAL ASSESSMENT

Water main breaks are a common occurrence in drinking water distribution systems. With regards to intrusion susceptibility, water main breaks provide information for two of the three susceptibility conditions, namely, susceptibility due to low pressure and susceptibility due to the presence of an intrusion pathway.

Water main breaks, depending on the size and duration, may result in a significant drop in the pressure surface. This pressure drop can occur at a location some

TABLE 18.3 Summary of Additional Spatial Refinement

Model item	Original skeleton	New skeleton
No. of nodes	489	1641
No. of links	600	1751
Total link length (ft)	149.8	149.4
Computational time (approximate seconds)*	5	15

Clock time to run model in EPANET on a Pentium III processor, under utility provided operating scenario, 168-h run duration, with a 30-min hydraulic time step (no quality modeling in this example).

distance away from where the break has occurred. Even a well-designed and well-operated system without pressure problems during normal operations may have critically low pressure for short periods during a main break situation. Besides pressure susceptibility, main breaks provide an indicator of which mains have a pattern of structural deficiency. Structurally deficient lines are more apt to experience intrusions during depressurization. Well-documented water losses due to main leaks can also provide an indication of which mains are pathways for contaminant intrusions. Using the main break data as a predictor for mains that are contaminant pathways implies that the mains with higher break rates have higher leak rates and water leaking out during positive pressure may be drawn back into the main during a depressurization event (this assumes that mains leak before they fail and repair is necessary).

Estimating the structural integrity of buried infrastructure is challenging (Hadzilacos et al, 2000; Fenner and Sweeting, 1999; Kleiner and Rajani, 1999; Habibian, 1992, AWWARF, 1986). Much of the structural analysis shown here is based on the guidance given in an AWWARF research report (1986) using available utility maintenance and repair data. The repair data was provided as a data file of maintenance events located through the service area. These points were then matched to repaired water mains and other environmental variables noted to influence water main failure (AWWARF, 1986). Main attributes of diameter, vintage, and material were investigated along with bedding slope and bedding soil type. Figure 18.2 shows the distribution of main breaks across the system. The main breaks represented in Fig. 18.2 show all repair events conducted in the service area from 1985 to 1999. Repair rates by diameter were tabulated directly from the repair history to the existence of reported repair diameter in the maintenance database. Repair rate analysis with respect to vintage and main material required matching the repair point location to the nearest water main within the GIS (this information was not included in the utility repair data). This matching yielded some locations where the repair date preceded the matched main's vintage (e.g., a repair occurred in 1987 and the nearest main was installed in 1996) suggesting the entire main had since been replaced. In this instance the vintage and material of the line experiencing the break could not be ascertained. From the 748 total breaks, 520 were cleanly matched to yield information on vintage and material. The environmental variables of bedding slope and soil type were matched to the total repair set (748 events). This matching of breaks to soil type and slope was done completely within the GIS.

For all main break variables the results were normalized against the length of pipe in the respective category in existence for each year of the break record [e.g., 6-in (15.2-cm) breaks in 1985 per miles of 6-in (15.2-cm) pipe existing in 1985]. This activity resulted in a break rate measured in terms of breaks per mile for each year and category.

The results suggest for the study area that diameter, material, and vintage play important roles in main breaks. However, those vintages with poor break histories were all of the same pipe material; thus the same information is available from two

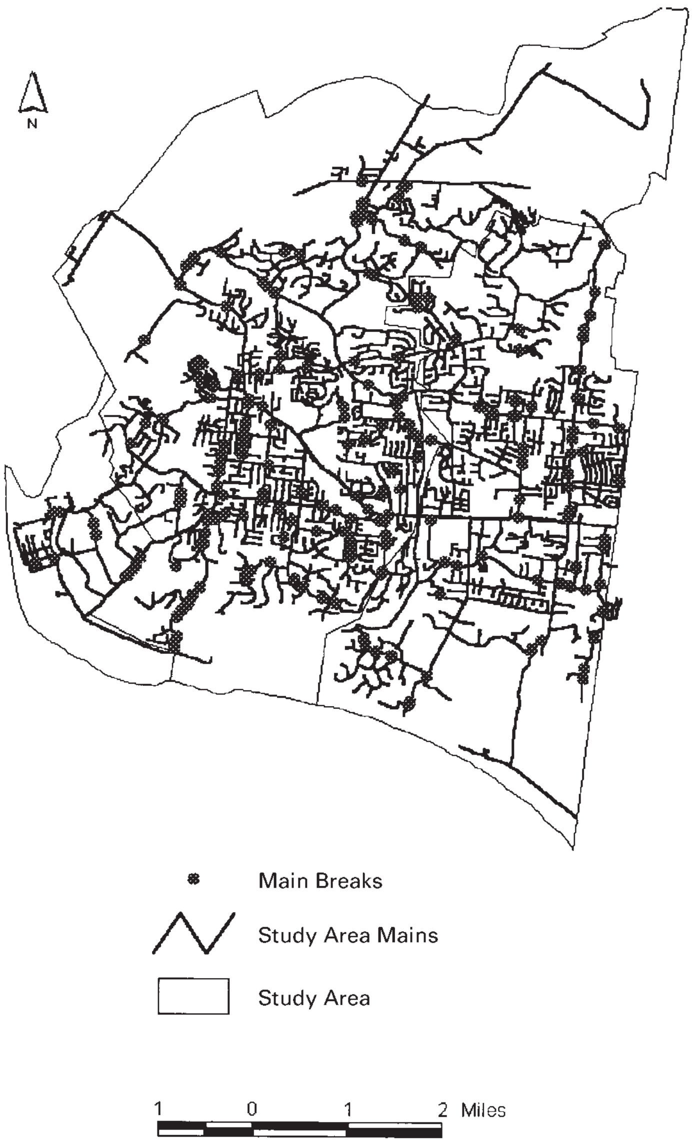


FIGURE 18.2 Spatial distribution of main breaks within study area.

key variables, diameter and vintage (in general, over 90 percent of the breaks were in cast-iron pipe all installed prior to 1975). The structural findings for this study area are in line with observations by the utility relative to water mains with less than favorable break histories. Six-inch-diameter pipes installed between 1940 and 1970 have break rates much higher than the nearest category investigated.

These mains were assumed as probable pathways for intrusions. It is noted that even mains with excellent repair histories may have low leak rates and could be pathways for intrusion during depressurization events.

A reduced set of repair events (281 events) applicable only to water mains represented in the model was analyzed in a similar fashion. These results were used to simulate a historically defensible main break scenario in the hydraulic model; that is, not all main breaks were used to simulate breaks due to the skeletal nature of the hydraulic model. Figure 18.3 shows the total monthly average break rates for modeled mains irrespective of diameter and vintage (based on 281 main breaks and 150 mi of modeled water mains). Figures 18.4 and 18.5, respectively, show the number of breaks per 100 mi by diameter and vintage. These results are consistent with the utility's observations and results elsewhere.

18.6 PRESSURE MODELING AND PRESSURE SUSCEPTIBILITY

Hydraulic simulations of the study area conducted under routine operating conditions (i.e., average patterns and no main breaks) resulted in no distribution system locations experiencing pressures less than 30 psi (207 kPa). (Certain locations do experience regular low pressure, but these are at pump intakes or in the treatment works; these locations were not considered). To investigate the possibility of lower pressures during times of system stress, the historic main break data was used to

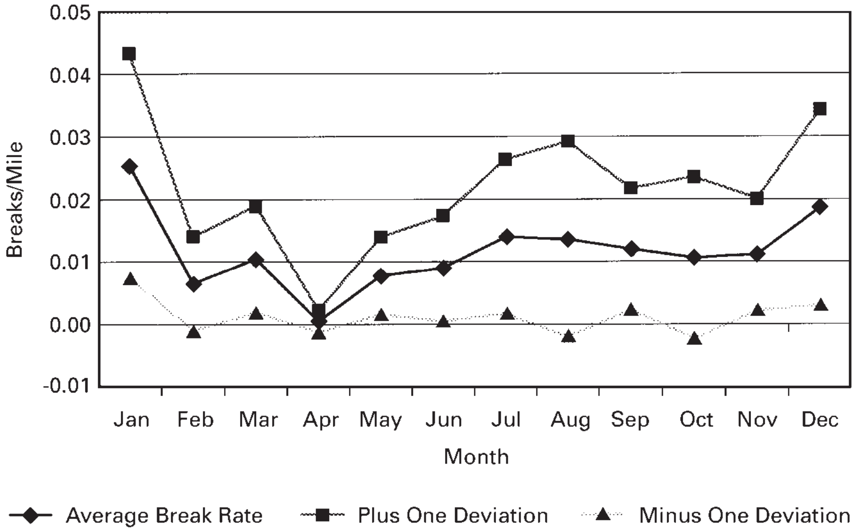


FIGURE 18.3 Distribution of break rates by month for water mains in the study area represented in the hydraulic model ($n = 281$ main breaks).

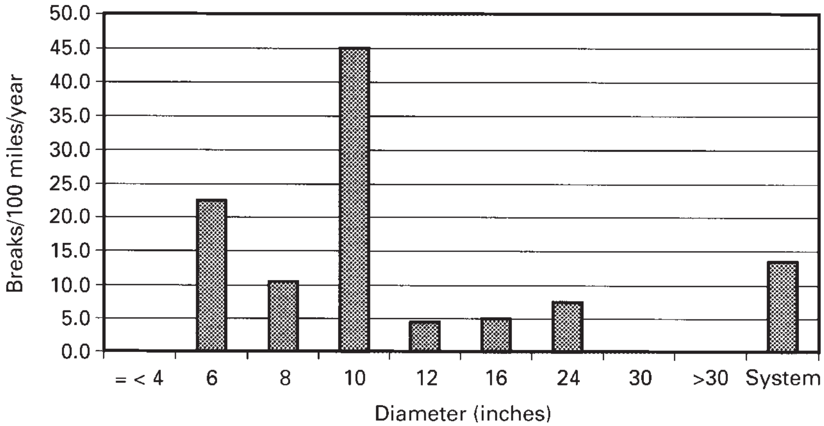


FIGURE 18.4 Average break rates for entire system as a function of diameter. The 10-in rate is obviously much higher than those of the other diameters. However, the study area has only 0.6 mi of 10-in pipe; the majority of the repair activity involves 6-in mains.

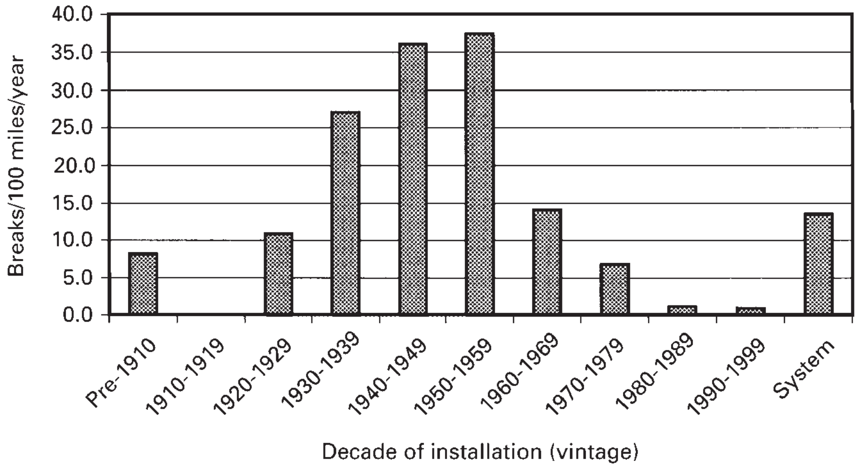


FIGURE 18.5 Average break rates for entire system as a function of vintage.

perturb the system with a large localized demand. These main break demands were coupled with a high systemwide demand pattern equivalent to approximately two times the average demand pattern. This operating scenario is possible during the summer months in the study area investigated.

A *stochastic modeling approach* was used for estimation of the intensity, duration, frequency, and location of low pressure due to the main breaks and summer demands. Stochastic modeling involves inputs of random variables rather than fixed values. The model inputs can take on a range of values based on a probability distribution; this approach captures the inherent variability of processes at work in many water resource applications.

Using the *EPANET Toolkit* (Rossman, 2000) and a Visual Basic programming algorithm, a long-term (1 month) simulation was performed. Water mains were broken in a fashion comparable to the system's main break history to induce pressure stresses expected during an actual break event in a real distribution system. The actual breaking of mains in the model occurs at the nodes. This is done during each simulation; the algorithm returns the number of breaks for the month (from Fig. 18.3), returns the diameter (from Fig. 18.4), and returns a vintage given that break diameter (the conditional probabilities are not provided here). The algorithm then searches the set of model nodes for those matching the diameter and vintage description; the matched node is now a break node. Nodes selected as break nodes were "turned on" or activated using a version of the emitter function in EPANET.

$$q = Cp^{0.5} \tag{18.1}$$

where q = flow rate, gpm
C = discharge coefficient, gpm / $\sqrt{\text{psi}}$
p = pressure, psi

In the algorithm, the discharge coefficient is also a random variable whose range of outcomes, when coupled with typical operating pressures per Eq. (18.1), has a basis in the flow rate of breaks observed by the utility and flow available for fire demands. Table 18.4 shows a summary of the random model inputs and typical outputs from the main break algorithm (larger or smaller values are possible).

The single-month simulation was repeated hundreds of times in the context of a Monte Carlo simulation, generating probability distributions for low pressures. Each 1-month simulation is a historical realization of events, including system demands and main breaks that occur in the distribution system. Figure 18.6 shows study area locations expected to have a high frequency of break occurrences and locations expected to have a high frequency of low-pressure occurrences. Figure 18.6 reveals that locations with higher tendencies for low pressure are not necessarily correlated in space to locations with high tendencies for breakage. In Fig. 18.6, frequency is on a monthly time scale; for an event (main break or low pressure) to be counted, it occurred at least once during a given 1-month simulation. The 1 percent level is equal to low-pressure occurrences in 5 simulations out of 500. The low-pressure locations in Fig. 18.6 are locations susceptible to intrusion events due to the potential for low pressure (assuming 20 psi is low enough to allow some kind of backflow or intrusion to occur).

