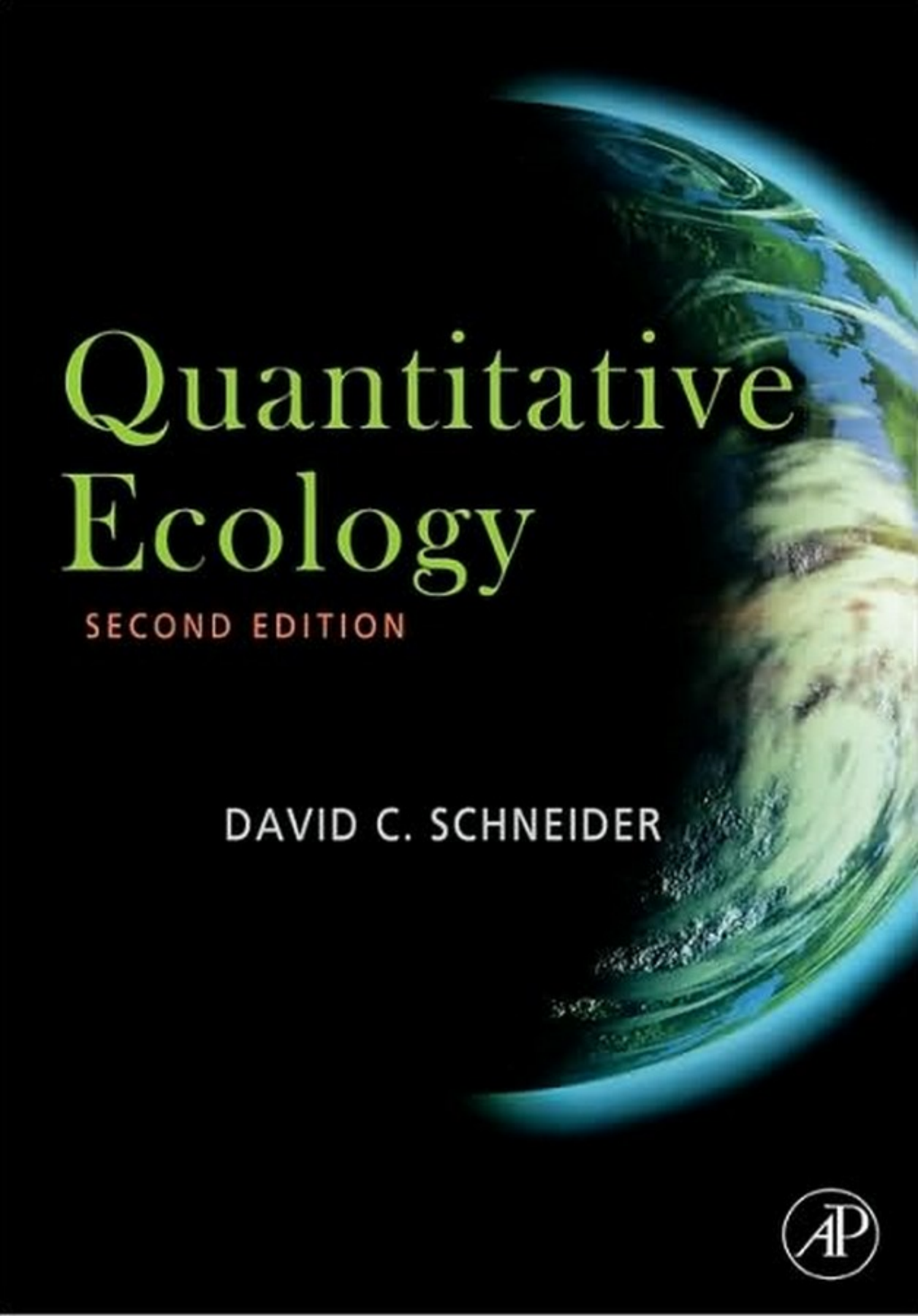
Quantitative Ecology

SECOND EDITION

DAVID C. SCHNEIDER



Academic Press is an imprint of Elsevier
32 Jamestown Road, London NW1 7BY, UK
30 Corporate Drive, Suite 400, Burlington, MA 01803, USA
525 B Street, Suite 1900, San Diego, CA 92101-4495, USA

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

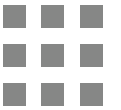
Library of Congress Cataloging in Publication Data
A catalog record for this book is available from the Library of Congress

ISBN: 978-0-12-627865-1

For information on all Academic Press publications
visit our website at www.elsevierdirect.com

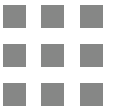
Printed and bound in the United States of America

09 10 11 12 13 10 9 8 7 6 5 4 3 2 1



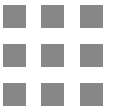
Dedication

To R., B., and the memory of M.



Preface to First Edition

This book advocates a style of quantitative thinking that I have found useful in ecological research and that I hope others will find of benefit in the design of their research or in reading and evaluating the work of others. The style relies heavily on the use of units and dimensions, which I believe are critical to the successful integration of "scale" into ecology. I have made every effort to develop concepts and generic methods rather than present a set of commonly used recipes or techniques that I happened to have mastered. It certainly would have been much easier to write a book on statistical biology, which I teach each year. The book is (I hope!) aimed at people who are thinking about ecology as a career. But I trust that ecologists with field or theoretical expertise will also find something of interest to them.



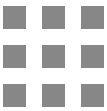
Preface to Second Edition

It is a privilege to have the opportunity to revise a book that appeared when the concept of scale was just emerging in ecology. This edition retains the first edition's emphasis on quantitative thinking founded on measurement, units, and dimensions. The presentation of the material has now been focused on two concepts: scope and power laws. These have in turn been used to organize substantial amounts of new material on research design and statistics, topics that were not touched on in the first edition. The audience is the same: people who are at the beginning of a career in ecology and the environmental sciences. But I trust that people already established in their careers will find something of interest to them as well.

Acknowledgments

The content of this book grew out of a postdoctoral collaboration (1980–1984) with George Hunt at the University of California. I thank Dick Haedrich and Memorial University for the extraordinary opportunity to develop scaling as a research direction during a two-year research fellowship. This experience became the core of the first edition, written during a sabbatical stay as a visiting scholar at Scripps Institution for Oceanography in 1991–1992. I thank Paul Dayton and SIO for a productive opportunity to apply "reasoning about scaled quantities."

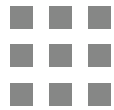
This second edition grew out of an extended workshop on spatial scaling organized by the New Zealand Institute for Water and Atmospheric Research in 1994. I thank Simon Thrush and NIWA for the opportunity for extended collaboration. Thanks also to David Duffy for the opportunity to apply scaling concepts to pressing issues on conservation biology during a sabbatical at the University of Hawai'i. Thanks also to Bruce Milne and Jim Brown for the opportunity broaden my research experience to landscape ecology during a sabbatical at the University of New Mexico. And thanks to Susan Williams for the opportunity to explore the temporal component of a biodiversity scaling during a sabbatical stay at the Bodega Marine Laboratory. I am indebted to Chris Bajdik for challenging me to teach the general linear model to undergraduates. This edition has been shaped by working with several graduate and honors students, most notably Elizabeth Bennett, Tammo Bult, Deneen Cull, Joanne Ellis, Charles Feng, AnnMarie Gorman, Ben Hammond, John Horne, Dan Ings, Ben Laurel, Stephen Mayor, David Methven, Mike Norris, Miriam O, John Rex, Lee Sheppard, Maria Thistle, Nadine Wells, and Richard Wong. Thanks also go to Robert Gregory for substantial logistical contributions that made possible field application of these concepts by students. Finally, warm thanks to Chuck Crumly, Andy Richford, and Emily McCloskey for their editorial patience and wisdom.



PART

I

Introduction



1 Quantitative Reasoning in Ecology

We have noticed that much of the confusion and misunderstanding in the contemporary literature of evolutionary theory and ecology, fields that have received more than their share of polemics, arise when the disputants can't measure it. In the past progress usually followed when ideas were abstracted into sets of parameters and relations that could be built into models or when new methods of measurement were invented.

—E. O. Wilson and W. H. Bossert, *A Primer of Population Biology,* 1971

1.1 The Need for Quantitative Reasoning

Important questions in ecology are more than a matter of biology; they affect our well-being and they have an ethical component. An example is species diversity on a tropical reef. This diversity poses one of the central questions in biology: Why are there so many species in such a small area, and what caused such diversity? This variety of species affects the well-being of island people who fish the waters of the reef. Their lives and culture are adjusted to the peculiarities of the species found on the reef. The well-being of island people who make a living from tourism depends on the continued diversity of the reef. At a larger scale, the well-being of still more people will be altered by therapeutics discovered among the many unique physiologies that coexist on a reef. The diversity of species on a coral reef also raises ethical issues. Reef inhabitants, tourists, and the beneficiaries of new therapeutics lose if diversity is eroded by destructive practices such as dynamiting for fish or discharging nutrients over a reef. The ethical issues created by fish dynamiters are clear-cut; the ethical issues raised by tourists are more complex via unintended consequences such as reef loss through increased discharge of nutrients in waste streams.

For the ecologists who study ecosystems, pressing questions have an additional characteristic—that of the complexity of multiple causation. Among the processes responsible for coral reef diversity are regional history, frequency of disturbance, and differential predation on common species. These processes often interact. For example, a sessile species may overgrow its neighbor at low rates of nutrient supply, but not at high rates. The interacting effects of nutrient supply and competitive capacity maintain diversity by repeatedly reversing the outcome of competition for space. Multiple causes act at different space and time scales. The number of species and their relative abundance on a patch reef will depend not only on the competitive interactions among neighboring organisms but also on the resupply of larval recruits from other patch reefs.

doi: 10.1016/B978-0-12-627865-1.00001-5

Pressing problems in ecology are often attended by substantial uncertainty. Some of this uncertainty will arise from measurement error. If we look at the role of fish predators on a reef by excluding them from sites that range from high to low prey density, the experimental outcome will vary in part because of error in measurement of initial and final prey density. The experimental outcome will also vary because of process error, the sum effect of all the unknown factors that change prey density. In a laboratory setting, process error can be reduced substantially by manipulative control. But in field experiments, process error will remain large in even the mostly tightly controlled experiment.

How do we address ecological problems characterized by complexity and uncertainty? We have no choice but to use a model to simplify and make sense of the situation. In biology, the classic solution to the problem of complexity is a verbal or graphical model obtained by the comparative method. A classic example of the success of the comparative method is Darwin's theory of coral atolls. Comparisons of Pacific islands uncover similarities that establish an historical sequence, which begins with a volcanic island fringed by reefs. These continue to grow upward into the light as the island erodes, leaving a ring-shaped atoll. Another example is the phyletic assignment of fossils from the middle Cambrian in the Burgess shale formation in eastern British Columbia. Comparisons of fossil fragments uncover similarities that establish the presence of sponges, echinoderms, chordates, four major groups of arthropods, and species that defy placement in known phyla. Yet another example of a model due to the comparative method is a diagram of the vertebrate nervous system, with its dorsal nerve cord and dendritic structure. The diagram is a model extracted from comparative dissections of organisms as diverse as fish and kangaroos. In these examples generalization resulted from qualitative comparison of units, without quantitative treatment.

The comparative method has a long record of success in biology and geology, most notably where the impress of history is strong, as in morphology, embryology, palaeontology, and stratigraphy. At the much shorter time scales of human actions and pressing ecological problems, the method of comparing and contrasting a series of cases will sometimes serve, but more often measurement will serve better. This in turn will require quantitative models to simplify and extract meaning from the data.

The first step toward a quantitative model is verbal—a statement of the relation among measurable variables (Figure 1.1). An example is the statement that release of nutrients into a lagoon will alter coral abundance. A verbal model can address both causality and uncertainty: the effect of nutrients on coral in a lagoon may occur with considerable local variation. The verbal model is the door to a graphical model, in which data are simplified to a graphical model, as in Figure 1.1. For example, we could sketch the expected form of the relation of coral growth to nutrient release rate from previous data showing that high nutrient levels slow the growth of coral. Data can be simplified

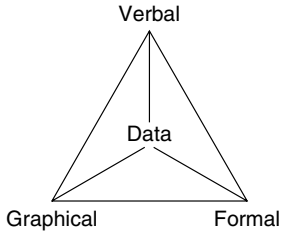


FIGURE 1.1 Data are Simplified to Verbal, Graphical, and Formal Models.

to a formal model. An equation describing growth in relation to nutrient levels in the lab could be used as a preliminary model to predict effects in the lagoon. The number of samples needed to detect the predicted effect could be computed from a model of uncertainty, such as normal error distribution.

At any level, whether verbal, graphical, or formal, a model is a simplification of the complex causality and inevitable uncertainty that attend pressing ecological problems. The simplification provided by a model is a necessary part of disentangling causality. A formal model can also address uncertainty. Such a model will have a structural component (the causal part) and an error component (the uncertainty part). Such models make best use of data in advancing understanding of the problem and in addressing uncertainty during decision making.

ANOTHER LOOK AT SECTION 1.1

1. State an ecological question of interest to you, then briefly state its societal importance, intrinsic biological interest, and ethical implications.
2. State an ecological question of interest to you, then list relevant variables and sources of uncertainty in estimates of these variables.
3. State several possible causal linkages between variables you listed. Can you identify causal pairs that interact? Which, if any, pairs interact at different time and space scales?

1.2 From Words to Graphs to Equations

Formulation of an ecological problem often begins with a verbal expression, usually in the form of a question. So let's open the triangle out into a sequence beginning with verbal expression (Figure 1.2). This figure represents an effective approach to any ecological problem, which is to begin with a question, sketch the suspected relation of key variables, graph the relation of these variables, then estimate the formal relation as a function, as shown. As experience accumulates, these empirical functions then yield to more broadly supported functions based on theory. Examples of conceptually founded functions include the Holling (1959) equation for prey encounter rate and the Ivlev (1968) equation for prey consumption.

The sequence from verbal through graphical to formal expression of a model usually proceeds by iteration rather than as a simple progression. A verbal model, formulated as a question, will be compared to data, leading to a revised question and new model (Figure 1.3). This iterative cycle of challenging models against data continues until an acceptable model emerges.

ANOTHER LOOK AT SECTION 1.2

For an ecological problem of interest to you, list at least four relevant variables. State two variables that are suspected to be related causally, one to the other.

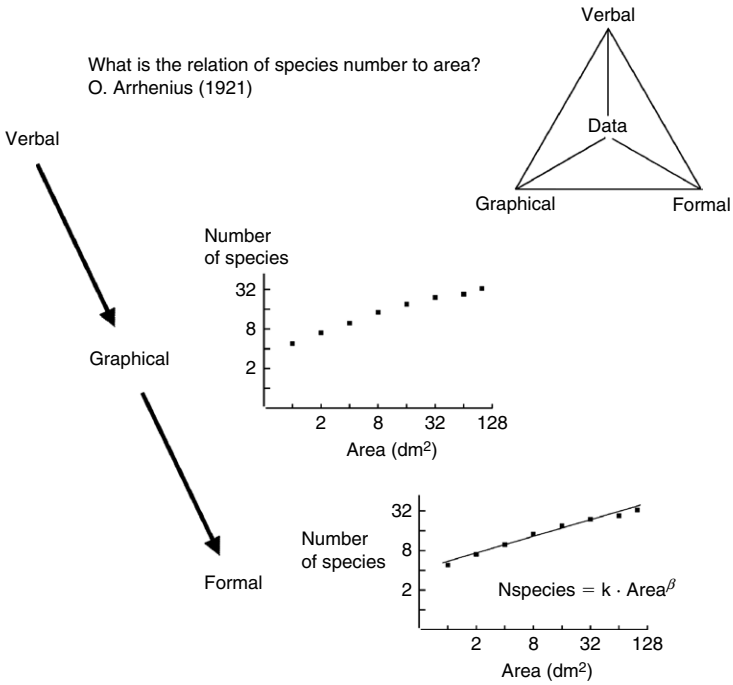


FIGURE 1.2 Typical Sequence of Verbal, Graphical, and Formal Models. Data and model of number of species in herb-*Pinus* wood plots from Arrhenius (1921).

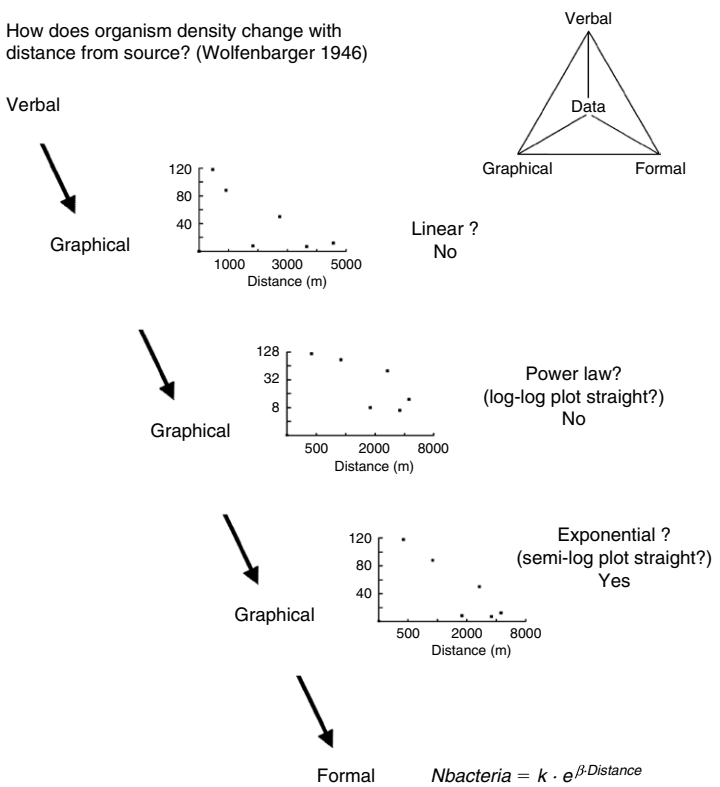


FIGURE 1.3 Verbal Model with Iterative Graphical Identification of Formal Model. Verbal model of dispersal with distance from source from Wolfenbarger (1946). Data are bacteria numbers versus altitude, redrawn from Figure 58 in Odum (1959).

1.3 Statistical Analysis

In ecology the statistical analysis of data is often presented as a set of recipes. But underlying all of the most widely practiced statistical techniques in ecology are models. The mean value is, for example, simply a single-value model that summarizes a larger collection of measurements. Here is that model in the form of an equation:

$$Data = \text{mean}(Y) + Residual \tag{1.1a}$$

Most of the data analyses that ecologists undertake focus on the relation of observed data to one or more additional variables. The most widely applicable techniques are regression, analysis of variance (ANOVA), logistic regression, and contingency tests. All these are based on models. In fact, these common techniques are all special cases of the generalized linear model (McCullagh and Nelder, 1989), which says that data can be partitioned into a structural model μ and an error term ε.

$$Data = Model + Residual \tag{1.1b}$$

$$Data = \mu + \varepsilon \tag{1.1c}$$

The structural model μ is expressed as a collection of means (Figure 1.4a), as one or more lines (Figure 1.4b), or sometimes as odds (such as "Odds of survival are 4:1"). One does not have to assume that the errors ε are distributed normally. One can use any distributional model (e.g., binomial distribution, gamma distribution, Poisson, or normal distribution) to estimate the means or lines in the structural model.

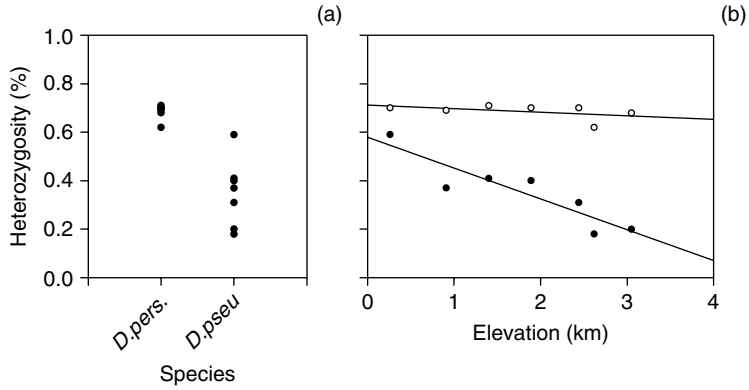


FIGURE 1.4 Genetic variability (H = heterozygosity) in Two Species of Fruit Fly *Drosophila pseudoobscura* and *Drosophila persimilis*. Data from Th. Dobzhansky (1948) as reported by Brussard (1984).
(a) Comparison of means.
(b) Genetic variability in relation to altitude.

The error component ε addresses questions of uncertainty. Given the variability in the data, does the difference between treatment and control lie within the realm of chance? How good is the fit of the model to the data? Is a normal error model correct, or should we use something else? The model component μ addresses questions of causality. It can be structured to separate several causal factors according to an experimental design. It can be augmented with factors that are known to cause variation in experimental outcome but cannot be fixed or controlled. It can incorporate previously established

models, such as an exponential growth rate. Model-based statistics allow one to separate the analysis of causation from questions of reliability and uncertainty. It permits a wider range of analysis than recipe-based techniques because of the great flexibility in choice of both model and error components (Crawley, 1993). One can write a model based on the biology of the case at hand, without the constraints of conforming to the assumptions of normal errors. Model-based statistics allow one to focus on models that express biological concepts, with attention to the scale of measurement and dimensional consistency of terms (Table 1.1).

Table 1.1 Problems with Statistical Practice in Ecology*

Elementary statistics courses for biologists tend to lead to the use of a stereotyped set of tests:

1. Without critical attention to the underlying model involved
2. Without due regard to the precise distribution of sampling errors
3. With little concern for the scale of measurement
4. Careless of dimensional homogeneity
5. Without considering the ideal transformation
6. Without any attempt at model simplification
7. With too much emphasis on hypothesis testing and too little emphasis on parameter estimation

*Crawley, 1993.