This result does not include other low-pressure initiation events such as maintenance flushing, fire flows, or hydraulic transients; transients cannot be addressed with a model such as EPANET.

TABLE 18.4 Stochastic Variables within Stochastic Main Break Algorithm

Model variable	Variable probability distribution	Typical output	Basis
Break rate	Truncated normal	2–3 breaks/month	Utility repair history
Break diameter	Historical proportions	22% 6 in, 61% 8 in, 6% 12 in, 10% >12 in	Utility repair history
Break time during the 1-month simulation	Uniform	Random time throughout month	Engineering judgment
Breakflow rates	Log-normal	150–300 gpm	Utility observations, fire flow available
Breakflow duration	Function of flow rate	2–6 h	Utility observations, crew response criteria

18.7 CONTAMINANT SOURCE ANALYSIS

The contamination sources investigated include sanitary sewer lines, septic systems, and known high-risk service connections. For the subsurface contaminant sources, the offset distance between the potable water mains and the sources was analyzed. Figure 18.7 shows the results of this analysis for sanitary sewer lines and septic systems.

A layer of customer service connection information was used to investigate high-risk service connections as a source of contamination. The utility maintains a rigorous cross-connection control program including process risk assessment at the time of connection and yearly inspections of control devices. Data collected by the utility at the time of connection (i.e., information about a customer's process) were not accessible in a digital format. Therefore, explicit identification of connections considered high risk was done by examining the data related to the service's branch connection diameter in the customer information attributes. This resulted in identification of 161 connections out of 18,900 [with service branches greater than 4 in (10.2 cm) in diameter]. These locations were then screened manually (generally by name) to determine if the facilities could be considered high-risk. Twenty-eight facilities were eventually identified as high-risk connections. They are generally industrial or commercial facilities with one or more large-branch connections [>4 in (10.2 cm)] and contain some type of internal process deemed a concern (e.g., hospitals, mortuaries, chemical process locations, manufacturing facilities). All of these are controlled in the utility's cross-connection control program. These locations along with the subsurface pathogen reservoirs (sewer lines and septic systems) near water mains are potential sources of contamination.

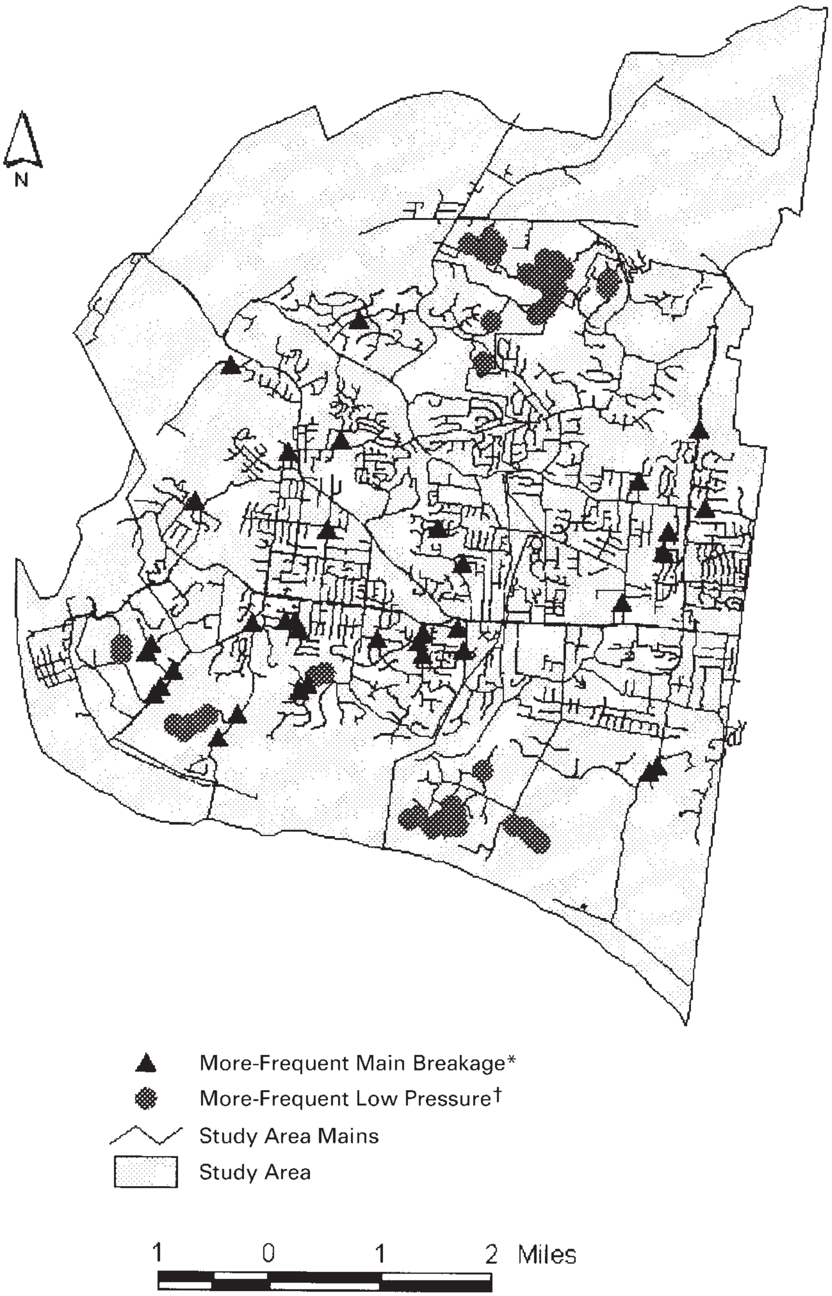


FIGURE 18.6 Locations with high break frequency and high frequency of low pressure occurrence. * = greater than 1 percent chance of break at this location in a month; † = greater than 1 percent chance of experiencing pressure less than 20 psi at this location in a month.

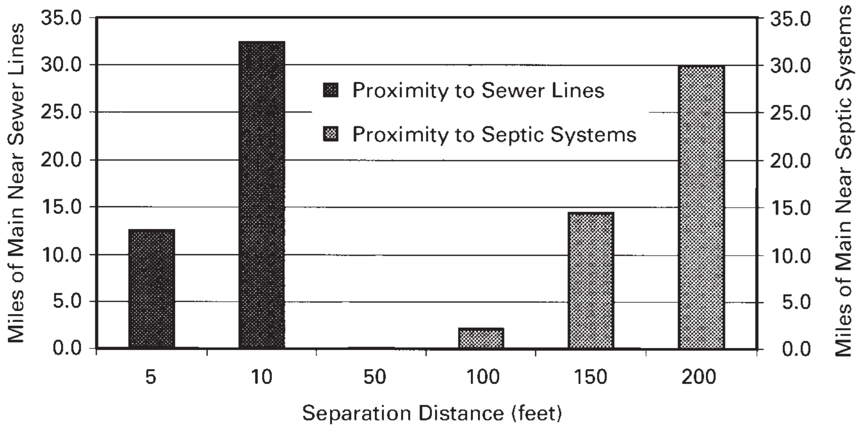


FIGURE 18.7 Proximity of study area water mains to a known subsurface pathogen source.

18.8 COLOCATION OF SUSCEPTIBILITY CONDITIONS

At this point, locations that have an identifiable hydraulic intrusion pathway, experience low pressure during system stress, or are associated with a contaminant source have been identified. These results are for only a single susceptibility condition; these locations were then combined giving a final set of locations deemed more susceptible to contamination by external intrusion because multiple conditions exist at the *same location.* As depicted in Fig. 18.1, the colocation of one, two, or three susceptibility conditions yields increasing potential for intrusion. The layering of the case study area locations having pressure, pathway, or contaminant source (hazardous connections or subsurface contaminant source) concerns was conducted within the GIS. Table 18.5 summarizes the query criteria for these susceptibility conditions and relates the susceptibility conditions to actual observed distribution system intrusion events (not in the study area but observed in other locations throughout North America).

Figure 18.8 shows the results of the colocation of the various susceptibility layers. The legend in Fig. 18.8 refers to the intrusion scenarios in Table 18.5.

18.9 SENSITIVE POPULATION ANALYSIS

In the event the susceptibility analysis yields multiple locations of concern (perhaps too many to address at once), consideration of the potentially influenced populations, specifically sensitive populations, serves as a prioritization mechanism. Influence was assumed given a hydraulic connection (via the distribution system)

TABLE 18.5 Colocation of Susceptibility Conditions and the Potential Intrusion Scenarios Addressed by the Analysis

Potential intrusion scenario	Modeled susceptibility conditions and criteria in the study area	Historical evidence of this type of occurrence*
1. Subsurface contaminant intrusion along buried infrastructure	(1) Pressure surface < 0 psi (2) 6-in mains constructed from 1940–1970 (3) Sewer within 10 ft, septic system within 200 ft	Evidence of fecal pollution entering operating water mains during 1946, 1959, 1983 in the United Kingdom (AWWARF, 2000).
2. Subsurface contaminant intrusion during repair	(1) 6-in mains constructed from 1940–1970 (2) Sewer within 10 ft, septic system within 200	Likely contamination during repair, contaminants from nearby sanitary sewer overflow entered repair location; 1989, Cabool, Missouri (Geldreich, 1996).
3. Backflow at uncontrolled (residential) connection	(1) High frequency of pressure < 20 psi (2) Numerous connections served by the location with low pressure	Backsiphonage of uncontrolled fluoride during low pressure caused by water main break; 1993, Hawaii (Craun and Calderon, 2001).
4. Backflow at controlled hazardous connection	(1) Occurrence of pressure < 20 psi (2) Any connection deemed hazardous	Backflow of cooling system chemicals into water system through fault backflow preventer; 1995, California (Craun and Calderon, 2001).

*Noted intrusion scenarioDid not occur in the study area but in other distribution systems with deficiencies.

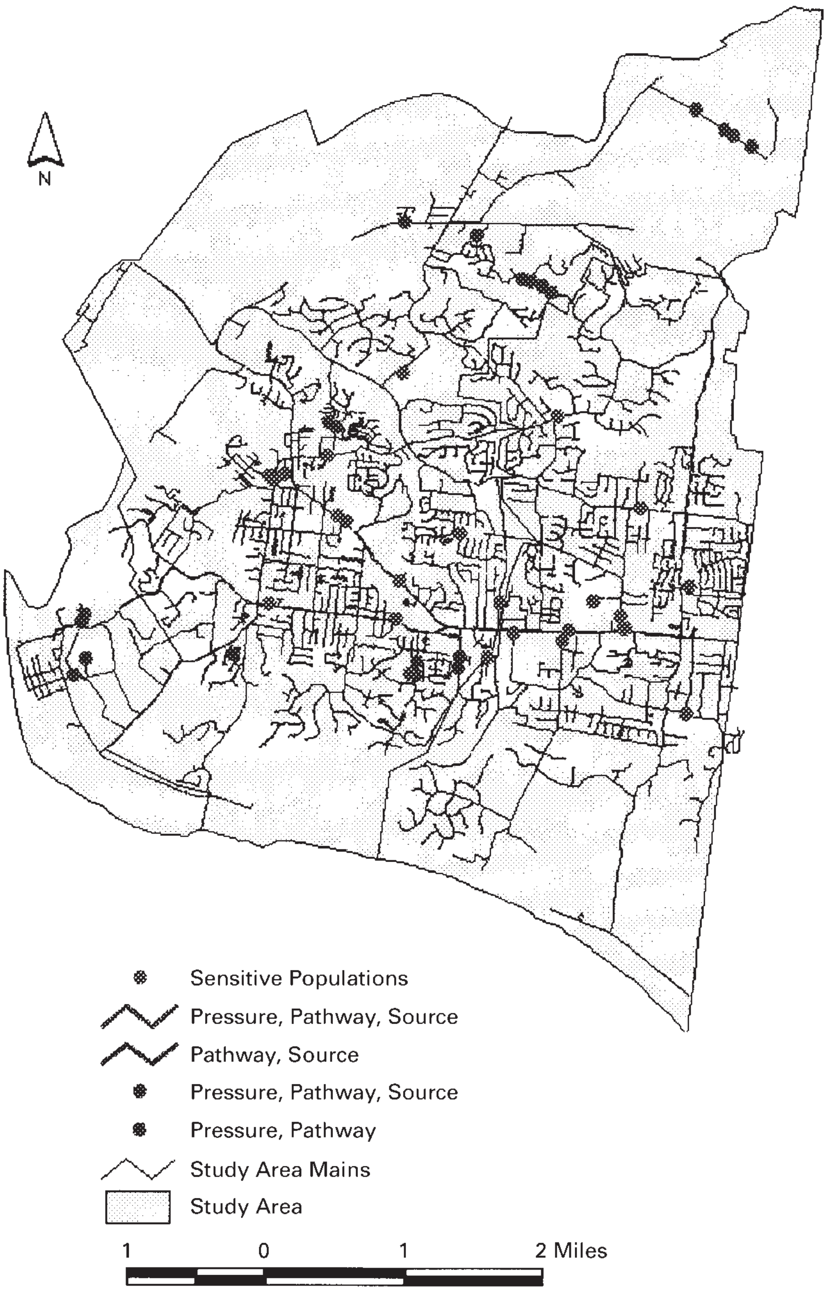


FIGURE 18.8 Colocated susceptibility conditions and sensitive populations.

between the intrusion susceptable location and the sensitive population. Sensitive populations (USEPA, 2000) associated with specific facilities or population centers were identified in the study area. These centers include day care centers, preschools, elementary schools, adult day care centers, retirement communities, and nursing homes (shown in Fig. 18.8). The locations were added to the GIS via geocoding against a matchable street layer (Clark, 1999). These facilities were also surveyed about the number of inhabitants at the location on a typical day and the length of time the facility has been located at its present address. Targeting these populations allows a utility to address a large percentage of the population more susceptible to water quality upsets while managing fewer locations. Hydraulic connectivity was investigated using the trace function in EPANET. Those locations identified as potential intrusion sources were set as source nodes and the number of receptors hydraulically connected to the source was totaled. Of the intrusion susceptible locations set as source nodes, only two were hydraulically connected to sensitive populations.