Model-based statistics allow effective treatment of the error component. One can choose the most appropriate error distribution, with no need to resort to irrelevant or uninterpretable transformations of the data (Table 1.1).

ANOTHER LOOK AT SECTION 1.3

"If your experiment needs statistics, you ought to have done a better experiment."

Evaluate this statement (attributed to the physicist Ernest Rutherford, 1871–1937) in light of ethical limitations on experimental treatment of human subjects. Evaluate this statement for large-scale environmental effects such as desertification or air pollution.

1.4 Scaled Quantities

The data produced by the act of measurement consist of scaled quantities—units attached to numbers. Yet ecological publications, like statistical texts, routinely treat data as numbers rather than as scaled quantities resulting from measurement. In statistics texts, the units of measurement are often omitted in the interest of generality; the results are meant to apply regardless of the units used to obtain the numbers. In ecology, units matter. For example, counts of saguaro cactus (*Carnegiea gigantea*) in 10 meter units can be expected to vary within the modest range of Poisson distribution, for which the variance is equal to the mean. But counts in 10 hectare units will almost certainly vary beyond that of a Poisson distribution, since clumping at this scale creates variance with values well in excess of the mean. Units matter; the results at the scale of square meters will not be the same as those at the scale of hectares. Nevertheless, the traditional analytic path in ecology leaves off the units, puts numbers into a formula, produces a

numerical result, and often omits the units in reporting the result. When units disappear, so does scale and the capacity to interpret results.

As an example of how units matter, imagine an experiment to determine the effects of an invasive weed on an endemic plant. A classic Fisherian experiment begins with definition of the experimental units (plots), then randomly assigns treatment (adding weed seeds) to half the plots, no treatment to the other half. The Fisherian principle of local control could be addressed in several ways, such as pairing plots based on endemic plant density prior to the experiment, then assigning the treatment randomly to one plot in each pair. Does the difference in density of endemic plants between paired plots increase after the experimental intervention? The conclusion will depend on the size of the unit (plot). The variance in endemic plant density among 1 m^2 plots will differ from variance in density among 1 hectare (100 m by 100 m) plots. Because the units of measure differ (plants per m^2 versus plants per hectare), the error variance in the analysis will differ, in turn altering the results of statistical analysis. The conclusions from an experiment based on m^2 units can easily differ from those resulting from an experiment with hectare units.

Another example showing how units matter is the spatial scale of units to quantify habitat. If one measures habitat area at fine resolution, such as a grid of m^2 units, the estimate of percent cover by each habitat type will certainly differ from the estimate on a grid of 1 hectare units. If one measures the length of an ecotone or habitat boundary at a 1 meter resolution, that estimate will certainly differ from an estimate made at 100 m resolution. Habitat maps produced by two government agencies will almost inevitably differ because different spatial units will have been used. The dependence of outcome on the unit of analysis is well recognized in geography as the modifiable area unit problem (MAUP).

Units matter because without them we cannot determine whether the results of two studies are comparable. An example is the variation in results among experimental studies of nutrient enrichment (Downing et al., 1999). To achieve comparability, these authors restricted meta-analysis to experiments lasting two to seven days. Experiments of a day or less exhibited time lags in response to nutrients, whereas experiments lasting more then seven days were confounded by nutrient depletion or grazing effects. Units also matter because without them, we cannot make accurate computations. An example is the intrinsic rate of increase, usually denoted by the symbol r for exponential population growth. An estimate of r is of no utility unless we know its units, whether they be percent per day or percent per year.

Reporting units might seem obvious, but an examination of the research literature reveals all too many studies in which the units cannot be determined. Table 1.2 shows a compilation of articles published in a single year (1990) in four journals representative of ecology. The journals and fields that they represent are as follows: ecology at the organismal level (*Behaviour*), ecology at the population level (*Ecology*), ecology at the community level (*Ecology*), and theoretical ecology (*Theoretical Population Biology*). A physics journal with a strong environmental focus (*Journal of Physical Oceanography*) was included for comparison. In Table 1.2 these journals are abbreviated BEH, ECOL-P, ECOL-C, TPB, and JPO. Each article was scored by use of units (Good = units used almost always; Fair = units used sometimes; Poor = units used rarely). We tallied the percentage of papers that use scale either explicitly or implicitly in the discussion. A tally was made of the percentage of papers using theory, using experimental data, and using statistical tests of hypotheses. In this compilation the theoretical ecology journal stood out in terms of poor use of units and absence of data, even though the use of theory was at the same level as a physics journal. An update on this survey is left as an exercise.

Table 1.2 Research Style in Behavior (BEH), Population Ecology (ECOL-P), Community Ecology (ECOL-C), Geophysical Fluid Dynamics (JPO), and Theoretical Ecology (TPB)*

	BEH	ECOL-P	ECOL-C	JPO	TPB
Articles	61	116	82	104	44
Units:					
Good (%)	68.9	75.9	75.6	41.3	18.2
Fair (%)	18.0	21.6	22.0	32.7	29.5
Poor (%)	13.1	2.6	2.4	26.0	52.3
% of articles with:					
Scale	42.6	68.1	69.5	97.1	25.0
Theory	32.8	64.7	58.5	95.2	93.2
Data	70.5	69.0	69.5	27.9	4.5
Statistical tests	91.8	95.7	87.8	12.5	2.3

*See text for source journals.

ANOTHER LOOK AT SECTION 1.4

Seabirds are positively associated with fish schools at the scale of 20 km of more. What do you think happens to this association if smaller units (say, 100 m segments of a transect) are used? Why?

1.5 Quantitative Reasoning Using the Principle of Similarity

Units are more than simply a matter of good scientific practice. They are key to quantitative reasoning based on the *principle of similarity*. A simple example of the use of the principle of similarity is to say that:

$$100^2 \text{ meters}^2 \text{ is to } 1^2 \text{ meter}^2 \text{ as a hectare is to a square meter}$$

or, more briefly:

$$(100 \text{ m} : 1 \text{ m})^2 = 1 \text{ hectare} : 1 \text{ m}^2$$

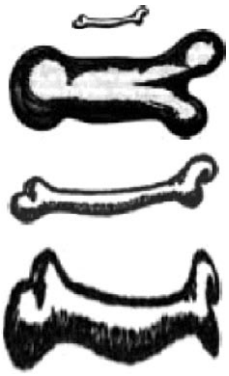
Reasoning is about the similarity between scaled quantities:

$$(100 \text{ m})^2 : 1 \text{ m}^2$$

Reasoning is not about numbers stripped of units:

$$100 : 1$$

Scaling arguments and the principle of similarity have long been used routinely in one area of biology, the allometric scaling of form and function to body size. The origin of allometric studies in biology is clear: D'Arcy Wentworth Thompson's 1917



Galileo began with a small bone, then drew a bone three times longer but wide enough to 'perform the same function which the small bone performs for its small animal'

Here is the small bone, magnified 2.8 times in length and 2.8 times in width. The shape is the same.

Here is the small bone again, now magnified 2.8 times in length and 8 times in width, to match Galileo's drawing. The change in shape is needed to provide the strength to support the larger animal.

FIGURE 1.5 Allometric Scaling of Bone Size and Shape, from Galileo (1638).

treatise *On Growth and Form* (Thompson, 1961). Thompson showed that a style of quantitative reasoning developed by Galileo and Newton leads to a new understanding of the form and functioning of organisms. Galileo reasoned from experience with buildings that the width of bones must increase relative to length, going from small to large animals (Figure 1.5). A large bone with the same proportions as a smaller bone would break under the weight of the animal. Thompson (1917) showed that scaling arguments such as this are widely applicable in biology.

Physiology is a second research area in biology in which scaling arguments are used routinely. Scaling arguments and the principle of similarity are used to put calculation of flows, clearances, and other rates on a sound physical basis. Scaling and similarity are used to check whether equations are dimensionally consistent (Riggs, 1963). An example is the application of the principle of similarity to obtain first-order approximations for a question, such as blood flow through a tube at a known pressure, given a viscosity of, say, 15 times that of water.

Scale arguments are used routinely in engineering research. Before being built, large structures are tested on smaller-scale models that have similar physical properties. Calculations based on the systematic application of the principle of similarity are used to mimic, in the smaller-scale model, the same balance of forces found in the full-scale structure or prototype (Taylor, 1974). The rescaling is not always intuitive. For example, real ice cannot be used to mimic the effects of ice on small models of ship hulls in a tank. A much more brittle material than ice must be used in a tank to mimic the balance of forces of ice on a full-sized ship.

Another example, one that is especially relevant to aquatic ecology, is geophysical fluid dynamics. Rapid progress in atmosphere and ocean sciences occurred when fluid dynamics were taken out of pipes and put into a geophysical grid (Batchelor, 1967; Pedlosky, 1979) with attention to time and space scales (Stommel, 1963). The principle of similarity is routinely used to guide research in the earth, ocean, and atmospheric sciences, from climate change to ocean circulation and the motion of plates floating on the earth's molten interior.

The consistent use of scaling arguments in geophysical fluid dynamics is one of the major differences in research style between this field and ecology. In the 1990 compilation (Table 1.2), one contrast that stands out is that virtually all papers in the *Journal of Physical Oceanography* use spatial and temporal scales, compared to only a quarter of the

articles in a leading ecology journal with similarly high theoretical content (*Theoretical Population Biology*). This is changing as explicit treatment of scale makes its way into ecological journals (Schneider, 2001b). For the most part this explicit treatment remains qualitative (e.g., comparing results from an experiment carried out at three spatial scales). Quantitative treatment (e.g., a ratio of two rates at the spatial scale of the study) can be expected to increase in the ecological literature.

The theory of measurement is another area in which scaling and similarity are routine parts of reasoning (e.g., Stevens, 1946; Ellis 1966; Kyburg, 1984). Inquiry in this area is directed at understanding the basis of measurement employed to understand natural phenomena. The questions asked are such things as, "What kind of measurement units are valid?" This line of inquiry brings out the key difference between mathematical reasoning and reasoning about scaled quantities, which is that the latter employs units and the principle of similarity.

Reasoning about scaled quantities is the route by which "scale" is incorporated into geophysical fluid dynamics, the allometry of body size, engineering research, measurement theory, and physiology. The most important characteristic of quantitative reasoning in this special sense is that it uses the principle of similarity and that it is directed at scaled quantities obtained by measurement or by calculations from measurement. Similarity statements apply to quantities, not to numbers, symbols, or mathematical expressions devoid of units. This point is important because the rules for mathematical work with scaled quantities are not the same as those for working with numbers or equations. This is the apple/orange principle, which says that unlike things cannot be compared: 3 apples cannot be subtracted from 4 oranges. The rules for subtraction allow 3 to be subtracted from 4 or X to be subtracted from Y, but the rules for units do not allow 3 apples to be added to 4 oranges. The apple/orange principle is an important part of reasoning about quantities.

Quantitative reasoning in this special sense differs from the meaning of "quantitative" that has developed in ecology. The meaning that comes to mind is the use of statistical and mathematical techniques, rather than the use of scaled quantities, to reason about ecological processes. To check this impression I searched the University of California library system (8.1 million titles in 2003) for all books with the words *quantitative* and *ecology* in the title. The search turned up seven general works (Darnell, 1971; Greig-Smith, 1983; Kershaw and Looney, 1985; Poole, 1974; Turner and Gardner, 1991; Watt, 1968; Williams, 1964), not counting second and third editions. All seven rely heavily on statistical analysis. Four of the books use a mixture of population biology and statistical methods to make calculations and address questions at the community and population levels (Greig-Smith, 1983; Kershaw and Looney, 1985; Poole, 1974; Watt, 1968). *Scale* appears as a topic in three of the books (Greig-Smith, 1983; Kershaw and Looney, 1985; Poole, 1974) and is a central theme in Turner and Gardner (1991). However, none of these books treat scaled quantities and the principle of similarity as key to quantitative reasoning.

Books that treat scaled quantities and the principle of similarity are confined largely to the topic of organism form and function, building on the tradition of Thompson (1917). These include Huxley (1932), Brody (1945), Vogel (1981), Pedley (1977), Peters (1983), Schmidt-Nielsen (1984), Calder (1984), Alexander (1989), and Pennycuick (1992). Of these, Vogel (1981), Peters (1983), Calder (1984), Schmidt-Nielsen (1984), and Pennycuick (1992) treat ecological topics. Mann and Lazier (1991)

use scaled quantities and the principle of similarity to present basic concepts in marine biology. Legendre and Legendre (1998) present a chapter on units and dimensions in a book devoted largely to statistical methods.

ANOTHER LOOK AT SECTION 1.5

The satirist Jonathan Swift used 1:1 similarity to draw dramatic conclusions about Lilliputians and Brogdingnags in *Gulliver's Travels*. A Brobdingnag, who is as tall as a steeple, sleeps in a bed 20 yards (18 meters) wide and 8 yards (7 meters) off the floor. How does the ratio of Brobdingnag to human height (60 feet/5.5 feet) compare to the ratio for bed width and the ratio for bed height? Discuss the consequences of a fall out of bed for a Brobdingnag who is larger than an elephant but who has bones shaped like those of a human.

1.6 Quantitative Ecology

In the broad sense, the term *quantitative ecology* applies to any mathematical or numerical treatment of the topic, whether or not units are used. For the purposes of this book, *quantitative ecology is defined as the use of scaled quantities in understanding ecological patterns and processes*. This definition arises from two facts and two beliefs. The facts are that scaled quantities are *not* the same as numbers and that the rules for working with quantities are *not* the same as for either equations or numbers. The first belief is that calculations based on reasoning about quantities are useful in solving ecological problems. The second belief is that the scaled part of a quantity (units and dimensions) is just as important as the numerical part in reasoning about ecological processes. The definition rests on the idea that sound research in ecology requires scaled quantities, not numbers or equations devoid of units and dimensions.

"Quantitative ecology" brings to mind thickets of statistics or partial differential equations bristling with Greek symbols. In this book the thickets are replaced by scaled quantities and quantitative reasoning. Equations are presented as ideas that express the linkage of processes measured by scaled quantities. Symbols stand for biologically interpretable quantities. To illustrate, the symbol for the gradient in density of N organisms is $\nabla[N]$. This is read as "the gradient in density of N organisms". A gradient in density is readily visualized. Thus quantitative ecology becomes the study of interesting quantities, such as density gradients. Quantitative ecology in the sense of working with quantities such as animal density, primary production, or gene flow is readily learned because the goal is to make calculations about quantities that are of interest to ecologists. Symbolic expression is emphasized because it is the language of scaled quantities. Symbols (which stand for quantities) and equations (which express ideas about how quantities are related) allow us to make calculations about the biological phenomena that interest us.

This book assumes that verbal and graphical treatment of a concept is the road to understanding equations, which are necessary to make calculations about any quantity of interest. This book takes the position that data are important but must be summarized (refer back to Figure 1.1) in verbal models (sentences), graphical models, and

ultimately, formal models (equations). The three lines radiating outward from data in Figure 1.1 represent summarization skills: the act of expressing data in the form of sentences, graphs, or equations. Each of the three peripheral lines represents translation skills: reading graphs and equations into words, drawing graphs from words or from equations, and writing equations from words or from graphs (refer back to Figures 1.2 and 1.3). All six skills are used in this book. Emphasis is on verbal treatment of concepts that are then developed as symbolic expressions of ecological ideas about the relation of one scaled quantity to another.

If I use a symbol or write an equation without stating it in words (which I try not to do), I ask the reader to translate the symbols into words. If I use an equation without drawing it (which I often do), I ask the reader to sketch the equation. These acts of translation will help considerably in understanding the ideas expressed in the terse form of symbolic notation. Facility in translation, which comes with practice, will help the reader considerably in his or her own research. My hope is that ecologists will become as effective as physiologists or oceanographers in the use of scaled quantities in their research. I also hope that many ecologists, not just a small number of "modelers," will become able to express their ideas about scaled quantities in the form of equations.

ANOTHER LOOK AT SECTION 1.6

Estimate or count the number of courses taken during your undergraduate career, then estimate the number that used equations to express ideas.

1.7 Tools

In ecology, it has been traditional to annex equations to appendices. The end of this tradition is now in sight as introductory ecology texts (e.g., Ricklefs and Miller, 2000) begin to use calculations based on equations that express important ideas in ecology. To encourage this trend, I have tried to reduce the math in this book to its essentials and to explain it in vivid and sufficient detail along the way. Opening the book at random will, more likely than not, turn up symbols and equations. The equations are merely the ghostly outlines of ideas that are explained in the text. The equations are present because they are necessary for making calculations from the ideas. In fact, the best way to understand the book from Chapter 4 onward, and to understand scale and scaling theory, is by doing the calculations represented by the equations.

What tools are needed for learning quantitative ecology? For the material in this book (except the material in Chapter 15), a handheld calculator or spreadsheet will suffice. The calculator must be able to take logarithms and raise a number to any power. A graphing calculator allows quantitative expressions to be graphed. An alternative to a graphing calculator is a spreadsheet program, using the function commands in combination with the graphics commands. Spreadsheets are easy to use, but they have the disadvantage that they do not display equations in standard symbolic format. Like most computer applications, they use line-code format. To illustrate the difference, here is the equation expressing the relation between the perimeter of an object and the length of the

step used to measure the perimeter. The equation is then written in line-code format and spreadsheet format:

$$\text{Symbolic:} \qquad \frac{Perim}{Perim_o} = \left(\frac{L}{L_o}\right)^{1-\beta}$$

$$\text{Line code:} \qquad \text{Perim} = (\text{Perim\_}o)*(\text{L/L\_o})**(1 - \text{beta})$$

$$\text{Spreadsheet:} = \$A\$1*(B1/\$a\$2)\text{^}(1 - \$a\$3)$$

The symbolic format matches what is used in textbooks. The relation of quantities can be read. Such notation is essential for comparing one equation to another or understanding how one expression is related to another. The relation of one quantity to another is less easily read from the line-code format, which a computer needs to undertake calculations based on the equation. The relation of one quantity to another is all but impossible to read from the line-code and spreadsheet formats. Spreadsheets, like all statistical packages and most mathematical packages, are restricted to line-code format. There are, however, several programs (MathCad, MatLab, and Mathematica) that allow calculation from symbolic format. Perhaps the easiest to use is *MathCad*, which displays symbolic notation that is easily edited. It translates the symbolic notation into line code (which the user does not see) to make calculations. MathCad is a good learning tool because the user interacts directly with the symbolic code and does not need to learn line-code conventions. *MatLab* works with line code but allows a window that displays symbolic code to be opened. Accordingly, the user must learn the conventions of line code to work with MatLab. *Mathematica* is organized around line code but allows input and editing of either form. Student purchase rates apply for all three mathematical packages. In deciding which tool to use with this book, the reader will need to consider past experience and whether it is worth investing time and money into one of these tools rather than a handheld calculator or a spreadsheet, with its attendant clumsiness.

1.8 Overview of Chapter Concepts and Sequence

Part I of this book presents introductory material consisting of the present chapter and the next chapter, which traces the increasing use of scale in ecology and then introduces the central scaling concepts in this book.

Part II covers scaled quantities and their use in models and quantitative reasoning. Chapter 3 defines a *quantity*, a concept that is key to the integration of scaling into ecology. The concepts of *zooming* (sequential changes in attention) and *panning* (roving viewpoint) are introduced in this chapter. Quantities on a ratio type of scale have *units*, described in Chapter 4. The rules for working with *scaled quantities*, listed in Table 4.4 and demonstrated in Box 4.2, differ from those for working with numbers. Chapter 5 describes logical, rigid, and elastic rescaling of quantities.

Chapter 6 introduces the concept of *dimension*, or similarity groups. Chapters 7 through 10 are a digression into spatially and temporally variable quantities. Chapter 7 develops the premise that ecological quantities have spatial and temporal *attributes*—their chronology, duration, location, and extent. The idea that quantitative reasoning requires clear and accurate notation makes its appearance in this chapter. Chapter 8 explores complex quantities derived from sequential measurements in space and time—fluxes, gradients,

divergences, and their relatives. The theme of *notational clarity* continues in this chapter; the themes of panning and zooming reappear. Chapter 9 introduces a new theme, that of *ensemble quantities*, formed either by juxtaposition or by superposition. Sums, weighted sums, and means are interpreted as scaled quantities rather than mathematical operations. Chapter 10 extends the theme by considering deviances and variances, again as scaled quantities subject to increase and loss rather than as operations on numbers.

Part III develops the concept of *scope*, which is the door to scaling theory and the principle of similarity. Chapter 11 applies the concept to measurements, instruments, quantities, and equation. The chapter describes the several types of scope diagrams found in the literature. Chapter 12 extends the concept of scope to *diagrammatic analysis* of research programs.

Part IV brings together the material on quantities (Part II) and scope (Part III). Chapter 13 takes up the relation of one quantity to another, treating equations as a way of making calculations based on ideas. Chapter 14 briefly treats *derivatives* from the point of view of scaled quantities rather than as a mathematical exercise. The examples and exercises in Chapters 13 and 14 are aimed at the development of skill in translation among verbal, graphical, and formal models expressed as equations. Chapter 15 introduces the theme of *data equations* and the *general linear model*, which form a bridge between statistical analysis (which will be familiar to most ecologists) and dimensional analysis (which will not). Chapter 16 presents a completely personal view of the future of spatial and temporal scaling in ecology.