Besides utilizing the trace function, the main break algorithm developed to assess pressure sensitivity was also used to simulate continuous intrusion events into the distribution system. During the simulation, an intrusion was assumed for any location that experienced pressures less than 1 psi (6.9 kPa). During the interval of low pressure, the intrusion was simulated via placement of a conservative mass inflow (distributed randomly between 0 and 1000 mL/min) into the line at this location. The mass inflow was shut off when the pressure exceeded 1 psi (6.9 kPa). This rule implies that only a low-pressure event is required to initiate a contaminant intrusion. The random mass injected at low-pressure locations spread to other locations connected to the source (no decay was assumed in either bulk flow or at the pipe wall). The amount of mass reaching other points in the network were ranked and plotted on normal probability paper as shown in Fig. 18.9 for one location. Each data point in Fig. 18.9 represents one mass event passing the node being observed. Plotting positions on the figure were identified by ranking each mass value (1 to n); the corresponding probability is determined from

$$p = \frac{m}{n + 1}$$

where p = probability
m = rank
n = number of values in set

The exceedence probability is calculated as $1 - p$. The results shown in Fig. 18.9 are for a location in the northern portion of the study area. A similar graph can be generated for every location receiving mass from nodes where intrusions have occurred.

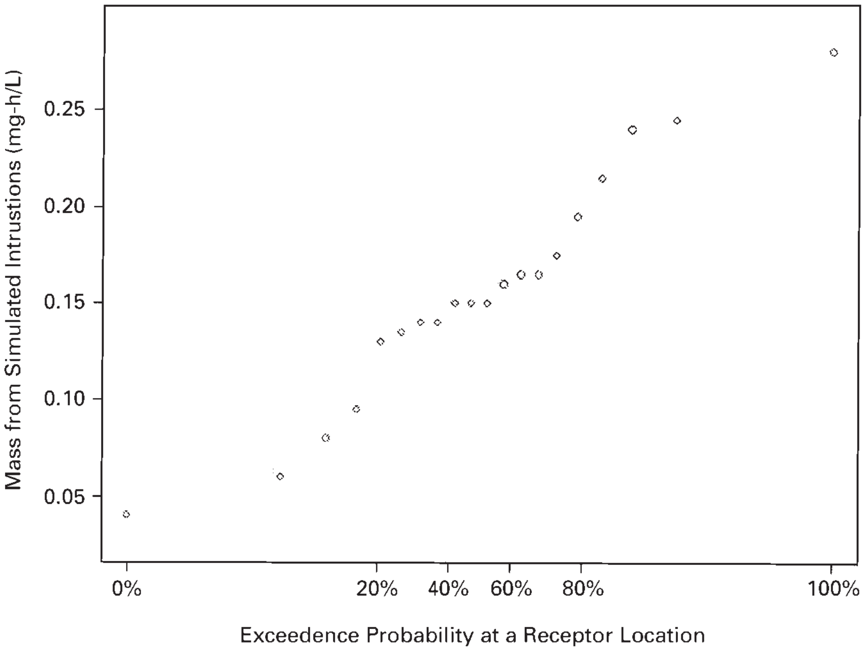


FIGURE 18.9 Results from random intrusion modeling due to low pressure during main breaks.

18.10 DISCUSSION

The majority of the locations experiencing infrequent pressure problems are located on the system periphery where the network is more branched than in the central part of the study area; this finding reiterates the value of a looped network (Fig. 18.6). The study area's resistance to pressure fluctuation eliminates a key susceptibility condition and makes contamination events practically impossible.

In the case study, susceptible conditions exist at some locations in the system, but few locations experience all conditions at the same time. Additionally, the sensitive populations considered tend to be located in the central region of the area, where the pressure buffering due to the looped network is most effective. This spatial serendipity (unique to the study area) means those locations with the highest potential for intrusion events have little effect (considering hydraulic connectivity as a metric) on the sensitive populations noted here.

It is clear in this case that the hydraulic modeling used to generate pressure information and subsequent influence analysis is closely connected to other spatial information in the GIS and vice versa. In the case study, historical maintenance data for main breaks was used to drive the hydraulic model. Those tasked with distribution system design and operations are encouraged to think of their hydraulic models and their GIS data as single, not separate, or even parallel systems. Traditional skeleton models with a handful of demand nodes connected by straight links (i.e., the links are not spatially correct except for length) will not readily support the type of analysis shown in the case study or other analyses beyond traditional planning applications. Proper spatial representation of the modeled links and nodes (which served in this case as main attribute placeholders) allows for wider use of distribution system models in conjunction with other spatial data.

18.11 CONCLUSION

Distribution system contaminant intrusions are attributed to three underlying susceptibility conditions: adverse pressure gradient intrusion pathway, and contaminant source. A hydraulic model and spatial information about network risk factors were integrated within a GIS to identify areas of the distribution system susceptible to intrusion events. Once identified, these areas were prioritized by considering how they influence (via hydraulic connectivity) local sensitive populations. With this information, distribution system managers can apply remedies to overcome or manage one or more of the conditions that result in system susceptibility.

The distribution system is the final barrier to waterborne illness. However, certain initiation events (i.e., pipe breaks coupled with large demands), which may be extreme but not improbable, can result in the occurrence of multiple susceptible conditions at a single location. The presence of the susceptible conditions does not mean intrusion is imminent; it simply means the necessary conditions exist for an intrusion to occur. Identification of these critical conditions has been suggested as part of a comprehensive operating plan or distribution system sanitary survey (AWWARF, 2000). The results of the framework may support utility capital improvement plans, infrastructure maintenance, improved cross-connection control, accreditation procedures, and provide a basis for regulatory sampling designs. This kind of approach may actually improve water quality and protect public health more than extensive sampling and monitoring efforts (Allen et al., 2000).

18.12 REFERENCES

Allen, M. J., Clancy, J. L., and Rice, E. W., 2000. The Plain, Hard Truth About Pathogen Monitoring. *J. AWWA* 92(9):64.

American Water Works Association, 1997a. AWWA Standard for Double Check Valve Backflow Prevention Assembly. AWWA C510. Denver, Colo.

American Water Works Association, 1997b. AWWA Standard for Reduced-Pressure Principle Backflow Prevention Assembly. AWWA C511. Denver, Colo.

American Water Works Association, 1996. AWWA Standard for Flexible Membrane Lining and Floating Cover Materials for Potable Water Storage. AWWA D130. Denver, Colo.

American Water Works Association, 1992a. AWWA Standard for Disinfecting Water Mains. AWWA C651. Denver, Colo.

American Water Works Association, 1992b. AWWA Standard for Disinfecting Water Storage Facilities. AWWA C652. Denver, Colo.

American Water Works Association. 1990. Manual M14, *Recommended Practice for Backflow Prevention and Cross-Connection Control.* Denver, Colo.

AWWARF (AWWA Research Foundation), 2000. Pathogen Intrusion into the Distribution System. AWWA, Denver, Colo.

AWWARF (AWWA Research Foundation), 1999. *Maintaining Water Quality in Finished Water Storage Facilities.* AWWA, Denver, Colo.

AWWARF (AWWA Research Foundation), 1986. *Water Main Evaluation for Rehabilitation/Replacement.* AWWA, Denver, Colo.

Barcellos, C., 2000. Health Risk Analysis of the Rio de Janeiro Water Supply Using Geographic Information Systems. Interdisciplinary Perspectives on Drinking Water Risk Assessment and Management [*Proc. Santiago (Chile) Symp.,* September 1998], *IAHS,* 260:95

Clark, K., 1999. *Getting Started with Geographic Information Systems.* Prentice-Hall, Upper Saddle River, N.J.

Clark, R., 2000. Water Quality Modeling—Case Studies, chap. 13 in L. W. Mays, (ed.), *Water Distribution Systems Handbook,* McGraw-Hill, New York.

Craun, G. F., and Calderon R. L. 2001.Waterborne Disease Outbreaks Caused by Distribution System Deficiencies in the United States. *J. AWWA* 93(9):64.

Fenner, R. A., and Sweeting, L., 1999. A Decision Support Model for the Rehabilitation of Non-Critical Sewers. *Water Science and Technology* 39(9):193.

Foundation for Cross-Connection Control and Hydraulic Research (FCCHR), 1993. *The Manual of Cross-Connection Control,* (9th ed.). University of Southern California, Los Angeles.

Geldreich, E. E., 1996. Microbial Quality of Water Supply in Distribution Systems. CRC Lewis Publishers, Boca Raton, Fla.

Great Lakes Upper Mississippi River Board (GLUMRB), 1997. Recommended Standards for Wastewater Facilities.

Haas, C. N., 1999. Benefits of Using a Disinfectant Residual. *J. AWWA* 91(1):65.

Habibian, A., 1992. Developing and Utilizing Data Bases for Water Main Rehabilitation. *J. AWWA,* 83(7):75.

Hadzilacos, T., Kalles, D., Preston, N., Melbourne, P., Camarinopoulos L., Eimermacher, M., Kallindromitis, V., Frondistou-Yannas, S., and Saegrov, S., 2000. UtilNets: A Water Mains Rehabilitation Decision-Support System. *Computers, Environment and Urban Systems,* 24:215.

Kleiner, Y., and Rajani, B., 1999. Using Limited Data to Assess Future Needs. *J. AWWA* 91(7):47.

LeChevallier, M. W., 1999. The Case for Maintaining a Disinfectant Residual. *J. AWWA,* 91(1):86.

Lindley, T. R., and Buchberger, S. G., 2002. Framework to Assess Intrusion Susceptibility in Water Distribution Systems. *J. AWWA* 94(6).

Maslia, M. L., Sautner, J. B., Aral, M. M., Reyes, J. J., Abraham, J. E., and Williams, R. C., 2000. A Protocol for Using Water-Distribution System Modeling as a Tool to Assist Epidemiological Investigations. *J. Water Resources Planning and Management,* 126(4):180.

McMath, S. M., and Casey, R. J., 2000. The Use of Network Modelling Techniques to Resolve Water Quality Issues in Distribution Systems. *Proc. AWWA WQTC,* November 2000, Salt Lake City, Utah.

Rossman, L. A., 2000. *EPANET 2 Users Manual.* USEPA, Office of Research and Development. EPA 600-R-00-057.

Trussel, R. R., 1999. Safeguarding Distribution System Integrity. *J. AWWA* 91(1):46.

USEPA, 2000. Report to Congress EPA Studies on Sensitive Subpopulations and Drinking Water Contaminants. Office of Water, EPA 815-R-00-015.

USEPA, 1989. *Cross-Connection Control Manual.* Office of Water, EPA 570/9-89-007.

World Health Organization (WHO), 2000. Global Water Supply and Sanitation 2000 Report. World Health Organization and United Nations Children Fund. Printed in the United States of America.

P · A · R · T · 5

SECURITY OF URBAN WATER SUPPLY SYSTEMS

CHAPTER 19

PROTECTING THE NATION'S CRITICAL INFRASTRUCTURE: THE VULNERABILITY OF U.S. WATER SUPPLY SYSTEMS*

Robert M. Clark
U.S. Environmental Protection Agency
Cincinnati, Ohio

Rolf A. Deininger
School of Public Health
University of Michigan
Ann Arbor, Michigan

19.1 INTRODUCTION

Terrorism in the United States used to be a rare phenomenon. It was generally not regarded as a serious threat because of the nation's military strength, relative geographic isolation, and secure borders. However, recent attacks against targets within the United States by domestic and foreign terrorists forced many government planners to consider the possibility that the nation's critical infrastructure may, in fact, be vulnerable to terrorist attacks. In response to this concern, the President's Commission on Critical Infrastructure was formed to evaluate the vulnerability of the following infrastructure categories to internal and external terrorism: Information and

*This chapter was originally published with the same title in Journal of Contingencies and Crisis Management, Volume 8, No. 2, pp. 73–80, 2000. (Permission to republish granted by Blackwell Publishing, Ltd., Oxford, U.K.)

Communication, Physical Distribution, Banking and Finance, Energy and Vital Human Services (President's Commission on Critical Infrastructure Protection, 1996). The rapid proliferation of telecommunication and computer systems, which connect infrastructures to one another in a complex network, compounds this vulnerability.

Vital Human Services include community water supply systems on local and state levels. In terms of public administration, water supply systems are generally governmental in nature. However, each supply system tends to be highly localized. Failures in one community may have little direct impact on other communities, although the problems and vulnerabilities may be similar. Water supply systems are vulnerable to the full range of terrorist threats including physical attack, cyber, and biological terrorism.

The potential of bioterrorism as a threat to public safety is becoming increasingly apparent (Henderson, 1998). For example, two epidemics of smallpox occurred in Europe in the 1970s. Each outbreak resulted from one infected individual. An aerosolized anthrax discharge from a Russian bioweapons facility in 1979 resulted in 77 cases of anthrax and 66 deaths. It is estimated that the release probably lasted no more than a few minutes and the weight of the aerosols released may have been as little as a few milligrams.

Table 19.1 summarizes potential biological weapons based on a report prepared by the U.S. Army Combined Arms Support Command (Burrows and Renner, 1998). Column 1 lists the agent and column 2 describes the type of organism, while the potential for "weaponization" is identified in column 3. Seven out of 27 agents are listed as being weaponized and 14 others are listed as either possible or probable weapons. A number of these organisms are listed as definite or probable threats in water.

This chapter will review the characteristics of water supply systems and their vulnerability to terrorist attacks, including the general potential for contamination of water systems. Possible physical and chemical countermeasures will be examined. An example of a naturally occurring variant of cholera that is resistant to chlorine will be presented as a possible organism that could be used in bioterrorist attacks. A case study that illustrates some of the difficulties inherent in tracking a contamination event in a drinking water system will be examined.

19.2 CHARACTERISTICS OF WATER SUPPLY SYSTEMS IN THE UNITED STATES[1]

Some of the common elements associated with water supply systems in the United States are as follows:

- A water source may be a surface impoundment such as a lake, reservoir, river, or groundwater from an aquifer.
- Surface supplies generally have conventional treatment facilities including filtration, which removes particulates and potentially pathogenic microorganisms, followed by disinfection.