The material in this book can be used in an upper-level undergraduate or first-year graduate-level course. Algebra and a knowledge of logarithms are required; some acquaintance with calculus will help, but that is not a prerequisite. A few sections of Chapter 13 assume some familiarity with calculus. Chapter 15 introduces statistical concepts at a level that is successfully grasped by fourth-year undergraduates who have taken only one course in statistics. Chapter 1 and the chapters in Part II form an introduction to quantitative ecology, defined as reasoning about ecological quantities, rather than as a collection of mathematical techniques. I routinely use material from Chapters 1, 3, 4, 5, 13, and 15 in a course in model-based statistics for fourth-year and beginning graduate students. These chapters stand apart, to some degree, from the particular occasion of this book, which was to develop a guide to spatial and temporal scaling in ecology. That guide is introduced in Chapter 2, then developed in Parts III and IV. It is a guide based on concept, not a collection of recipes or techniques that currently prevail in the literature. It is based on working with honors and graduate students in applying scaling concepts in their research.

1.9 Exercises

Exercises from the first edition have been updated and new exercises have been added. These can be found online at www.elsevierdirect.com/companions/9780126278651.

2 Scale in Ecology

We can no longer … cling to the belief that the scale on which we view systems does not affect what we see. … This is quite a different way of viewing the world than that which was in vogue a decade ago.

—J. Wiens, *Landscape Ecological Analysis*, 1999

2.1 The Problem of Scale in Ecology

Ecological concepts can be found in the writings of Aristotle, but the word "ecology" is recent, coined by Ernest Haeckel in 1869 from the Greek words for "house" and "study". Scaling concepts appear early in the 20th century (Johnstone, 1908; Mercer and Hall, 1911), at about the same time that ecology become a recognizably distinct field of biology. The use of the word "scale" in ecology is more recent, dating from the late 1970s in biological oceanography (Smith, 1978; Steele, 1978) and from the 1980s in terrestrial ecology (Starr and Allen, 1982). Recognition of the problem of scale in ecology became widespread around 1990 (Wiens, 1989; Steele, 1991a; Levin, 1992).

The problem of scale has three components. The first is that pressing problems in ecology are often at the scale of decades and large ecosystems. An example is habitat loss, which reduces population viability when it is chronic (at the scale of generation times) and widespread (at the scale of populations). The second component is that measurements are usually at scales far smaller than pressing problems. Most variables, and in particular most rates, can only be measured in limited areas during brief periods. Remote sensing allows a few variables, such as ocean color, to be measured over large areas at fine resolution, but only during brief periods. Automated devices allow variables such as temperature or sea level to be recorded continuously over long periods at high resolution, but only in small areas. The third component of the problem of scale is that patterns and process prevailing at small scales do not necessarily prevail at larger scales. Habitat loss, for example, alters animal distribution and movement at local scales, whereas at larger scales it can lead to species extinction. When pattern and process depend on spatial or temporal scale, variables at the scale of decades and ecosystems cannot necessarily be computed directly from local measurement, which includes almost all experimental manipulations.

The classic example of the problem of scale is loss of biodiversity. The causes of species extinction occur at the scale of ecosystems, whereas measurements of species

doi: 10.1016/B978-0-12-627865-1.00002-7

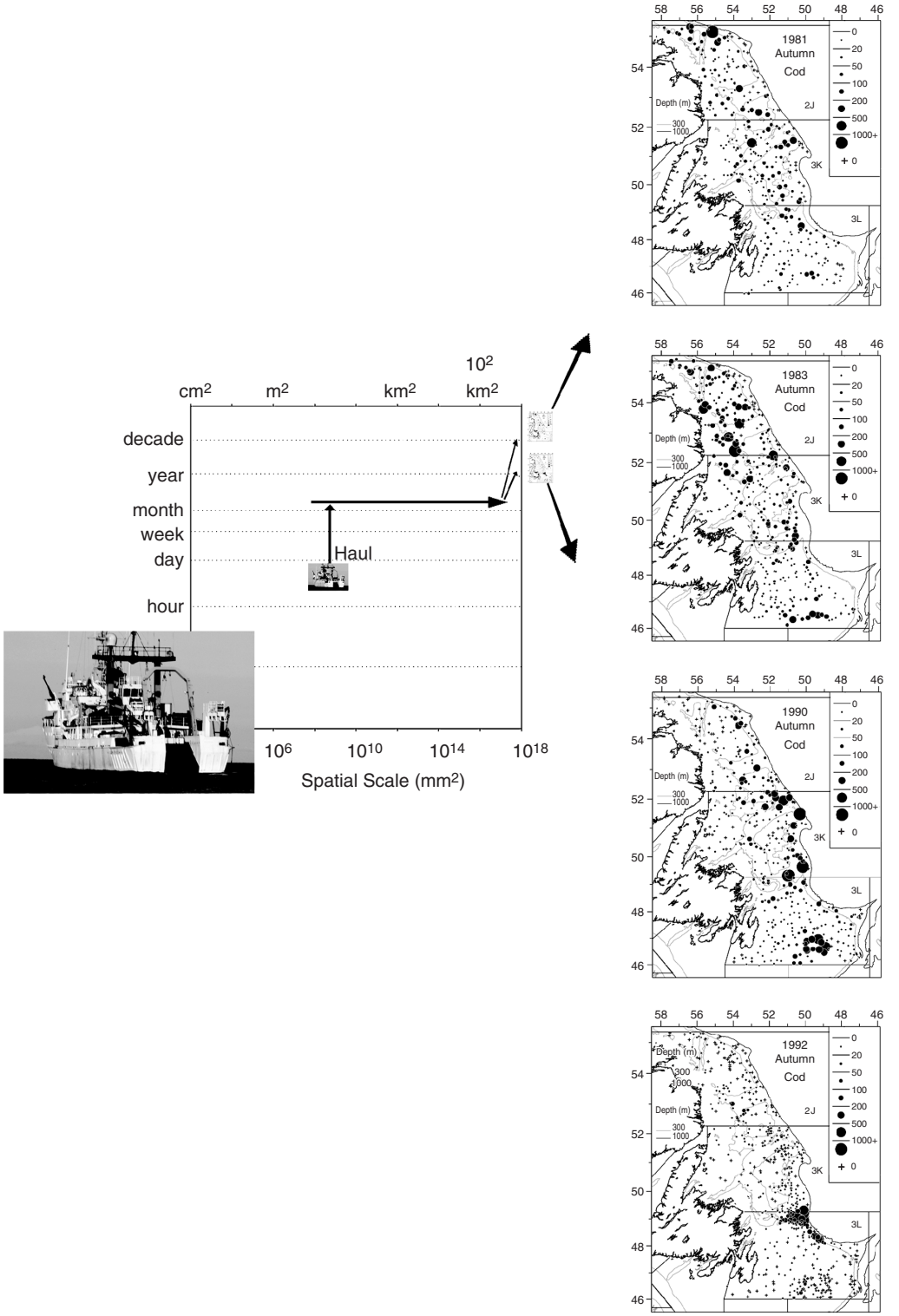


FIGURE 2.1 Scale-up in a Fisheries Survey. Sample consists of 900 hauls covering 22 square kilometers over 200,000 square kilometers of ocean. Maps courtesy of G. Lilly, Department of Fisheries and Oceans, St. John's, Canada.

number are of necessity confined to smaller areas. Yet no biologist would extrapolate directly from local samples to calculate extinction rate at a larger scale because it is well known that species number does not scale directly with area. It has become increasingly clear that the same principle applies to any ecological problem. Another example is tracking change in stock size of a fishery such as cod (*Gadus morhua*) in the northwest Atlantic (Figure 2.1).

The cod fishery brought European settlement to North America, from the Vikings and Basques to John Cabot and the Pilgrims of Cape Cod (Kurlansky, 1997). The fishery extended along thousands of kilometers of coast for hundreds of years, spreading out from the coast and, by 1900 extending across the entire continental shelf. After World War II technological innovation allowed fishing vessels to pursue cod into increasingly inaccessible spawning grounds, until the last large fishery off Newfoundland collapsed in 1991. On the second day of July 1992, this centuries-old fishery was closed, putting 30,000 people out of work.

During the decade before the closure, changes in stock size were tracked by calibrating catch statistics against annual research surveys. Commercial catch statistics cannot be used by themselves to measure changes in stock size for a variety of reasons–one of the most important is a matter of scale. Catch rates remained locally high, even as the stock disappeared from most of its formerly larger range (Figure 2.1), because vessels converged on local aggregations. Systematic surveys address this problem by random sampling from the entire stock area, resulting in accurate estimates in the long run. But in the short run, estimates are made from samples that cover a relatively small area and thus remain vulnerable to the problem of scale. In any one year a completely unbiased survey can be expected to miss large aggregations (leading to an underestimate in that year at the scale of the stock), or it can hit several large aggregations (leading to an overestimate that year). The latter happened in 1995 off the south coast of Newfoundland, when catch from just one tow (out of 161) accounted for 87% of the biomass of fish caught during the entire survey (Brattey et al., 1999). This single tow increased the estimate at the scale of the stock by nearly eightfold.

Figure 2.2 shows in diagrammatic form the three components of the problem of scale for the problem of monitoring pollution induced variation in benthic invertebrate numbers against a background of natural variation in Manukau Harbor, New Zealand (Thrush et al., 1997). Pollutants are dumped at point sources (lower left part of Figure 2.2a). Continued dumping and tidal mixing extended the problem to larger scales (upper right corner, Figure 2.2a). Surveys to monitor changes in benthic populations are confined by cost to a limited number of samples, each covering a small area. Experiments to separate effects of pollution from naturally occurring changes are similarly confined by cost to areas far smaller than the scale of the problem (Figure 2.2b). At these smaller scales animal movement exceeds mortality (Figure 2.2c), and hence mortality due to pollution is difficult to detect, because it is swamped by the movement of animals across the boundaries of survey or experimental sites.

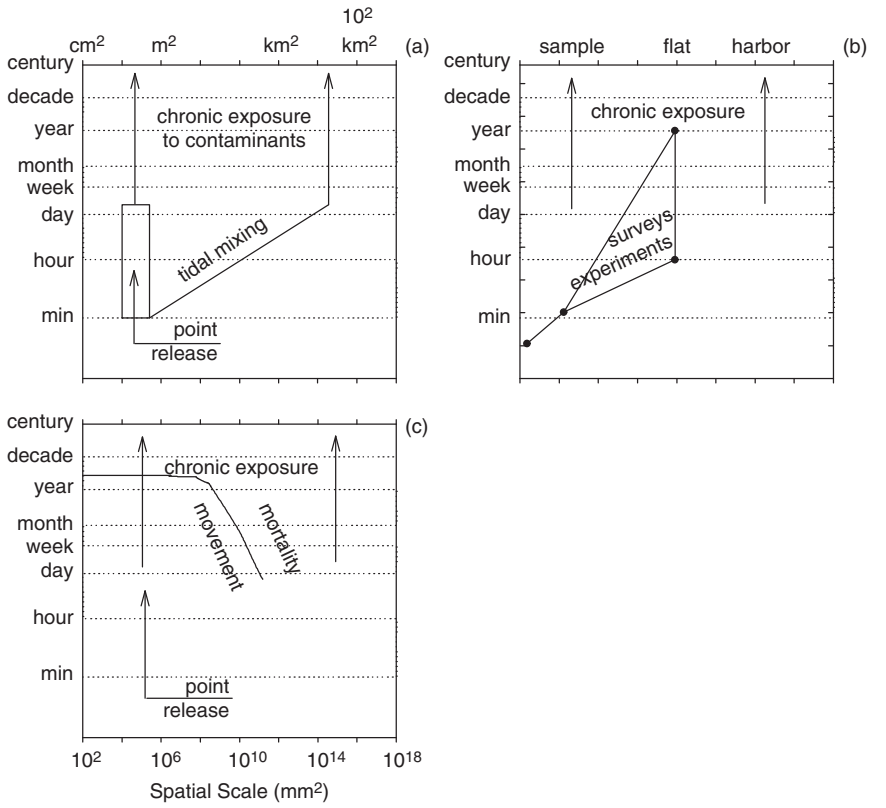


FIGURE 2.2 Space-time Diagrams for the Problem of Monitoring Effects of Chronic Release of Contaminants in Manukau Harbour, New Zealand. Redrawn from Schneider *et al.* (1997). (a) Spatial and temporal scale of the problem of point release of contaminants; (b) Space and time scales of survey and experiments addressing the problem; (c) Space and time scales of prevailing dynamics.

ANOTHER LOOK AT SECTION 2.1

State in words an ecological problem of interest to you, then sketch a space-time diagram showing the space and time scales of the problem and of the measurements that can be made.

2.2 Definition of Scale

The word *scale* has multiple meanings, contributing to its ambiguous usage in the ecological literature. The Oxford English Dictionary distinguishes 15 different meanings arising from two different roots. The Old Norse root in *skal,* or bowl, gives rise to fish scales, the scales of justice, and hence, by extension, measurement via pairwise comparison of objects. Using an old-fashioned scale, one can assign a mass of 110 grams to a bird that balances standard masses, one of 100 grams and the other of 10 grams. The Latin root in *scala*, or ladder, gives rise to musical scales, scaling a wall, and, by extension, measuring distances by counting steps or subdivisions. Using a ladder known to be 2 meters in height, one can assign a height of 1/2 ladder (1 meter) to a small tree.

In ecology the word *scale* has acquired several common technical meanings (Box 2.1). This diversity in technical definition, added to the diversity in origin of the word, works against attempts at standard definition. Instead, the word needs to be defined for the situation at hand and then used consistently, usually with an appropriate qualifier (Schneider, 1994b; Withers and Meentemeyer 1999).

Box 2.1 Common Technical Definitions of Scale

The *type of measurement scale* (Stevens, 1946) distinguishes variables quantified on a nominal scale (presence/absence), ordinal scale (ranks), interval scale (equal steps, such as degrees Centigrade), and ratio scale (equal steps and known zero, such as degrees Kelvin).

Cartographic scale is the ratio of the distance on a map to the distance on the ground. A world map of about a meter in width will have a scale of 1:39,000,000.

Ecological scaling (Calder, 1983; Peters, 1983) refers to power laws that scale a variable (e.g., respiration) to body size, often according to a nonintegral exponent. Respiration typically scales as $\text{Mass}^{0.75}$ and hence a doubling in body size increases oxygen consumption by $2^{0.75} = 1.7$ rather than by a factor of 2.

The *spatial scale of a measured variable* refers to the spatial resolution relative to the spatial extent (Wiens, 1989; Schneider, 1994b). Similarly, the *temporal scale of a measured variable* refers to the temporal resolution relative to the temporal extent. Measured variables have a minimum resolvable area or time period (resolution, grain, or inner scale) within some range of measurement (range, extent, or outer scale). For example, a tree-coring device measures changes in growth at a resolution of one year over a range of thousands of years. This definition can be extended to the *scale of a measured variable*, defined as the resolution relative to the largest value. For example, body mass can be measured to the nearest gram over the range of sizes, such as from mice to elephants.

In *multiscale analysis* the variance in a measured quantity or the association of two measured quantities is computed at a series of different scales. This is accomplished by systematically changing either the range or the resolution. The resolution is changed by increasing the separation (lag) between measurements or by increasing the averaging interval (window size) for contiguous measurements (Platt and Denman, 1975; Schneider, 1994b; Milne, 1997).

Powell (1989) defines scale as the distance before some quantity of interest changes. An example is the scale of a hurricane, defined as the distance across the weather system. This use of the term refers to the *scale of natural phenomena*.

Of these definitions, one of the most comprehensive is that *scale denotes the resolution within the range or extent of a measured quantity.* This definition of scale can be applied to the space and time components of any quantity. For example, an investigation of litter fall to the forest floor might have a spatial scale resolved to the area occupied by a single tree, within the spatial extent of a 100 hectare study area. The temporal scale of the study might be resolved to weekly measurements over a temporal extent of two years. This definition can be extended to any variable, regardless of its spatial or temporal resolution and range.

For the leaf litter example, the scale of the measured variable the operational definition might be the precision of gravimetric measurement of litter accumulation beneath a single tree during a week, within a range set by the total accumulation of mass in the study area.

Both range and resolution can be altered. In the example of the investigation of litter fall, we could increase the spatial scale (extent) by enlarging the study area beyond 100 hectares. We could also increase the scale (resolution) by sampling at a finer scale under each tree. The temporal scale could be increased by extending the range beyond two years or by making daily rather than weekly measurement. The mass scale could be altered from large units to units small enough to resolve the fall rate of components (branches, leaves, bark).

Systematically altering either the range or the resolution leads to *multiscale analysis*, which tells us more than analysis at a fixed scale. An early example of multiscale analysis comes from Greig-Smith (1952), who showed that plants are more clustered at some spatial scales than others. Trees may occur in small clusters scattered over the landscape, yet within each cluster trees may be spaced in a relatively regular way due to competition. This intuitive idea of pattern can be quantified as a series of means describing contrasts in density from place to place. Other statistical measures can be used, such as a variance to measure the strength of contrast among several means. Whatever the statistic chosen, it will vary with the range and resolution at which examination takes place. *Scale-dependent patterns* can be defined as a changes in some measure of pattern with change in the range (extent) or resolution (separation or unit size) of measurement.

Scale-dependent processes, like scale-dependent patterns, can be defined relative to change in range or resolution of measurement. However, we encounter problems if we simply say that a process depends on the range and resolution of measurement. A physical process such as gravity acts at any spatial scale we care to consider. Similarly, a biological process such as mutation operates at time scales from seconds to millennia. We cannot say that a process is restricted to any particular scale. But we can point to specific time and space scales at which one process prevails over another. For example, rates of change in animal density depend on movement if we consider small areas, but demographic rates (births, death) prevail over rates of immigration and emigration across the boundaries of larger areas. In the ocean, viscous forces prevail over gravitational forces at small spatial scales typical of bacterial sizes, whereas gravitational forces become prominent for animals large enough for us to see. A convenient definition is that *processes are scale dependent* if the ratio of one rate to another varies with the range (extent) or resolution (separation or unit size) of measurement. An example is shading of the forest understory. This process occurs on a regular seasonal cycle at mid-latitudes, it varies annually, and in forests it can change suddenly at any point due to events such as fires or deadfalls. So we cannot say that shading depends on time scale. But we can measure the process of shading as a ratio of two rates: the flux of light to the forest floor relative to the flux of light to the earth. This ratio will depend on temporal and spatial scale, with important effects on growth rates of plants.

ANOTHER LOOK AT SECTION 2.2

Name three variables of interest to you. For each, state the temporal resolution, measured as the time it takes to complete a single measurement. State the temporal extent, defined as the typical duration of a study in which the variable occurs.

2.3 The Rise of the Concept of Scale in Ecology

The idea that scale matters might now seem obvious, but in fact it is a relatively recent concept that is still evolving. Thirty years ago, the idea that scale matters went largely unrecognized in ecology. Of 28 plenary contributions covering the major ideas in ecology at the First International Congress of Ecology (van Dobben and Lowe-McConnell, 1975), none used the word "scale" or stated that scale matters. Most ecology texts still do not present the concept of scale. The first major text to do so was Ricklefs and Miller (2000), in which scale is a key concept listed in the first chapter.

Did the appearance of scaling concepts in the research literature occur suddenly as a "paradigm shift"?

> *Kuhn's concept of a paradigm shift is a useful way to interpret the annual meeting [1988] of the E[cological] S[ociety] [of] A[merica]. … Every symposium or session I attended featured, included, or was structured by the concepts of scale and spatial patterns. I left feeling I had observed one of those rare creatures of the intellectual bestiary, a paradigm shift.*
>
> —*F. Golley, in Landscape Ecology 3:65, 1989.*

There is little question that the explicit use of the concept of scale appeared suddenly and grew rapidly. In a digital version of the journal *Ecology* the term "spatial scale" makes its first appearance in the early 1970s (Marten, 1972; Wiens, 1973). Frequency of appearance grew exponentially in the 1980s at about 10 times the rate of growth in number of articles per annual volume (Figure 2.3). During the 1980s the term usually appears in this journal to qualify a result as being found at a particular spatial scale. Since then use of the term "spatial scale" has continued to grow more rapidly than the number of articles per year in this journal. Use of the term "temporal scale" first appeared in this journal in the late 1970s (Erwin, 1977) and grew rapidly in late 1980s and early 1990s. Use of both terms has increased in *Ecology* into the early 21st century, at a higher rate for spatial than temporal scale.

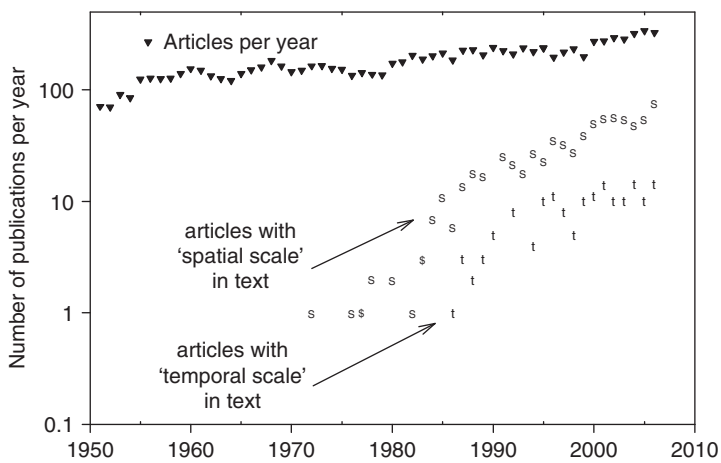
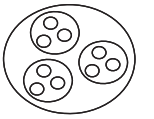


FIGURE 2.3 Verbal Expression of the Concept of Scale in the Journal *Ecology.*

A similarly rapid increase in treatment of the concept of scale during the 20th century occurred within a specialized and relatively old research area, the pelagic ecology of seabirds. In a comprehensive reference list on the topic (Schneider, 1991), the two oldest articles were Collins (1884) and Murphy (1914). The number of articles increases at a rate of 5.5% year^{-1} from 1946 to 1980, then levels off. Explicit use of the concept of scale first appears in 1980, followed by exponential increase at 18.9% yr^{-1} from 1980 to 1990 (Schneider, 2002).