TABLE 19.1 Potential Threat of Biological Weapons Agents: Summary*

Agent	Type	Weaponized	Water threat	Stable in water	Chlorine† tolerance
Anthrax	Bacteria	Yes	Yes	2 years (spores)	Spores resistant
Brucellosis	Bacteria	Yes	Probable	20–72 days	Unknown
C. perfringens	Bacteria	Probable	Probable	Common in sewage	Resistant
Tularemia	Bacteria	Yes	Yes	Up to 90 days	Inactivated, 1 ppm, 5min
Glanders	Bacteria	Probable	Unlikely	Up to 30 days	Unknown
Meliodosis	Bacteria	Possible	Unlikely	Unknown	Unknown
Shigellosis	Bacteria	Unknown	Yes	2–3 days	Inactivated, 0.05 ppm 10 min
Cholera	Bacteria	Unknown	Yes	"Survives well"	"Easily killed"
Salmonella	Bacteria	Unknown	Yes	8 days, freshwater	Inactivated
Plague	Bacteria	Probable	Yes	16 days	Unknown
Q fever	Rickettsia	Yes	Possible	Unknown	Unknown
Typhus	Rickettsia‡	Probable	Unlikely	Unknown	Unknown
Psittacosis	Rickettsia-like	Possible	Possible	18–24 h, seawater	Unknown
Encephalomyelitis	Virus	Probable	Unlikely	Unknown	Unknown
Hemorrhagic fever	Virus	Probable	Unlikely	Unknown	Unknown
Variola	Virus	Possible	Possible	Unknown	Unknown
Hepatitis A	Virus	Unknown	Yes	Unknown	Inactivated, 0.4 ppm, 30 min
Cryptosporidiosis	Protozoan§	Unknown	Yes	Stable days or more	Oocysts resistant
Botulinum toxins	Biotoxin¶	Yes	Yes	Stable	Inactivated, 6 ppm, 20 min
T-2 mycotoxin	Biotoxin	Probable	Yes	Stable	Resistant
Aflatoxin	Biotoxin	Yes	Yes	Probably stable	Probably tolerant
Ricin	Biotoxin	Yes	Yes	Unknown	Resistant at 10 ppm
Staph enterotoxins	Biotoxin	Probable	Yes	Probably stable	Unknown
Microcystins	Biotoxin	Possible	Yes	Probably stable	Resistant at 100 ppm
Anatoxin A	Biotoxin	Unknown	Probable	Inactivated in days	Unknown
Tetrodotoxin	Biotoxin	Possible	Yes	Unknown	Inactivated, 50 ppm
Saxitoxin	Biotoxin	Possible	Yes	Stable	Resistant at 10 ppm

*Burrows and Renner (1998).
†Ambient temperature, <1 part per million (ppm) free available chlorine, 30 min., or as indicated.
‡Parasites that are pathogens for humans and animals.
§Consisting of one cell or of a colony of like or similar cells.
¶Toxic to humans.

- Transmission systems which include tunnels, reservoirs and/or pumping facilities, and storage facilities.
- A distribution system carrying potable water through a network of water mains and subsidiary pipes to consumers.

Community water supplies are designed to deliver water under pressure and generally supply most of the water for fire-fighting purposes. Loss of water or a substantial loss of pressure could disable fire-fighting capability, interrupt service, and disrupt public confidence. This loss might result from a number of different causes. For example, sabotaging pumps that maintain flow and pressure or disabling electric power sources could cause long-term disruption. Many of the major pumps and power sources in water systems have custom-designed equipment, and in case of a physical attack it could take months or longer to replace them (President's Commission on Critical Infrastructure Protection, 1996).

Contamination of public water supplies generally occurs accidentally. Recently *Giardia* and *Cryptosporidium* have been implicated in water supply contamination events (Fox and Lytle, 1996). Both of these parasites are resistant to chlorination, and there is concern that microorganisms with similar resistance could be introduced deliberately into water systems and cause illness before remedial action could be taken.

Many common organisms are either removed by effective filtration or inactivated by common oxidants such as chlorine. However, there are strains of bacteria, viruses, and other pathogens that can survive water treatment to cause sickness and death among water consumers. The President's Commission concluded that there is a credible threat to the nation's water supply system from certain known biological agents (President's Commission on Critical Infrastructure Protection, 1996). In addition, newly discovered or emerging pathogens may pose a threat to water supply systems. One such pathogen isolated during a United States Environmental Protection Agency (U.S. EPA) study in Peru will be discussed.

Several chemical agents have also been identified that might constitute a credible threat against water supply systems (Berger and Stevenson, 1955). Although much is known about chemical and biological agents dispersed in air, almost nothing is known about these agents in potable water. Contamination has occurred from surface runoff, leaching from toxic waste dumps, or toxic materials leaching from underground pipes and tanks.

The amount of material needed to deliberately contaminate a water source (such as a reservoir or aquifer) is large and generally exceeds what an individual or small group of terrorists could easily acquire, produce, or transport. However, contaminants introduced into a distribution system would be less susceptible to dilution and would reside in the system for shorter times, thus diminishing the effects of disinfectants and chemical decomposition and oxidation.

Many urban water systems rely on an aging infrastructure. Temperature variations, large swings in water pressure, vibration from traffic or industrial processes, and accidents often result in broken water mains. Distribution systems in many major cities operate with very little margin for error. Planning for main breaks is usually based on historical experience. However, breaks could be induced by a systemwide hammer effect, which could be caused by opening or closing major control valves too rapidly. This could result in simultaneous main breaks that might exceed the community's capability to repair these in a timely manner, causing widespread outages. Recognizing this vulnerability, water systems have been incorporating valves that cannot be opened or closed rapidly. However, many urban systems still have valves that could cause severe water hammer effects.

Interrupting the water flow to agricultural and industrial users could have large economic consequences. For example, the California aqueduct, which carries water from northern parts of the state to the Los Angeles–San Diego area also serves to irrigate the agricultural areas in midstate. Pumping stations are used to maintain the flow of water. Loss of irrigation water for a growing season, even in years of normal rainfall, would likely result in billions of dollars of loss to California and significant losses to U.S. agricultural exports.

Another problem associated with many community water systems is the potential for release of chlorine to the air. Most water systems use gaseous chlorine as a disinfectant, which is normally delivered and stored in railway tank cars. Generally, there is only minimal protection against access to these cars. Accidental release of chlorine gas could cause injury to nearby populations.

19.3 POTENTIAL FOR CONTAMINATION OF WATER SUPPLIES

Textbooks on military chemistry and chemical and biological warfare list the criteria for sabotage by poison and the circumstances that might surround such an event (Franke, 1977). Botulinus toxin is likely the most deadly naturally occurring poison. Table 19.2 lists some of the natural poisons and their estimated lethality (Deininger and Meier, 2000); the "LD50" level means the dosage at which 50 percent of the exposed population of mice would be dead.

If one includes human-made substances, as well as inorganic substances, one can make a relative ranking of the poisons based on their solubility in water and their toxicity. The underlying equation is

$$R = \frac{\text{solubility in mg/L}}{\text{lethal dose in mg/human} \cdot 1000}$$

The fator of 1000 in the above equation is to ensure a conservative estimate and to account for a variability in sensitivity to the toxin. Results from this calculation are shown in Table 19.3. The most toxic contaminant in Table 19.3 is the Botulinus toxin.

TABLE 19.2 Comparative Toxicity of Natural Poisons

Substance	LD50 in μg/kg (mouse)
Botulinus	0.00003
Tetanus	0.001
Tetrodotoxin	9
Saxitoxin	9
Batrachotoxin	2.7
Palytoxin	0.15

The World Health Organization has described a hypothetical contamination scenario for a city of 50,000 people with a daily water use of 400 L per person per day (World Health Organization, 1970). Each person drinks 0.5 L of water per day, while a lethal dose is 1 μg, according to the scenario. The total dose required to contaminate 20 million L of water is 40 g, assuming a homogeneous solution. Allowing for a factor of 6 to compensate for unequal distribution and dilution, the total amount of poison required would be 240 g. The toxin would have to be delivered over a period of time, most likely from midnight to 6 A.M. The affected population would show effects about 8 h later, or around the middle of the afternoon. There are many ifs associated with this scenario, such as travel time and water use in the mains. Consumers closer to the injection point would show symptoms earlier than those located farther away, as the travel time of the water in the system could be days.

TABLE 19.3 Relative Toxicity of Poisons in Water

Compound	*R*
Botulinus toxin A	10,000
VX	300
O-ethyl-S-(N_3N-dimethylaminoethyl)	250
Methyl thiophosphonate	
Sarin	100
Nicotine	20
Colchicine	12
Cyanide	9
Amiton	5
Fluoroethanol, sodium fluoroacetate	1
Selenite	1
Arsenite, arsenate	1

There are few reported attempts to poison a water supply in the United States. One occurred in New York City (*New York Times,* 1986), as low levels of plutonium were found in the drinking water [(in the order of 20 femtocuries (fCi)]. The usual background is below 1 fCi. However, a person would have to drink several million liters of water to acquire a lethal dose estimated at about 100 μCi. A femtocurie is nine orders of magnitude smaller than a microcurie.

Another case was the contamination of salad bars in Dalles, Oregon, by the Rajneeshee religious cult, using vials of *Salmonella typhimurium* (McDade and Franz, 1998). *S. typhimurium* is a highly toxic bacteria frequently carried by birds. The cult also contained a city water supply tank using *Salmonella.* A community outbreak of salmonellosis resulted in which at least 751 cases were documented in a county that typically reports fewer than five cases per year. The cult apparently cultured the organisms in its own laboratories.

19.4 COUNTERMEASURES

There are several steps that a water utility can take to protect the system against deliberate contamination. Some of these measures are listed below.

19.4.1 Physical Countermeasures

The most effective manner of proactive planning is to construct a highly protected system, which in the event of partial destruction would still be operational at some minimal level. Access to a free water surface such as exists in a water reservoir should be eliminated. For example, the ventilation devices in an underground reservoir must be constructed in such a way as to prevent contamination of the reservoir. The intakes, pumping stations, treatment plants, and reservoirs should be fenced to secure them against casual vandalism. Beyond that, intrusion alarms should be installed to notify the operator that an individual has entered a restricted area. An immediate response might be to shut down a part of the pumping system until the appropriate authorities determine that there is no threat to the system.

The water supply system in Zurich, Switzerland, is an example of the type of protective design that enhances physical security (Schalekamp, 1984). In 1981 it was decided to enlarge the capacity of the Zurich water system by expanding the groundwater plant at Hardof, which is one of several facilities that provide water to Zurich and the surrounding area. Three additional wells were constructed at the Hardof complex as a complement to the existing wells. These vertical wells were up to 22 m in depth and 4 m in diameter. Each well was equipped with a submerged pump with a capacity of 220 L/s. The groundwater required treatment with aeration and subsequent chlorination and deacidification. The equipment for operating and maintaining the wells is located in the tops of the wells which are covered by a layer of earth that is landscaped with shrubs, providing the wells with protection from intrusion and from surface sources of contamination. The entire site was designed as a protective zone. Security measures such as video cameras were installed to protect against sabotage at all facilities. All of the construction complied with Switzerland's civil defense standards, including the installation of diesel generators for emergency power. Provisions were made for continuous monitoring of the security system, temperature of the water, and raw and treated water quality; measurement of oxidants, chlorine dioxide, pH, redox potential, conductivity ultraviolet (UV) extinction, and oxygen; and finally, monitoring of water quality through fish test facilities.

19.4.2 Chemical Countermeasures

There are five forms of Botulinum Toxin (labeled A, B, C, D, and E) that are of importance. Virtually all of them are deactivated after a few seconds exposure to chlorine (Brazis, 1959). A lethal dose for a human is about 100 mL of such a solution.

The second line of defense would be to maintain a chlorine residual of about 0.5 mg/L free chlorine in the distribution system, and to increase this dosage in times of perceived or real danger. This might increase the total trihalomethane levels. In the United States, the controlled level is 80 μg/L. However, this total trihalomethane limit is based on a lifetime exposure. Most bacteria are destroyed by common disinfectants such as free chlorine or chloramines. It is, of course, possible to deactivate the chlorine by adding sodium thiosulfate, but this would require large amounts and a reliable injection system.

Monitoring of chlorine residuals is not a universal practice, but it can be done at minimal cost. On July 10, 1986, the water supply to the presidential rooms of the White House was cut off after a monitor indicated a lack of chlorine. President Reagan got his morning coffee anyhow, using bottled water. The supply to the West Wing was not shut off, but the staff was warned not to drink the water (*New York Times,* 1986). Strategically placed monitors in a distribution system provide one solution for the protection of a water supply system.

19.5 PERSISTENT MICROBIL CONTAMINANTS: THE RUGOSE VARIANT

Some systems are very vulnerable to microbiological contamination, but chlorine residuals are assumed to be protective. There are, however, natural microorganisms that are highly resistant to chlorine, as illustrated by the entries in Table 19.1. In March, 1991, the U.S. EPA was requested by the Peruvian Ministry of Health to send a team of water supply experts to Peru. The team provided technical support to the field epidemiology staff of the U.S. Centers for Disease Control, which had been in Peru for some time studying a cholera outbreak (Craun et al., 1991).

During the course of the investigation, the EPA team began a series of disinfection experiments on several cholera strains (Rice et al., 1993). One strain was an isolate recovered from a patient during the Peruvian epidemic. This isolate exhibited both a "rugose" variant as well as a "smooth" type of culture. The smooth culture, which is normally associated with cholera, exhibits a round, smooth colonial morphology. The rugose variant represents an extremely rough form of *Vibrio cholerae* and receives its name from the corrugated appearance of the colonies on nonselective agar media. The EPA investigators found that the rugose variant was more resistant to disinfection than the smooth culture.