In terrestrial ecology, hierarchically nested levels have been used to express the concept that "scale matters" (Box 2.2). Publication frequency on hierarchy concepts thus reflects the rise in verbal treatment of the concept of scale. In a comprehensive list of 231 articles compiled by Jurek Kolasa (McMaster University, Hamilton, Ontario, Canada), the publication rate grew exponentially from 1969 to 1990 (Schneider, 2001b). If the 63 nonecological publications in the list are removed, publication rates grew exponentially from 1973 through 1991 at 18.2% yr^{-1}.

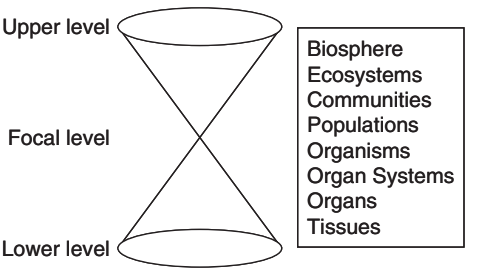
Box 2.2 The Concept of Hierarchy in Ecology.



The concept of nested organizational levels (cell, tissue, organ, organism) has a long and productive history in biology.

Egler (1942) used a nested icon to represent the idea of species nested within genera within families, atoms nested within molecules within cells, and cells within tissues within organisms within individuals.

E.P. Odum (1959) extended the concept of level to populations and communities



Allen and Starr (1982) combined the concept of nested hierarchical level with the idea that large scale dynamics (Upper level) constrain local dynamics (Levin 1976, Levandowsky and White 1977) at a focal level.

O'Neill *et al.* (1986) proposed that focal levels beyond that of the organism arise from disparate rates of energy dissipation in ecosystems.

Allen and Hoekstra (1992) and O'Neill and King (1998) distinguished the qualitative concept of level from the quantitative concept of scale.

Jagers (2008) distinguished at least 4 criteria for defining hierarchies, noting that spatial and temporal scale are generally not used. Liddicker (2008) noted that hierarchically defined organizational complexity is at best weakly related to spatial scale.

Increasing use of spatial scale reflects the rapid growth in recognition of the concept in the late 20th century, but the concept of scale in ecology is far older. Early in the century, Johnstone (1908) questioned whether catches from a limited area could be used to calculate the density of organisms across the entire Irish Sea. At about the same time, Mercer and Hall (1911) recognized that experimental results from small areas differ from those in larger areas. Greig-Smith (1952) quantified change in plant patchiness at multiple scales, showing that patchiness and strength of association with soil characteristics depend on spatial scale.

The rise in recognition of the concept of scale depends on the way the search is defined. Sudden appearance followed by exponential growth was found when a particular phrase (spatial scale, temporal scale) or concept (hierarchy) was considered. The first edition of this book aimed at comprehensive coverage of the concept of scale in ecology, regardless of whether the word "scale" was used. In a reference list of 237 articles that considered more than one spatial or temporal scale (Schneider, 1994b), the publication rate per year showed sustained growth throughout the 20th century, interrupted only by World War II. The number of articles grew exponentially at 7.5% yr^{-1} from 1950 to 1990 (Schneider, 2001b).

Associated with this rapid growth in explicit use of the concept of scale (Figure 2.3) was the appearance of classificatory schemes (Box 2.3) in terrestrial ecology (Delcourt and Delcourt, 1988), marine ecology (Haury et al., 1978; Steele, 1991a), and evolutionary biology (Whittaker, 1960, 1977). One notable feature of the terrestrial versus marine classification scheme is that, at any given spatial scale, the time scales are longer in the terrestrial than the marine scheme. This reflects the persistence of habitat features on land compared to the fluid environment of the ocean (Steele, 1991b). In an evolutionary context, Whittaker (1977) defined *inventory diversity* at four spatial scales or levels: the point sample, the habitat, the landscape, and the region. *Differentiation diversity* is defined as the change in diversity between these four levels (*cf.* Box 2.3). *Point diversity* is the number of species in a small or microhabitat sample within a community regarded as homogeneous. It is also called *internal alpha* or *subsample diversity. Pattern diversity* is a differentiation diversity: the change going from one point to another within a habitat. It is also called internal *beta (β) diversity. Alpha (α) diversity* is the within-habitat diversity: the number of species in a sample representing a community regarded as homogeneous (despite its internal pattern). It measures the number of potentially interacting species. *Beta (β) diversity* is the between-habitat diversity differentiation: the change along an environmental gradient or among the different communities of a landscape. It was developed by Whittaker (1960) as a measure of packing competing species along a gradient; it was defined as the ratio of regional (gamma) diversity to average within-habitat (alpha) diversity. Wilson and Shmida (1984) review other measures of beta diversity. *Gamma (γ) diversity* is the landscape diversity: the number of species in a set of samples including more than one kind of community. *Delta (δ) diversity* is geographic differentiation diversity: the change along a climatic gradient or between geographic areas. *Epsilon (ε) diversity* is the regional diversity: the number of species in a broad geographic area including differing landscapes. The distinction between inventory and differentiation diversity is often lost in the literature, which tends to use the term *beta diversity* for what Whittaker meant by gamma diversity. Whittaker

et al. (2001) highlight the distinction between inventory and differentiation diversity but advocate replacing the terms *alpha, beta,* and *gamma diversity* with the terms *local, landscape*, and *macro-scale*.

Box 2.3 Classification of Space and Time Scales in Evolutionary Biology, Terrestrial Ecology, and Marine Ecology

	Evolutionary Biology		***Terrestrial Ecology***		***Marine Ecology***	
	Whittaker, 1977, Figure 14		**Delcourt and Delcourt, 1988**		**Haury et al., 1978**	**Steele, 1991a**
-10^{15}	Area		Area	Time	Area	Time
–						
–			Mega	10^6 + yr	(Mega Macro)	
-10^{12}						
–			Macro	10^4–10^6 yr	Meso	1–10 yr
–						
-10^9	ϵ diversity					
–			Meso	500–10,000 yr	Coarse	days–yr
–		δ diversity				
-10^6	γ diversity					
–						
–						
-10^3			Micro	1–500 yr	Fine	1–10 days
–		β diversity				
–	α diversity					
$1\,m^2$		Pattern				
–	point					
–					Micro	
-10^{-3}						

The growing recognition of the importance of scale in the 1980s accompanied major changes in the amount of data generated by electronic devices. Examples include high-resolution satellite imagery, acoustic surveys in the ocean, and digital recording of environmental data such as temperature. Digital devices typically generate continuous records that replace point measurements in space and time. This has been accompanied by rapid evolution of software that was made possible by exponential increases in cheap computing capacity over the last two decades. The recognition of the problem of scale (refer back to Figures 2.1 and 2.2) and the growth of scaling concepts (Figure 2.3) have been propelled as much as anything by technological innovation.

In the literature prior to 1980 (Schneider, 1994b), the concept of scale is uncommon, it is highly heterogeneous, and there is little evidence that one article influenced

another. As a group, though, these early publications introduce key concepts: the very different environments inhabited by large, medium-sized, and very small aquatic organisms (Hutchinson, 1971); concomitant effects of "slow" and "fast" processes (Levandowsky and White, 1977); the practical problem of choice of space and time scales in survey design (Wiens, 1976; Smith, 1978); the role of fluid dynamics in generating variability in marine organisms over a range of time and space scales (Haury, McGowan, and Weibe, 1978; Steele, 1978); linkage of time and space scales in paleontology (Valentine, 1973) and terrestrial ecology (Shugart, 1978); and multiscale analysis (Greig-Smith, 1952; Platt and Denman, 1975).

The rapid recognition of the importance of scale in the 1980s (Figure 2.3) was followed by a proliferation of concepts. Peterson and Parker (1998; Table 22.1) list 78 scaling concepts expressed as phrases, each with one or more sources, nearly all from 1989 onward. These authors grouped scaling concepts into four categories: scale-dependent patterns, scale-dependent processes, multiscale analysis, and a miscellaneous collection of applications and concepts.

Several distinguishable themes appear in the 20th-century ecological literature. One theme, now widely recognized, is that patterns depend on the spatial and temporal scale of analysis (Piatt and Schneider, 1986; Wiens, 1989; Hengeveld, 1990). This idea can be traced to the early part of the century (Johnstone, 1908). It was quantified by Greig-Smith (1952), who showed that plant patchiness and association with soil characteristics depended on the size of the area of the unit of analysis. Application of the concept grew rapidly in the late 1980s (Meentemeyer and Box, 1987; Menge and Olson, 1990; Rastetter et al., 1992).

A second recurrent theme appearing in the latter half of the 20th century is that of a "characteristic scale" at which best to study a particular pattern and its dynamics. This concept can be traced to geophysical fluid dynamics, where the scale of a physical feature (e.g., the width of the coastal upwelling strip) can be calculated from the balance of forces that generate the phenomenon. In the case of upwelling, the balance between buoyancy forces (causing lighter water to remain on top of denser water) and the Coriolis force (causing rotation of a water mass once it is in motion relative to the earth) determines the width of the strip in which upwelling will occur along a coast. Weak Coriolis forces (as one approaches the equator) or strong vertical contrasts in buoyancy (as in tropical waters heated by the sun) widen the band of upwelling water. Knowing the vertical buoyancy difference and the latitude, one can calculate the characteristic scale of the band of upwelling water along the coast. The same approach has been less successful in biology, something noted by Lauren Haury (from Scripps Oceanography Institute) in the early 1990s. In biology the classic application was the calculation of the "characteristic scale" of phytoplankton patches. Skellam (1951) and Kierstead and Slobodkin (1953) calculated the size of patches due to the opposing effects of cell division (which amplifies existing patchiness) and eddy diffusion (which disperses patches by folding and stretching the fluid). Verification of this "characteristic patch size" was not uniformly successful (Harris, 1980), even after the addition of spatially variable (random) growth rates (Bennett and Denman, 1985). The lack of success may be due in part to the interaction of biological with physical processes. In the example of phytoplankton patch size, density-dependent growth rate may depend on eddy diffusivity so that growth balances eddy diffusivity at a shifting rather than fixed scale. When biological

process interact with physical processes, as is often the case, a shifting rather than "characteristic" scale can be expected (Levin, 1992). When fast processes interact with slow processes, the result is a smearing of outcome over multiple scales. An example of this interaction of the fast with the slow is the episodic collapse of trees, which releases the potential energy stored above the forest floor by slow growth of massive trunks. Episodic collapse tears open the canopy, leaving holes that range widely in size rather than being confined to a characteristic scale (Sole and Manrubia, 1995). Other examples are earthquake magnitudes (Bak and Chen, 1995) and epidemic frequencies (Sugihara, Grenfell, and May, 1990). These phenomena range in size and frequency. They lack a characteristic time or space scale.