19.5.1 Chlorine Inactivation

Chlorine inactivation experiments were conducted in triplicate for each condition as previously described. Both the Gulf Coast and Peruvian smooth strains were readily inactivated by free chlorine. The smooth cultures exhibited classical first-order inactivation. At 1.0 mg/L free chlorine the smooth strains were inactivated by greater than four orders of magnitude in less than 20 s. Rates of inactivation did not differ significantly between the two smooth strains.

Disinfection of the rugose culture displayed a deviation from first-order kinetics. After 80 s of exposure to free chlorine in samples containing both smooth and rugose organisms, the smooth strain was inactivated by approximately three orders of magnitude. The rate of inactivation of the rugose variant portion of the sample was significantly lower than that of the smooth culture and showed only limited inactivation. These results suggest that there are naturally occurring pathogens that are resistant to chlorination. Such organisms can be cultured and grown in a laboratory with ease, and the normal chlorination practices can thus be defeated.

19.6 CONTAMINANT PROPAGATION IN DRINKING WATER SYSTEMS

The potential for developing or discovering pathogens that are resistant to chlorine or other disinfectants is only one issue in assessing the vulnerability of water systems to sabotage. Another problem in dealing with deliberate contamination of a water system is the complex water movement within a system. Flow patterns in distribution systems can be highly variable, and these patterns can have a significant impact on the way contaminants are dispersed in a network. It is difficult to predict where a given parcel of water will be at a given time, because there is a large random element to the movement of water in a distribution system. Moreover, water distribution systems are far-flung and made up of different hydraulic elements. The pattern of contaminant movement in a system was studied for the North Penn Water Authority (NPWA) in Lansdale, Pennsylvania (Clark, 1990).

At the time of the case study (1985), the NPWA served 14,500 customers in 10 municipalities, supplying an average of 5 million gallons of water per day (18.9 million liters). A surface water source provided a daily supply of 1 million gallons of treated water (3.78 million liters), purchased from the Keystone Water Company. Groundwater from 40 wells operated by the North Penn Water Authority supplied 4 million gallons of water per day (15.1 million liters). The NPWA distribution system was modeled in a network consisting of 528 links and 456 nodes.[2]

The model assumed the injection of 100 units of a contaminant at the Keystone tie-in, in order to study its movement through the network. Initially, the contaminant was localized in the southeastern portion of the system near the point of entry. After 9 h, the contaminant had spread to the central portion of the water distribution system. At 16 h after the injection, the contaminant had begun to disperse. After 24 to 26 h, the contaminant was well dispersed and was beginning to disappear. After 34 h, the contaminant still persisted at very low levels, but was virtually gone from the system (Clark, 1990).

Another way of examining this effect is the calculation of contaminant concentrations at given points in the system. A time trace of contaminant movement was calculated at three points. The analysis demonstrated that it is difficult to track a contaminant in a water distribution system. An outbreak of a waterborne disease in a small community in western Missouri illustrates this point.

19.7 ILLNESS PROPAGATION IN A WATER DISTRIBUTION SYSTEM

Gideon is a small town located in southwestern Missouri with a population of 1104. On November 29, 1993, two high school students from Gideon were hospitalized with culture confirmed salmonellosis, which came to the attention of the Missouri Department of Health (Clark et al., 1996). Within 2 days five additional patients from Gideon were hospitalized with salmonellosis (one student, one child from a day care center, two residents, and one visitor of a nursing home). The State Public Health Laboratories identified the isolates as dulcitol negative *Salmonella,* and the Centers for Disease Control laboratories identified the organism as serovar *typhimurium.* Interviews conducted by the Missouri Department of Health suggested that there was not a similar food source shared by a majority of the patients. However, all of the ill persons had consumed municipal water.

The Missouri Department of Natural Resources was informed that the Department of Health suspected a water supply link as the cause for the outbreak. Water samples collected on December 16th were positive for FC (fecal coliform bacteria). On December 18th, the city of Gideon issued a boil water order, as required by the Missouri Department of Natural Resources. Signs were posted at the city hall and in the grocery store, while two radio stations in the area announced the boil water order (Angulo et al., 1997).

By January 8, 1994, the Missouri Department of Health had identified 31 cases with laboratory confirmed salmonellosis, associated with the Gideon outbreak. The State Public Health Laboratories identified 21 of these isolates as dulcitol negative *Salmonella* serovar *typhimurium.* Fifteen of the 31 culture confirmed patients were hospitalized, including two patients hospitalized for other causes, who developed diarrhea while in the hospital. The patients were admitted to 10 different hospitals. Two of the patients had positive blood cultures. Seven residents of the nursing home, who exhibited diarrhea, subsequently died. All of the culture confirmed patients had been exposed to Gideon municipal water. It is estimated that nearly 500 of the 1104 residents became ill as a result of the contamination event.

The investigation clearly implicated the consumption of Gideon municipal water as the source of the outbreak. Speculation focused on a sequential flushing program conducted on November 10, 1993, involving all 50 hydrants in the system. The flushing program was conducted in response to taste and odor complaints about the municipal water. A thermal inversion could have caused stagnant or contaminated water in the upper part of a water tank to mix with the deeper water. Thus the dirty water could have become discharged into the system, resulting in taste and odor complaints (Fennel, James, and Morris, 1974). The water authority initiated a citywide flushing program in an attempt to get rid of the problem. Turbulence from the flushing program could have stirred up sediments in the water tank, subsequently transported into the distribution system. It is likely that the bulk water and/or the sediments were contaminated with *Salmonella* serovar *typhimurium.*

During a field visit of the U.S. EPA to Gideon, a large number of pigeons were observed roosting on the roof of the municipal water tank with a capacity of 100,000 gal (378,000 L). The municipal system also has a smaller water tank, having a volume of 50,000 gal (189,000 L). A third tank, privately owned, is located on the Cotton Compress property. It also has a volume of 100,000 gal. Shortly after the outbreak, a tank inspector found holes at the top of the Cotton Compress tank, rust on the tank, and rust, sediment, and bird feathers floating in the water. The water in the tank looked black and was so turbid he could not see the bottom, according to the inspector. Another inspection confirmed the disrepair of the Cotton Compress tank and also found the large municipal tank in such a state of disrepair that bird droppings could, in the opinion of the inspector, have entered the stored water. Bird feathers were in the vicinity or in the tank openings of both the Cotton Compress and the large municipal tank. The private tank was drained accidentally during an inspection after the outbreak of the disease, so it was impossible to sample its water. However, its sediment contained *Salmonella* serovar *typhimurium* dulcitol negative organisms, also found in a sample from a hydrant and in culture confirmed patients. The *Salmonella* found in a hydrant matched the serovar of the patient isolate, by comparing DNA fragments in the laboratory of the Centers for Disease Control. The isolate from the tank sediment, however, did not provide an exact match with the other two isolates. No *Salmonella* isolates were found elsewhere in the water distribution system.

An evaluation was conducted to study the effects of system design, operations, water use, and hydraulic characteristics on the possible propagation of contaminants within the system. A water distribution and water quality model, EPANET, was used for this evaluation (Rossman, Clark, and Grayman, 1996). It was concluded that bird droppings in the large municipal tank were the most likely source of contamination. The analysis subsequently concentrated on water from the large municipal tank in conjunction with the flushing program. Data from the simulation study, the microbiological surveillance, and the outbreak itself were investigated. The water movement patterns showed that the majority of the special samples, which were fecal coliform positive, occurred at points that lie within the zone of influence of the small and large tanks. These areas were predominately served by tank water during the flushing program and during normal operation. Hence, the tanks were the likely source of the fecal contamination as fecal coliform samples were taken prior to chlorination.

An overlay was created of the areas served by the small and large water tanks during the first 6 h of the flushing period and the earliest recorded cases of salmonellosis. It was concluded that the water, which reached the residence and the Gideon School during the first 6 h of the flushing period, was almost totally derived from the large tanks. The illness propagated rapidly and nearly 500 of the 1104 residents became ill. However, by the time it was recognized as a waterborne outbreak, none of the routine surveillance samples yielded positive samples. This illustrates the difficulty of tracking and identifying the course of contamination in municipal water distribution systems. The outbreak of salmonellosis started in December 1993 and was over by the middle of January 1994.

19.8 CONCLUSIONS

There is growing concern about the vulnerability of the U.S. infrastructure to deliberate sabotage, including the susceptibility of water distribution systems to terrorist attacks. There are a number of factors that make water supply systems particularly vulnerable: they are spatially diverse, susceptible to intrusion, and contain many components. Water tanks and storage reservoirs are particularly vulnerable to deliberate sabotage. All water distribution systems use pumps to transfer water and maintain pressure. Besides normal water supply functions for domestic and industrial consumption, most municipal water systems must provide sufficient flow and pressure for fire-fighting purposes. All of these functions could be disrupted with well-placed and well-timed attacks.

Water distribution systems are particularly vulnerable to deliberate microbiological contamination, although residual chlorine provides some protection. A contaminant would move rapidly through the system and could cause a widespread outbreak of disease among the inhabitants of the area served by the water distribution system. Moreover, it would be difficult to track the contaminant, as shown in the above example from Gideon, Missouri.

It may be possible to poison a water supply system, but it would take a large quantity of most contaminants to have a major effect on the system's consumers. Most toxins can be neutralized by keeping a certain concentration of disinfectant in the water such as chlorine. There are pathogens, however, which exhibit a high degree of resistance to chlorine. It is essential for water system managers to maintain or increase the level of physical security to prevent access by unwanted people. The Hardof water utility plant in Zurich provides an example of the incorporation of physical security measures as part of its upgrading program. Water utility managers in the United States and throughout the world might consider similar action when such opportunities arise.

An important extension of the security concept would be the planning and construction of separate water lines that are fed from a protected water supply source, which would only be activated during an emergency. Many of the older cities in the United States have separate water lines that have been installed for fire protection in heavily developed downtown areas. These water lines might be upgraded for possible use to supply the population with safe water in emergency conditions. Such proactive planning for water security, including the continuous maintenance and monitoring of chlorine residual in the water, will help to ensure the safety of most water supply systems. Nevertheless, it is of vital importance that system planners and managers are constantly on the alert to prohibit deliberate sabotage of municipal water supply systems.

19.9 NOTES

1. The authors would like to acknowledge the assistance of Ms. Jean Lillie and Mr. Steven Waltrip of the Water Supply and Water Resources Division in preparing this manuscript.
2. A link is a water pipe and a node is an intersection. Water use represented conditions during the period May–June 1984.

19.10 REFERENCES

Angulo, F. J., Tippen, S., Sharp, D. J., Payne, B. J., Collier, C., Hill, J. E., Barrett, T. J., Clark, R. M., Geldreich, E. E., Donnell, H. D., and Swerdlow, D. L. (1997), 'A Community Waterborne Outbreak of Salmonellosis and the Effectiveness of a Boil Water Order,' *American Journal of Public Health,* Volume 87, Number 4, pp. 580–584.

Berger, B. B., and Stevenson, A. H. (1955), "Feasibility of Biological Warfare Against Public Water Supplies," *Journal American Water Works Association,* Volume 47, Number 2, pp. 101–110.

Brazis, R. A. (1959), "Effectiveness of Halogens or Halogen Compounds in Detoxifying Botulinum Toxins," *Journal American Water Works Association,* Volume 51, Number 7, pp. 902–912.

Burrows, W. D., and Renner, S. E. (1998), *Biological Warfare Agents As Potable Water Threats,* U.S. Army Combined Arms Support Command, Fort Lee, VA.

Clark, R. M. (1990) "Predicting Exposure to Water Contaminants in Distribution Systems," in Craun, G. F. (Ed.), *Methods for the Investigation and Prevention of Waterborne Outbreaks,* U.S. Environmental Protection Agency, Office of Research and Development, Washington, D.C., EPA/600/1-90/005[a], pp. 83–125.

Clark, R. M., Geldreich, E. E., Fox, K. R., Rice, E. W., Johnson, C. W., Goodrich, J. A., Barnick, J. A., and Abdesaken, F. (1996), "Tracking a *Salmonella* Serovar *Typhimurium* Outbreak in Gideon, Missouri: Role of Contaminant Propagation Modeling," *Journal of Water Supply Research and Technology—Aqua,* Volume 45, Number 4, pp. 171–183.

Craun, G., Swerdlow, D., Tauxe, R., Clark, R., Fox, K., Geldreich, E., Reasoner, D., and Rice, E (1991), "Prevention of Waterborne Cholera in the United States," *Journal American Water Works Association,* Volume 83, Number 11, pp. 43–48.

Deininger, R. A., and Meier, P. G. (2000), "Sabotage of Public Water Supply Systems," in Deininger, R. A., Literathy, P., and Bartram, J. (Eds.), *Security of Public Water Supplies,* NATO Science Series, Number 2—Environment, Volume 66, Kluwer Academic Publishers, pp. 241–249.

Fennel, H., James, D. B., and Morris, J. (1974), "Pollution of a Storage Reservoir by Roosting Gulls," *Journal Society for Water Treatment and Examination,* Volume 23, Part 1, pp. 5–24.

Fox, K. R., and Lytle, D. A. (1996), "Milwaukee's Crypto Outbreak Investigation and Recommendations," *Journal American Water Works Association,* Volume 88, Number 9, pp. 87–94.

Franke, S. (1977), *Textbook of Military Chemistry: Volume 2, Sabotage and Phytopoisons,* Military Publishing House of the German Democratic Republic, Berlin.

Henderson, D. A. (1998), "Bioterrorism as a Public Health Threat," *Emerging Infectious Disease,* Volume 4, Number 3, pp. 488–492.

McDade, J. E., and Franz, D. (1998), "Bioterrorism as a Public Health Threat," *Emerging Infectious Disease,* Volume 4, Number 3, pp. 493–494.

New York Times (1986), "White House Water Cut off Temporarily," July 10, p. 16.

President's Commission on Critical Infrastructure Protection (1996), *Critical Foundations: Protecting America's Infrastructure,* Available at *www.pccip.gov.*

Rice, E. W., Johnson, C. H., Clark, R. M., Fox, K. R., Reasoner, D. J., Dunnigan, M. E., Panigrahi, P., Johnson, J. J., and Morris, J. G. (1993), "Vibrio cholerae 01 can assume a 'rugose' survival form that resists killing by chlorine, yet retains virulence," *International Journal of Environmental Health Research,* Volume 3, Number 2, pp. 89–98.

Rossman, L. A., Clark, R. M., and Grayman, W. M. (1996), "Modeling Chlorine Residuals in Drinking Water Distribution Systems," *Journal of Environmental Engineering,* Volume 120, Number 4, July/August, pp. 803–820.

Schalekamp, M. (1984), "The Hardof Groundwater Complex in Zurich," *Journal American Water Works Association,* Volume 76, Number 6, June, pp. 78–83.

World Health Organization (1970), *Health Aspects of Chemical and Biological Weapons,* Geneva.

CHAPTER 20
SECURITY ANALYSIS AND RESPONSE FOR WATER UTILITIES*

Nicholas L. Burns
Christine A. Cooper
David A. Dobbins
Jessica C. Edwards
Les K. Lampe
Black & Veatch
Kansas City, Missouri

20.1 OVERVIEW

The events of September 11, 2001, force the intensified consideration of the terrorist threat. Water utilities have, until now, considered the likelihood of an attack on their systems to be very low, and few measures have been implemented to mitigate such a threat. This chapter provides guidance on the assessment of human-induced security risks and preparation of response plans appropriate for the risks identified in the security assessment.

Two points cannot be overemphasized regarding the use of guidance in this chapter or other security checklists. First, involvement of outside professionals in completion of security assessments is usually invaluable because they are trained to look for weaknesses in a system and can provide an objective evaluation of threats and vulnerabilities. Second, a completed security assessment and response plan is useful only if the utility staff and supporting agencies maintain continued vigilance.

*This chapter is a slightly modified version of the American Water Works Association report. Security Analysis and Response for Water Utilities, 2001.

20.1.1 Hazard Assessment

The sources of human-induced security risks to water utilities can include vandals, past and present employees and other individuals, domestic extremist groups, or foreign-based terrorist organizations. The two main types of threats discussed in this chapter are physical and contaminant threats.

No specific set of hazards or threats is applicable to all utilities but instead depends on the unique characteristics of each utility. Although a checklist of potential biological and chemical threats is included in this section, and more detailed checklists may be made available from other sources, each utility must use such lists with caution, making sure that differences in source of supply, treatment processes, distribution system configuration, and control systems are fully considered.

20.1.2 Vulnerability Assessment

The *vulnerability assessment* defines priorities for response activities. All elements necessary for a utility to produce safe and sufficient water must be evaluated for vulnerability to disruption. These elements of the utility supply system generally fall into the categories of

- Raw water supply facilities
- Treatment facilities
- Distribution facilities
- Operation and control facilities and systems
- Staff and personnel systems
- Support systems, including chemicals, power, and communication

It is recommended that public trust be considered a critical asset of the utility. A separate aspect of the security assessment is evaluation of the potential for disruption of the quantity and quality of the water supply system at various locations of vulnerability within the system. Special emphasis is given to vulnerability to biological or chemical contamination of water and the agents associated with this contamination. In addition, the utility now must consider the consequences of *intentional* chemical or biological contamination of the water, injury to facility personnel, or damage to the system.

20.1.3 Mitigation

Vulnerable components of a water supply system can be rendered less susceptible to harm through mitigation actions—actions intended to eliminate or reduce the damaging effects of disasters. Mitigation actions cover a wide variety of activities and can be as complex as retrofitting a treatment plant or as simple as erecting fences or changing locks and passwords after an employee is terminated. Mitigation actions will depend on the hazard and vulnerability analyses. This section discusses mitigating vulnerability through refined information sharing practices and reduced access to the "target."

20.1.4 Development of a Response Plan

Depending on the characteristics of vulnerability for a given water utility, a response plan must be prepared to define measures that will be implemented to either minimize the likelihood of an undesirable event or mitigate its impact. Changes in security systems, monitoring practices, physical facilities, and operations are dependent on the nature of the particular threats. A critical component of all response plans is an emergency response plan that defines responsibilities and resources both within the utility and external to the utility. Those resources external to the utility may be local (police, fire departments, emergency response teams, medical and public health agencies and facilities, etc.), regional or state (regulatory agencies, health departments, emergency preparedness offices), or national [U.S. Environmental Protection Agency (USEPA), the Federal Bureau of Investigation, the Centers for Disease Control and Prevention (CDC), the Federal Emergency Management Agency (FEMA), etc.]. The systems established pursuant to the Presidential Decision Directive 63 on Critical Infrastructure Protection are given special coverage.

20.1.5 Crisis Communications

Part of the response plan must be avenues and timing of communication with the utility customers. These customers will fall into different categories depending on the amounts of water used and the extent to which their operations are dependent on a potable water supply. Examples of large water users with critical dependence on water supplies are hospitals, schools, and fire departments. Communication with the general public is also critical, particularly related to health and safety issues. A comprehensive communication program and responsibility for implementation of the program must be established prior to a disruption or contamination of the water supply.

20.1.6 Summary

The American Water Works Association (AWWA) and the authors of this chapter recognize that water utility security issues are receiving more attention and importance than they did prior to September 11, 2001. The reliability and safety of public water supplies are of highest concern to the citizens whose health and well-being are dependent on this water; the professionals whose vocation is to obtain, treat, distribute, and protect this water; and all units of government responsible for protecting and maintaining the public health. Because of this heightened concern and awareness, it is likely that water system security will be enhanced in coming years through the development and application of many new tools and practices, particularly in the areas of monitoring/early warning, treatment technologies, distribution system controls, and security systems. This means that utility managers must continue to review and upgrade their approaches and technologies to better serve the nation's water system security needs.

While security assessments, response plans, and security systems are critically important and valuable, the effectiveness of these tools depends on the skills and diligence of the staff applying the tools. The effectiveness of water utility staff in implementing a security response plan can be measured by the degree to which the culture of the staff has become more security-conscious. The mission of a water utility is to deliver water of unquestioned safety to its customers. This means that the security of the water system must be considered at least equally with any other events that disrupt utility operations.

20.2 HAZARD ASSESSMENT

According to the President's Commission on Critical Infrastructure Protection (1997), three attributes are crucial to water supply users:

1. There must be adequate quantities of water on demand.

2. It must be delivered at sufficient pressure.

3. It must be safe to use.

Actions that affect any of these three factors can be debilitating for the infrastructure. The first two attributes are directly influenced by physical damage. The third attribute, water quality, is susceptible to physical events as well as the introduction of microorganisms, toxins, chemicals, or radioactive materials.

Terrorism can strike all the system components individually or in various combinations. Damages associated with most types of natural hazards can also result from terrorist activity. These can range from major events that cause severe damage and disruption of system operations to minor incidents that may not impact normal activities.

A *hazard* is defined here as any individual, group, or event that could destroy the facility, halt or suspend operations, otherwise threaten the public health, harm employees, publicly embarrass a utility, require a utility to expend a great deal of time or money, or cause general panic. In preparing the hazard (or threat) assessment, security assessment specialists generally look at who or what constitutes a threat and how the identified individual, group, or event could attack a particular system. It is important to consider all potential threats, because assuming the worst may result in overlooking more likely attempts.

20.2.1 Who Poses a Threat?

The terrorist's goal is to achieve notoriety for his or her cause. This can be through massive loss of life as was the case in the plane hijackings of September 11, 2001. Beyond the obvious human tragedy, the net result of these actions includes significant economic impact on industry, the federal government, and state and local governments as they scramble to provide additional law enforcement and National Guard presence. The key to terrorism's success lies not in the act itself but in the ensuing lack of confidence or feeling of insecurity for the citizens at large.

The actual goal of the terrorist is usually not directly associated with the target selected. The usual desire is to instill a general sense of fear in citizens to disrupt their normal activities and way of life. Consumer uncertainty can significantly impact a national economy. Public perception of the safety of infrastructure, as well as the infrastructure itself, are both important assets to be protected. The following paragraphs describe who may pose a threat to the potable water system.

Vandal. A vandal often has a goal in mind, but not necessarily a target; thus, the crime may be of opportunity. Some typical examples of such a crime are graffiti or broken windows. However, a vandalism problem can become much more serious if the paint used to write graffiti on a reservoir is then dumped into the reservoir, posing a threat to the health of the consumer.

Individual. An individual is someone working independently. Although the motivations of individuals can vary widely, the target is clearly defined as the final water user or the facility itself. Individuals who purposely threaten infrastructure are often mentally ill and may target victims based on ethnicity, beliefs, or other characteristics.

In 1974, a Yugoslav immigrant, Muharem Kubergovic, was the first person to acquire and threaten to use chemical agents against the United States. When apprehended in his home, police found 100 lb (45 kg) of explosives and multiple chemicals used in the manufacturing of chemical weapons. Kubergovic bought the chemicals and equipment from ordinary supply houses (Smithson and Levy, 2000).

Insiders. With their detailed knowledge of the facility and water system, past or present employees and contractors pose some of the most serious threats. Motivations of this group may include revenge or the venting of anger manifested from a real or imagined problem.

Domestic Extremist Groups. Cults and extremist groups with a political agenda pose a threat to water systems. In the past, food has been contaminated with salmonella with the objective of affecting voter turnout. The Aum Shrinrikyo cult in Japan spent $30 million on a poisonous gas plant for use in terrorist activities. Better known for the sarin gas attack, the cult also developed anthrax and attempted to use it as a weapon against a civilian population in downtown Tokyo (Smithson and Levy, 2000).

State-Supported Terrorist Organizations. These organizations usually have a large number of followers and the greatest financial and technological resources. The use of some chemical and biological weapons of mass destruction are limited to these groups because of the significant resources required for their development. Known nations with the capabilities to produce weapons of mass destruction are North Korea, China, India, Pakistan, Iran, Iraq, Syria, Libya, and Russia (Department of Defense, 2001).

20.2.2 Types of Threats

This section describes two types of threats against a water utility: physical and contaminant threats.

Physical. Many observers believe that a physical event that destroys or disrupts a water system's components is a much more likely scenario than a contamination event. For instance, explosive materials are readily available and require a lower level of education compared to the development and deployment of contaminants. Potential types of physical attacks are listed here.

Aerial attack. Includes physical attacks on the treatment facility or the use of airplanes to drop contaminants into open reservoirs or sources (the latter example combines physical and contaminant attacks).

Cyber-terrorism. Attacks on the data acquisition system [supervisory control and data acquisition (SCADA)] are another threat. A terrorist could disguise the data, neutralize the chlorine, or add no disinfectant, thus compromising disinfection and allowing the addition of microbes usually not considered a threat, such as salmonella (when chlorine residual is present). Alternatively, attacks on the central control systems could create a large number of simultaneous main breaks by opening and closing major control valves too rapidly. Because many SCADA and control networks are not connected to the Internet, this threat is most likely to come from a disgruntled employee with access to the system.

Explosives. Could be used at any number of locations to compromise the pumping, storage, or transmission of water. Explosives, which can be developed or obtained, pose less of a risk to the attacker compared to biological and chemical weapons. Explosives also require a comparatively lower level of education. A bomb explosion within the distribution system will require immediate response and redirection of water to prevent contamination and draining of the system.

Fire. May be one of the easiest methods of sabotage because needed materials are readily available. It is effective because destruction of the computer control system, pumps, or motors or compromising of the structure pose a significant barrier to the operation and timely restart of a plant. Furthermore, once the city water supply is reduced, the ability to fight other fires is compromised, seriously impacting the safety of other critical infrastructure.

Personnel. Attacks on the plant staff could lead to multiple personnel injuries that would leave a plant without a skilled operational workforce. A hostile takeover could also allow for a cyber-terrorism attack.

Contaminants. The following section outlines biochemical toxins; microbial agents; industrial chemicals; nerve, blood, and blister agents; and radioactive materials that could potentially contaminate potable water systems. The chemistry of each toxin, chemical, and microbial agent is specific. Some are neutralized by chlorine, others are effectively removed through the drinking water treatment process, and all have different thresholds for the appearance of symptoms, infection, and lethality.

Biochemical toxins. A very small volume is required compared to other chemicals. Still, many are difficult to develop in quantities large enough to pose a lethal threat to municipal water systems; however, smaller, nonlethal doses may be used to induce sickness or terrorize the population. Table 20.1 includes a list of potential toxins and their sources.

Microbial agents. Includes bacteria, virus, protozoa, and other microbes. Experts and the U.S. government believe biological weapons are within the reach of terrorists. However, the education level, monetary resources, and risks required to produce these agents are higher than those required for physical methods of terrorism. The high microbial concentrations place the developer of such weapons at high risk. Table 20.2 provides a partial list of potential bacterial and viral agents.

Industrial chemicals. These are yet another threat due to large, readily available supplies. Several factors are important in the analysis of a chemical threat: volume of water to be contaminated, solubility of contaminant, lethal dose, and volume of water that must be ingested to constitute a lethal dose. Fortunately, the vast majority of industrial chemicals make poor candidates as a lethal, undetectable agent. Often the lethal dose required to contaminate a water supply requires a very large quantity or is even insoluble at the required concentration. Furthermore, many

TABLE 20.1 Biological Toxins

Agent	Source
Abrin	Plant (rosary pea)
Aconitine	Plant (monkshood)
alpha-Conotoxin	Cone snail
alpha-Tityustoxin	Scorpion
Anatoxin A(s)	Blue-green algae
Batrachotoxin	Arrow-poison frog
Botulinum toxin	Bacterium
C. *perfringens* toxins	Bacterium
Ciguatoxin	Marine dinoflagellate
Diphtheria toxin	Bacterium
Maitotoxin	Marine dinoflagellate
Microcystin	Blue-green algae
Palytoxin	Marine soft coral
Ricin	Plant (castor bean)
Saxitoxin	Marine dinoflagellate
SEB (Rhesus/Aerosol)	Bacterium
Shiga toxin	Bacterium
T-2 toxin	Fungal myotoxin
Taipoxin	Elapid snake
Tetanus toxin	Bacterium
Tetrodotoxin	Puffer fish
Textilotoxin	Elapid Snake

Source: U.S. Army Medical Research Institute of Infectious Disease, 2001.

toxic chemicals have disagreeable colors, tastes, and odors that would alert the consumer to their presence. Although poisoning the water supply through the use of industrial chemicals is difficult, it is not impossible to make the water unfit for consumption or to simply terrorize the target population.

War agents such as nerve, blood, choking, and blister agents. These include sulfur, mustard, and sarin gases among many others. These have been developed by multiple countries for use generally as incapacitating or discomfort agents. They are not considered as likely toxic water contamination threats due to the high concentrations required. Nerve agents are the most deadly within this category, as they are 100 to 1000 times more lethal than pesticides made with organophosphorous chemicals (Smithson and Levy, 2000).

TABLE 20.2 Potable Water Pathogens

Scientific name	Common name	Availability in environment
Bacillus anthracis	Anthrax	Infected cattle, goats, swine, sheep, horses, mules, dogs, cats, wild animals, and birds
Brucella melitensis & *Brucella suis*	Brucellosis	Infected cattle, goats, swine, sheep, horses, mules, dogs, cats, fowl, deer, and rabbits
Vibrio cholerae	Cholera	Human excrement and shellfish
Clostridium perfringens	Clostridium perfringens	Soils, water body sediment, intestinal tracts of fish and mammals, crabs, and other shellfish
Cryptosporidium parvum	Cryptosporidiosis (Crypto)	Calves
Encephalomyelitis virus	Encephalomyelitis (VEE)	Rodents and horses
Picornaviridae & reoviridae	Enteric viruses	Humans
Burkholderia mallei	Glanders	Horses
Yersinia pestis	Plague	Prairie dogs, chipmunks, black rats, deer mice, certain species of ground squirrels, and coyotes
Chlamydia psittaci	Psittacosis	Birds
Coxiella burnetii	Q fever	Cattle, sheep, and goats
Salmonella typhimurium	Salmonella	Fowl, swine, sheep, cattle, horses, dogs, cats, rodents, reptiles, birds, and turtles
Shigella dysenteriae	Shigellosis	Sewage
Variola major & variola minor	Small pox	Centers for Disease Control and Russia Biological Lab
Francisella tularensis	Tularemia	Wild rabbits and most other wild and domestic animals
Ebola & hantaviral among others	Viral hemorrhagic fever (VHF)	Ticks and rodents

Source: Prescott, L. M., J. P. Harley, and D. A. Klein. 1999. *Microbiology* (4th ed.). Boston: McGraw-Hill.

Radioactive material. The primary radiological threat is the use of conventional explosives to spread radioactive contamination over a limited area or strategic terrain. This could include highly radioactive materials, such as spent fuel traveling to the Yucca Mountains for containment or low-level radioactive materials, including uranium-238, iridium-192, cesium-137, strontium-90, or cobalt-60. Using radioactive materials to contaminate drinking water presents a challenge, requiring large quantities of materials, many of which are insoluble in water, heavy, and would settle out before reaching the target or would be trapped in filters. Furthermore, radioactive material poses a safety concern for the attacker.

20.3 VULNERABILITY ASSESSMENT

This section adds consideration for terrorist activities to the four basic steps in a vulnerability assessment:

1. Identify and describe the separate components of the water supply total system.
2. Estimate the potential effects of probable disaster hazards on each component of the system.
3. Establish performance goals and acceptable levels of service for the system.
4. If the system fails to operate at desired levels under potential disaster conditions, identify key or critical system components responsible for the condition.

20.3.1 Step 1: Identify Major System Components

Key elements of the total system should be listed and described as components under the following general headings: (1) administration and operations, (2) source water, (3) transmission system, (4) treatment facilities, (5) storage, (6) distribution system, (7) electric power, (8) transportation, and (9) communications.

Describe system components with as much detail as possible. Typical items included in a general description are pressure zones, location of pressure-relief valves, pipe sizes, pipe material and ages, typical distance between hydrants, and major valve locations.

In addition to these physical components, an additional item to be considered is the public confidence and reputation of the utility. An important asset of water treatment and distribution systems is the public perception and confidence in the end product. This must also be considered vulnerable to terrorist attack.

20.3.2 Step 2: Determine the Effects of Probable Disaster Hazards on System Components

The effects of a terrorist event on a water utility can result in a wide range of consequences. For example, an explosive device detonated at a noncritical location may not cause any appreciable damage to the facilities and therefore not endanger the capability of the facility to process water. Conversely, chemical contamination of the system can result in long-term disruption of service until it can be cleaned and returned to service. The range of consequences that could be attributed to terrorist activities include

- Disruption of water treatment, storage and delivery, and delivery components
- Introduction of biohazards and toxins into the water system
- Injury to facility personnel
- Injury to the general public
- Damage to utility property or equipment
- Damage to private property
- Hazardous waste disposal problems—what happens to contaminated water if it is flushed from the system?

In addition, terrorist activity may focus on more than one part of the system, using damage in one area to divert the response team or to magnify the consequences of damage to another element of the system.

20.3.3 Step 3: Establish Performance Goals and Acceptable Levels of Service for the System

A water system is considered a lifeline because water is essential to the safety and health of the population it serves. A utility should develop specific goals and acceptable levels of service under disaster and recovery conditions. The acceptable goals for system service should consider the effects of terrorist activity. Taken individually, the effects are identical to those caused by various natural and human-induced hazards. Specific goals to consider are life safety, fire suppression, public health needs, and commercial and business uses.

20.3.4 Step 4: Identify Critical Components

Identifying the critical components of the system or its subcomponents is the final step in the vulnerability analysis. Critical components are those most vulnerable to failure or partial failure because of an intentional act or natural disaster. Failure of a critical component will reduce the system's ability to meet minimum health and safety performance goals. To identify those components that would fail in an intentional attack, run a desktop exercise of an attack scenario and then focus on those components whose failure would render the entire system inoperative—these are the most vulnerable components.

Consideration of critical components should include public perception of the value of a safe water system. In most cases, confidence in the safety of the water system and financial support are closely linked. The public will most likely choose to provide support for security measures if it is perceived that a well-conceived security master plan is in place.

20.3.5 Critical Review of Existing Security Systems

The purpose of security systems is to limit vulnerability. It is important to assess existing security measures and their integration with previous steps to determine the effectiveness of these systems. Do the existing security measures protect major and critical system components, and do they minimize risk? The security review should include facility inspection, a review of documents and operations (policies, plans, and standard operating procedures), and interviews with employees.

20.4 MITIGATION

Mitigation actions to reduce system vulnerability to terrorism are very similar to methods discussed at length in other publications about mitigating vandalism or natural hazards. However, mitigation actions also include elements that prevent unwarranted access to a system's components, which in some cases may be contrary to the ease of access (e.g., reaching an intake by boat) that is necessary for mitigating an unintentional disaster. The extent to which these measures are applied will directly depend on the acceptable amount of risk and the likelihood of the hazard materializing, which are determined by the hazard analysis and vulnerability assessment.

20.4.1 Mitigation at the Source

Standard mitigation measures relating to source water and transmission range from providing alternate sources and protecting wellheads to retrofitting dams or aqueducts. Mitigation actions for watershed damage or widespread contamination include providing automated monitoring equipment, using alternate sources or intakes, and modifying source water treatment at the plant. Controlling access, identifying alternate sources, and providing flexible treatment facilities can also mitigate the effects of deliberate contamination of reservoirs. Access to reservoirs and other outlying system components, such as pump houses and tanks, can be controlled by installing fences, gates, and signs; closing unnecessary roads; and increasing security patrols.

20.4.2 Preventing Access to Facilities

To prevent access to facilities, install adequate locks, window security, and lighting. Install intrusion-prevention devices, such as electronic keys, identification-card checkers, and 10-key code units, to control access to the strategic facilities.

If you have the personnel to perform constant monitoring or can contract with a security firm, you can also install closed-circuit television monitoring systems or alarm systems with ultrasonic, heat, or beam sensors and magnetic switches to detect intruders. Be sure to change passwords when employees are dismissed or a contractor's job is done, both on electronic keys and on computer systems.

Don't forget about controlling access to chlorine and other chemical systems, and design and construct these systems with automatic control systems that can indicate the extent and locations of a leak, actuate chlorine scrubbers, close valves, shut down equipment, and isolate affected areas.

20.4.3 Distribution System Issues

Because of their large numbers and widespread dispersal, controlling access to distribution system components can be difficult. One key to accessing or locating any attacks on the distribution system will be accurate and up-to-date system plans, which should be maintained in numerous locations.

Work with your local fire department to be sure there is adequate alternative pumpage capability and water supply in the event of loss of flow to fire hydrants in an emergency situation. Consider what action should be taken to provide these redundancies if a pump house becomes disabled, either by loss of power or through destruction of the pumps, and where the water would come from if a storage tank were drained or a main break interrupted the flow.

Mitigation activities should take into account what happens to contaminated water if it needs to be flushed from the system. Does exposing it to the air cause even more potential for dispersal of the contaminant? Is the water normally released to a storm drain or waterway?

20.4.4 Staff Role in Preventing Terrorism

In addition to the mitigation techniques identified, facility personnel can also play a very important role in preventing terrorism and acts of sabotage to facilities. In the wake of recent terrorist actions, a new sense of patriotism has emerged. Facility personnel who previously may have been hesitant to take an active role in security activities are now generally eager to be involved in facility security. Most Americans are recognizing that they can play a role in protecting the citizens of the United States at home. Include the frontline employees when discussing what could happen and what action to take—these people are the eyes and ears of the system and may be the first to spot trouble.

Water utilities should be much more careful in disseminating information regarding facility operations, plant and system layouts, and emergency response and crisis management plans. These could prove useful to terrorists wishing to identify system vulnerabilities.

Two factors that will reduce water system susceptibility are the reduction of information regarding facility operations and preventing access to the target through countermeasures.

Reduction of Information. Information concerning the plant and distribution system layouts and emergency response and crisis management plans could help terrorists identify system vulnerabilities. These sources of information should be removed from public libraries, the Internet, and other available locations. Furthermore, those that legitimately claim to need access to such documents should be scrutinized.

Countermeasures

Access control. Includes a variety of systems designed to control the movement of persons or vehicles. Access control may include guards, locks and keys, or access cards.

Physical barriers. These may have already been established to prevent an intrusion or attack but usually only function to hinder or slow the attack. Such measures may include hardened construction, vandal-resistant glazing, and doors. Fencing is not a good barrier, as it is only considered to provide a 6- to 10-s delay. Physical barriers in regard to contaminants may include backflow protection, filters for airborne particulates on plant and reservoir vents, or specially designed reservoir vents to prevent the pumping of a contaminant into the distribution system.

Detection. Some type of monitoring system provides notification that an undesired intrusion (physical or contaminant) has occurred. Examples include name badges to identify unauthorized personnel, closed-circuit cameras, motion detectors, security guard scrutiny, or sensors in the distribution system. Detection should be as far away from the asset as possible to allow time for the response. Closed-circuit television systems make excellent assessment devices. Assessment refers to the verification of the detection system. For instance, personnel can use monitors to help assess the situation after an alarm goes off.

Once the utility decides which countermeasures to implement, any physical improvements should be combined with updated policies and procedures to ensure full optimization of the system.

20.5 RESPONSE PLANNING FOR PUBLIC DRINKING WATER SYSTEMS

All elements of a water treatment plant are susceptible to human disruption, including the raw water source, the treatment facility, the operations and control facility and systems, and the support facilities and systems such as chemical feed, power supply, and communications equipment. A preparedness plan and a response plan can be developed based on the results of the vulnerability assessment. The exact content of each response plan hinges on the vulnerability and the risk for any given water utility. A response plan must define the measures that will be implemented to minimize the likelihood of an event or to mitigate its impacts.

According to the Studies in Urban Security Group (SUSG) at the College of Architecture and Urban Planning in Ann Arbor, Michigan, the development of an emergency response plan for the contamination of a public water system may be subject to federal, state, or local regulations or guidelines (Rycus, Snyder, and Meier, 1989). Presidential Decision Directive 63 required that federal agencies develop and implement plans to protect the nation's critical infrastructure (President, 1998). The Safe Drinking Water Act Amendments of 1986 and the Emergency Planning Community Right to Know Act of 1986 require that each state appoint a State Emergency Response Commission, whose responsibilities include designating emergency planning districts within the state.

Following the designation of the emergency planning districts, local emergency planning committees (LEPCs) should be formed. These committees should consist of representatives of public agencies, such as water and wastewater utilities, fire departments, health officials, law enforcement, and government officials. The LEPC then reviews the responsibilities of the water service during storms, floods, earthquakes, fires, explosions, nuclear reactor spills, aircraft crash, hazardous material incidents, power failures, and civil disturbances when formulating the response plan for human disturbance.

The American Water Works Association publishes a manual on emergency planning (M19: *Emergency Planning for Water Utilities*) that provides utilities with information helpful in formulating a response plan. In addition, the SUSG study states that a response plan should contain

- A legal and administration basis
- A classification of emergency conditions
- Provisions for command and control, communications, emergency supplies and distribution, threat management, and plan review and revision (Rycus, Snyder, and Meier, 1989).

Many utilities may already have a response plan. The authors of the new or revised response plan should verify that the plan contains the most current contact names, phone numbers, and e-mail addresses; that the content is the product of a proper and complete vulnerability and risk assessment; and that it contains the components listed in this supplement.

20.5.1 Define Emergency Status

In developing a response plan, the term *emergency* must be defined, preferably with variations in degrees of alert/emergency status. The SUSG study defines four levels of severity:

1. *Normal operations/minor emergencies.* Required responses do not extend beyond the water department, so an emergency response is not warranted.
2. *Alert condition.* In situations where a major emergency may be forthcoming, the system director may declare an alert condition. An alert condition triggers the assembly of key decision makers and operational personnel to assess and monitor the situation.
3. *Emergency condition.* In a situation where disruption or contamination is imminent or has occurred and where the full resources of the system, augmented by external resources (e.g., fire, police, public health), are required for appropriate response, the system director should declare an emergency condition. Under this condition, a full response plan would be implemented.
4. *State of emergency.* At this most serious level, especially involving the community at large, declaration of a state of emergency would be appropriate. Typically, only the governor of the applicable state can declare this condition, and it implies that the broadest resources available will be applied to the problem (Rycus, Snyder, and Meier, 1989).

20.5.2 Developing an Emergency Response Plan

This section details the nine general steps that should be considered in the development of a response plan for a public drinking water system. When developing a response plan for a specific water utility, it is important to evaluate each of the following steps, investigate the relevance of each step to the specific water utility, expand on the information, and provide any useful information.

Step 1: Gather Command Group at a Designated Location. The SUSG study recommends that the command group include persons with specific expertise, such as the chief operator of the water utility, a staff member capable of providing technical support, staff member for administrative support, the head of engineering, the director of laboratories, and the director of security. The group should be organized according to an established chain of command.

The command group members should meet at a prearranged location, or emergency operations center (EOC), that contains working communications equipment, including a computer with information on distribution system plans, source water plans, any secondary or tertiary utility plans, the Internet, a telephone, and a radio (Rycus, Snyder, and Meier, 1989). This location will be the main operations center. A backup location should also be established in the event the primary location cannot be used.

Step 2: Conduct Preliminary Assessment to Determine the Nature, Extent, and Severity of the Disturbance. A plan for assessing a disturbance should be prepared in advance. The assessment should address control around the perimeter of the utility's assets and an internal review of water treatment plant data encompassing the laboratory, water treatment plant performance, disinfection contact times, chemical feed, and other relevant data. The personnel who participate in the assessment should have an intimate understanding of plant operations and the significance of the data.

Water utilities are not required to monitor their facilities for biological, chemical, or radiological contaminants or cyber interference. However, continual monitoring, routine testing, and careful observation are recommended to prevent contamination of supplies or interference with operation and to facilitate prompt mitigation of any disturbance. The local department of health or the laboratory that performs the analyses should immediately report any contamination or suspicious test results to a designated member of the command group, most likely the water utility supervisor. Any information from the laboratory data will enhance the assessment.

One way to determine the threat posed if a chemical contamination occurs within the utility is a computer-aided management of emergency operations (CAMEO) database. Many SERCs have access to a CAMEO database, which includes information on chemicals, transportation of the chemicals, and other information to better prepare for response to chemical emergencies. The USEPA's Chemical Emergency Preparedness and Prevention Office and the National Oceanic and Atmospheric Administration Office of Response and Restoration developed CAMEO. The USEPA also published a document about SERCs and the involvement of the CAMEO program titled *Secrets of Successful SERCs* (USEPA, 1993).

Based on the results of the assessment, the command group should decide whether to declare a water utility alert or a water utility emergency and, if necessary, implement the appropriate emergency response.

Step 3: Assemble Specialized Groups. Specialized groups should be assigned before the emergency occurs. This will save time and expedite the group's response actions. The assignments of the specialized groups may include

- Situation assessment
- Laboratory analyses
- Law enforcement and security
- Public information/media communications
- Emergency water supply
- Emergency evacuation
- Human health reporting/assistance
- Repair/recovery

A secured and current membership list for each group should include each member's position, name, and phone numbers. This list should also be included within the response plan.

Step 4: Alert Other Officials. If an emergency situation is declared, the command group should request further assistance from the appropriate law enforcement or other government agencies, including

- *Police.* Exterior investigation, directing traffic, security
- *Fire department.* Emergency fire protection
- *Emergency medical system.* Emergency health care
- *Public Health Department.* Emergency health investigations
- *USEPA.* Environmental impact aid
- *FEMA.* Emergency response aid
- *Department of Transportation.* Infrastructure aid
- *State National Guard.* Additional personnel and security
- *Other water agencies*

Step 5: Communicate with the Media. At the onset of an alert, emergency condition, or state of emergency, the designated media spokesperson should be notified. This assignment should be given to a person who can communicate effectively with the media. A media center should be established for both written and verbal press releases. It is also important to monitor media coverage. Sensible public information and communications during an emergency are crucial to the implementation of any type of response plan. See Sec. 20.6 for more information.

Step 6: Consider Human Health. If an alert, emergency condition, or state of emergency has been called, it is important to identify the areas most likely to be affected. If health concerns exist, local government officials, local water departments, hospitals, health departments, emergency medical teams, and fire departments must be notified. Depending on the type and severity of the situation, federal agencies such as the department of public health, FEMA, USEPA, and the CDC may also have to be alerted.

Step 7: Determine Alternative Sources for Emergency Water Supplies. An emergency situation may require the use of water from other sources, such as unaffected reservoirs, fire stations, or independent water supplies. It is important to identify the available sources of emergency water supplies before the need arises. A plan for emergency water distribution and for conveying the water to the consumers should be developed. It is also important to set priorities for where the emergency water will be distributed and how it will be used.

Utilities may choose to partially neutralize the risk of a break in service through a number of avenues:

- Partitioning the distribution system so that certain areas can be shut down without affecting a large proportion of the entire system.
- Estimating tank truck availability and inflatable water storage units (internal and external to the utility) for providing emergency water.
- Estimating hospital and sensitive-user water needs and ensuring water supply to these users in case of an emergency.

Step 8: Establish Relationships with Nearby Water Utilities and Supporting Utilities. Maintaining relationships with other water suppliers will expedite a utility's ability to tap into alternative sources of supply in an emergency. Such resources may include independent or private water utilities and purveyors, bottling companies, and some large water-using industries that have their own supplies.

Step 9: Plan and Implement Countermeasures and Recovery Measures. Determine the anticipated types and extent of laboratory tests and establish a testing plan.

Additional Information for the Response Plan

- Formulate an extensive list of the area laboratories, with phone numbers and individual contact names.
- Formulate a list of contacts with appropriate state and local government agencies and other companies such as power utilities and bottling companies.
- Annually review and update the response plan.
- Maintain evaluation forms and require proper completion if contamination has occurred.

20.5.3 Threat Management

Threat management is a major part of preventing an emergency situation. In the event of a threat, either written or oral, it is necessary to record the contents of the threat and quickly notify the command group to evaluate the threat. Each water utility should have an evaluation form that should be filled out by the employee who receives the threatening call or letter. See the sample threat evaluation form in Fig. 20.1.

Type of threat indicated by caller: ______________________________

Specific details of threat: ______________________________
____
____
____

Caller's sex (circle one) Male Female Not sure

Caller's age (circle one) Under 10 10–20 Over 20 Not sure

Describe the voice of the person placing the threat: ______________________________
____
____
____

Describe any background noises or unusual sounds during the phone call: ______________
____
____
____

Did the caller name any organization? ______________________________
____
____
____

Did the caller give any other information? ______________________________
____
____
____

Name of person completing this form: ______________________________

Date and time threat was received: ______________________________

FIGURE 20.1 Sample threat evaluation form. (Based on Rycus, Snyder, and Meier, 1989)

20.6 CRISIS COMMUNICATION

In times of extreme crisis, such as the events of September 11, 2001, citizens appear more trusting of governments (and, by extension, utilities). They look to mass media, particularly television, to bring them vital information about how to behave and protect themselves. Therefore, it is critical for utilities to have an effective, efficient plan in place to guide communications with the media, and hence with their public, in the event of crisis. Utilities should not rely on politicians to communicate crises but should keep state, local, and federal agencies continually updated on the crisis.

It is critical for the utility to communicate in a way that fosters trust and confidence. The utility must be first to release information about how the crisis has impacted the drinking water supply. Delayed release of information can result in loss of life, accusation of cover-up, and ultimate loss of public trust. In a crisis situation, the guiding principle for utilities is, "Be the first to deliver the bad news." Effective crisis communication has no room for images and egos—it is only about saving lives and protecting the public health.

While the crisis communication plan is being developed, its proposed content and direction should be reviewed with the public (through public meetings, discussion groups, civic club meetings, public libraries, etc.) to confirm that it meets specific public needs for information. Each segment of the utility's public will have its own needs for information and information delivery, and every effort should be made to accommodate these special needs. For example, the frail elderly may have one set of crisis concerns, pregnant women another, infants and young children yet another. The utility should develop a list of sensitive stakeholders, contacts, and phone numbers based on the public review findings. The list must be kept current.

One beneficial outcome of the public review is building a constituency of credible sources with which to partner in case of crisis. Potential candidates for credible sources include local universities, the League of Women Voters, parent–teacher organizations, the medical profession, and environmental groups.

20.6.1 Preparing for a Crisis

The following steps should be taken to prepare for a crisis:

1. Develop a communication policy and a plan of action for use in case of a crisis.
2. Form a "crisis team" and define member roles and responsibilities. The team should include representatives from top management, operations, public affairs, government affairs, legal, insurance, human resources, finance, and others.
3. Compile a list of contact numbers (e-mail might not be available in a crisis) for each team member and all top management members.
4. Identify primary and backup spokespersons for the organization. The spokesperson should be authoritative without appearing arrogant.
5. Notify all utility personnel that only the designated spokesperson should speak with the media.
6. If the utility uses an outside public relations firm, define its roles and functions. If the utility does not maintain an in-house staff of public communicators, pre-

arrangements should be made to use an outside firm in the event of a crisis. Should a crisis occur, the utility's public relations staff will be very busy. Staff members must be prepared ahead of time with such basic tools as scripts covering various crisis scenarios, lists of frequently asked questions, and media kits, including background information on all possible crisis threats (such as various biological elements).

7. Develop media (television, newspaper, and radio), community, and government contact information (phone, fax, and names of key contacts) and keep it current.

8. Create a "media" office location and backup location. Provide the office with appropriate supplies and equipment to handle electronic communication, faxing, phoning, and overnight mailing. Also provide television, radio, and VCR capability for monitoring media coverage and recording it. Determine security and access rights for both nonemergency and emergency situations. Be prepared to set up a podium with an appropriate backdrop. Set up a work area with phones for the press.

9. Involve the public to determine which groups might require specific direct communication and if the plan can be improved so as to best serve customers' health and safety needs in a time of crisis.

During a Crisis. These steps should be taken during a crisis:

1. Activate the "media" office. Communicate activation with security and implement appropriate access rights.
2. Collect everything known about the crisis situation and put it in writing for review by the crisis team. Update the report regularly.
3. Contact the external public relations firm if appropriate.
4. Immediately after review (within 1 h of the event), develop and distribute the initial media statement. Immediately update the utility's web page with this and future press releases. Provide the media with the name and phone number of the spokesperson to reduce the potential for rumors and inaccurate information.
5. Communicate individually with specific special interest groups of the utility, including:
   - Stockholders
   - Advisory groups
   - Governmental entities and elected officials
   - City council
   - Medical community
   - Others as identified in the public involvement effort

6. Maintain files of all hard copy, electronic, media, and video communications during the crisis event.
7. Provide crisis response office staff with lists of phone numbers and contacts.
8. Provide frequent updates to employees regarding the crisis event and the utility's responses and press releases. Post the list of phone numbers for crisis team members.
9. Request that all employees refer media requests to the designated spokesperson.
10. Establish a schedule of crisis team meetings (morning and afternoon) to update information and issue responses to the media and employees.
11. Provide the media with background information, including information about how the utility protects public health and safety.
12. Request support from the AWWA and other appropriate federal, state, and local organizations.
13. Work with previously established partners from the medical community to issue joint statements and press releases. Make joint television appearances.

Some general guidelines for dealing with the media include:

- The spokesperson must be in control of his or her emotions, remain calm, appear authoritative but not arrogant, and be extremely polite.
- Anticipate likely questions. Understand the perceptions and fears of the public. Address fears by offering facts, not conjecture. Do not answer "what if" questions. Do not use emotional statements or industry jargon.
- It may be necessary to communicate with the public before all facts are known. Be humble. State priorities (e.g., getting the situation under control) and assure the audience that the utility staff is doing everything it can.
- Assume that there is no such thing as "off the record."
- Avoid conjecture or assigning blame. Both could result in rumors, embarrassment for the utility, and future litigation against the utility.
- Keep communications succinct so as not to confuse or dilute the main message.

20.7 REFERENCES

Denileon, G. P. 2001. Introduction. The Who What, Why, and How of Counterterrorism Issues. *Journal AWWA* 93(5):78–85. *http://www.awwa.org/public%5Fep/may.pdf* (accessed November 20, 2001).

Department of Defense. 2001. Introduction. *Department of Defense Chemical and Biological Defense Program; Annual Report to Congress and Performance Plan: 5-8.*

http://www.defenselink.mil/pubs/chem_bio_def_program/2001_CBDP_Annual_Report.pdf (accessed October 25, 2001).

Prescott, L. M., J. P. Harley, and D. A. Klein. 1999. *Microbiology* (4th ed.). Boston: McGraw-Hill, pp. 728–729.

President. 1998. White Paper. The Clinton Administration's Policy on Critical Infrastructure Protection: Presidential Decision Directive 63. *http://www.ciao.gov/CIAO_Document_Library/paper598.htm* (accessed November 20, 2001).

President's Commission on Critical Infrastructure Protection. 1997. Appendix A, Sector Summary Reports. *Critical Foundations: Protecting America's Infrastructures:* A-45. *http://www.ciao.gov/PCCIP/PCCIP_Report.pdf* (accessed November 21, 2001).

Rycus, M. J., J. C. Snyder, and P. G. Meier. 1989. Strategic Security Planning for Public Drinking Water System—Prepared for USEPA, Office of Research and Development. The University of Michigan Studies in Urban Security Group, College of Architecture and Urban Planning. Ann Arbor, Michigan.

Smithson, A. E., and L. A. Levy. 2000. Grounding the Threat in Reality. *Ataxia: The Chemical and Biological Terrorism Threat and the U.S. Response.* Stimson Center Publications. *http://www.stimson.org/cwc/ataxia.htm* (accessed November 20, 2001).

U.S. Army Medical Research Institute of Infectious Disease. 2001. *USAMRIID's Medical Management of Biological Causalities Handbook. http://www.usamriid.army.mil/education/bluebook.html* (accessed November 20, 2001).

U.S. Environmental Protection Agency. 1993. *Secrets to Successful SERCs.* EPA 550-F-93-002: 5.