The absence of any single "right" scale at which to investigate a population or community (Levin, 1992) forces environmental biologists to adopt a multiscale approach. This stands in contrast to the characteristic space and time scales of many of the genetic, behavioral, and physiological processes investigated by organismal biologists. For these biologists, rates set by the mass of the organism make the choice of scale obvious. One can, for example, use body size to make reasonable calculations of running speeds in long-extinct species that can never be directly clocked (Alexander, 1989). In contrast, there is no single obvious scale at which to compute the interaction between populations of organisms. Similar problems have attended other efforts to use ecological data to identify a "characteristic" scale at which to carry out ecological research.

A third recurrent theme, antithetical to the idea of a characteristic scale that is "best" for research, is that of fractal geometry to quantify environmental features from streams to coastlines (Mandelbrot, 1977). Streams and coastlines have a geometric dimension somewhere between 1 (a line) and 2 (a plane). The fractal dimension, lying between the Euclidean dimensions of 1 and 2, expresses the degree of convolution. Early applications (Burrough, 1981; Frontier, 1987; Sugihara and May, 1990; Williamson and Lawton, 1991) were followed by a comprehensive treatment (Hastings and Sugihara, 1993). The concept of a fractal can be extended to dynamics by considering the change in fractal dimension over time. An example is the sea surface, which can be described as a two-dimensional object only on a windless day. As the wind strengthens, small waves build on larger waves, generating an increasingly complex structure that can no longer be described as a two-dimensional Euclidean plane. This more convoluted surface generated by the wind can be described as a fractal. That is, we adopt the idea that the sea surface has a dimension greater than a plane of Euclidean dimension two (Length^2), though less than a volume of Euclidean dimension three (Length^3). This geometry of fractal rather than integral dimensions is at first strange, but it becomes familiar through practice in viewing an object at more than one spatial scale. Examples of fractal objects range from root and branch systems of trees to habitat boundaries on land, land features such as archipelagos, or fluid phenomena such as the nesting of eddies within eddies in the atmosphere and the hydrosphere. The processes that generate these fractal shapes can themselves be viewed as fractal (West and Schlesinger, 1990). Fractal processes operate over a range of scales; they have non-Euclidean dimensions; they generate convoluted structure within convolutions; they exhibit episodic events within longer-term episodes; and these episodic events account for a disproportionate fraction of the total activity. Examples of fractal rates include rainfall (Lovejoy and Schertzer, 1986, 1991),

measles infection (Sugihara, Grenfell, and May, 1990), the frequency of turns by moving individuals (Frontier, 1987), and the flight speed of birds. Because of the disproportionate concentration of activity into episodes, one hallmark of a fractal process is extreme variability (Lovejoy and Schertzer, 1991).

A related theme in the late 20th century was that spatial variability, or patchiness, in natural populations is a dynamically interesting quantity rather than a statistical nuisance to be overcome (e.g., Thrush, 1991). An example of focusing on variance as a dynamical quantity is to consider the patch generation by tree falls in a forest (Shugart, 1984). Further examples are rate of decay of patchiness of fish eggs (Smith, 1973), rate of dispersal and coalescence of zebra populations during wet and dry seasons, and the production and erosion of prey patchiness by predators (Schneider, 1992). The time scales in these examples range from days to decades. The spatial scales range from several meters in the case of fish eggs to hundreds of kilometers in the case of coalescence of zebra population at water holes during the dry season. In all three examples the focus on variability calls attention to the spatial and temporal scale (resolution within range) of this variability. In contrast to fractal geometry in ecology, the theme of production and loss of patchiness did not developed as widely.

A fourth recurrent theme is that organisms respond to environmental change at a range of time scales. The time scales of the response of a bird to changes in the surrounding environment range from minutes (behavioral sheltering from storms), to months (seasonal migration), and years (lifetime fitness). The time scale of response by a population depends on whether the response is behavioral, physiological, or genetic; this in turn is often related to the time scale of the environmental change. The time scale of genetic response depends on generation time (Lewontin, 1965), which in turn depends on body size. The time scale of response of communities to change ranges from hours (daily variation in primary production) to millions of years (faunal changes due to vicariant effects of plate tectonics on the biogeography of species). Heterogeneous capacity for response to the environment draws attention to the issue of temporal scale in terrestrial (O'Neill et al., 1986) and aquatic (Harris, 1980) ecosystems.

A fifth recurrent theme is that larger-scale processes interact with local processes to maintain diversity. The classic example is reduction in species diversity through local competitive exclusion (Gause, 1934). Diversity is maintained in space-limited systems when a disturbance opens gaps that allow species squeezed out by competition to survive globally by larger-scale patterns of recolonization (Watt, 1947; Levin and Paine, 1974). Local diversity is maintained through a variety of larger-scale mechanisms that episodically reverse competitive advantages within a patch (Levin, 1981). Another example is larger-scale patterns of frequency-dependent predation by birds on intertidal invertebrates (Schneider, 1978). As a generalization, diversity arises from processes that act at different rates (Levandowsky and White, 1977). The idea that diversity results from rate heterogeneity (e.g., Hutchinson, 1961) produced a substantial literature in the last decades of the 20th century (Carpenter and Turner 2000; Levin, 2000).

A sixth recurrent theme is that effects at one scale propagate to other scales. In a fluid environment, large-scale events propagate to smaller scales (Kolmogorov, 1941, 1962; Stommel, 1963; Mackas, Denman, and Abbott, 1985) with important consequences for the organisms inhabiting the fluids (Dayton and Tegner, 1984; Barry and Dayton, 1991). Our earth consists of three fluid layers: the hot and highly viscous mantle, the

cooler and less viscous envelope of water over nearly 70% of the globe, and the even less viscous envelope of atmospheric gases. Because of differences in viscosity, the time scales of comparably sized structures in the three fluids differ markedly: atmospheric storms last for days and Gulf Stream eddies last for months, whereas the Hawaiian archipelago was created by fluid processes in the earth's mantle over millions of years. Even terrestrial and marine soils have many fluid properties (Nowell and Jumars, 1984; Kachanoski, 1988). Soils creep and mix under the influence of physical processes (e.g., frost heave) as well as biological processes (e.g., bioturbation by annelid worms). The idea that soils are fluids or that terrestrial organisms inhabit fluids (the soil or the atmosphere) does not commonly occur in ecology texts, even though fluid dynamics are as important to terrestrial organisms (Aidley, 1981; Hall, Strebel, and Sellers, 1988; Rainey, 1989) as to aquatic organisms. Energy propagates from large to small scales through chaotic dissipation of energy by fluid processes in the earth's mantle, in the sea, and in the atmosphere. The transfers of energy across scales within these fluid envelopes are best understood as a function of space and time scale in the fluid environments inhabited by life.

A variant on the theme of patterns created by the cross-scale cascade of energy is that mobile organisms extract kinetic energy from large-scale fluid motions to generate local spatial and temporal variability (Schneider, 1991). Flying or swimming animals can time their activity in ways that use larger-scale fluid motions to converge into an area or to remain in place rather than being dispersed by random fluid motions. An example is the seasonal increase in migratory restlessness in birds during passage of a weather front. Increased restlessness timed to the arrival of a high-pressure system carries migratory populations southward in the fall (Richardson, 1978), reducing the costs of migration (Blem, 1980; Alerstam, 1981). Selectively timed movements of migratory birds act like a ratchet, extracting the kinetic energy of storm systems.

A second example is selective tidal stream transport in bottom-dwelling fish (Harden-Jones Walker and Arnold, 1978). Increased swimming activity at one phase of the tide converges fish populations onto their breeding grounds (Arnold and Cook, 1984) or juvenile feeding areas (Boehlert and Mundy, 1988). A third example is activity timed to maintain position against the dispersing effects of fluid motions. Vertical migration timed to the stage of the tide allows zooplankton to extract energy from vertical flow gradients to maintain position against the dispersing effects of estuarine flushing (Cronin and Forward, 1979; Frank and Leggett, 1983). Generation of patchiness at small scales by timed extraction of energy from larger-scale fluctuations focuses attention on propagation of effects across scales.

An important theme in late 20th-century ecology was the application of scaling concepts to experimental research in terrestrial and intertidal ecosystems. The treatment of scale was typically in categories, usually at least three. A typical scale-sensitive experiment would report results from repeating it at a fine scale, an intermediate scale, and a coarse scale. This approach establishes whether results are scale sensitive within a limited range but provides too little information to examine the rate of change in a result with change in scale. The reporting of results on a continuum of scales has a long history in oceanography, extending over half a century. In ecology it is more recent (Legendre and Fortin, 1989; Mayor et al., 2007).

Yet another recurrent theme, perhaps the most important, is that major environmental problems such as global warming, desertification, and acid rain arise through

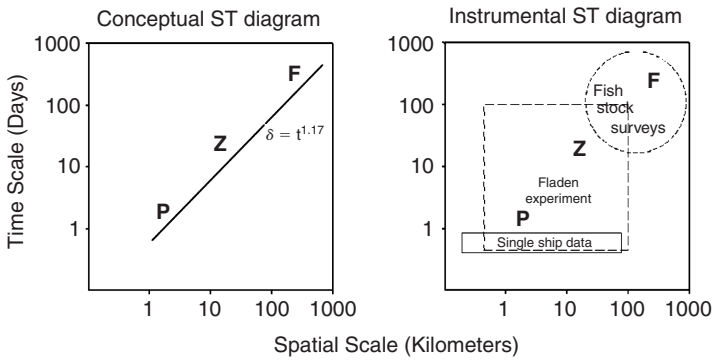
the propagation of effects from one scale to another (Ricklefs, 1990; Loehle, 1991; May, 1991; Holling, 1992). Consequently these problems cannot be attacked at a single scale of investigation (Dayton and Tegner 1984, Steele 1991a, Levin 1992).

ANOTHER LOOK AT SECTION 2.3

Of the seven or so recurrent themes noted in this section, how many have you encountered in either your own research or coursework?

2.4 Graphical Expression of the Concept of Scale

Graphical expression of the concept of scale dates from 1978, when John Steele modified a diagram used in physical oceanography (Stommel, 1963). Stommel's three-dimensional diagram showed variability in sea level against axes of space and time. Steele (1978) used Stommel's space and time axes to construct two novel diagrams. Steele's first diagram showed the space and time scales of patchiness of phytoplankton, zooplankton, and fish (Figure 2.4). This *conceptual space-time (ST) diagram* showed the scales of named phenomena. A conceptual ST diagram first appeared in terrestrial ecology five



Conceptual and instrumental space-time diagrams redrawn from Steele (1978). **P** = Phytoplankton. **Z** = Zooplankton. **F** = Fish.

Graphs with space and time axes increase exponentially after 1980.

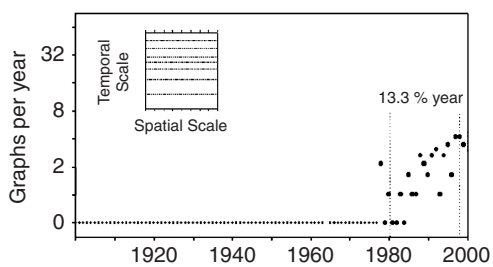


FIGURE 2.4 Graphical Expression of the Concept of Scale

years later (Delcourt et al., 1983). Steele's second diagram compared the time and space scales of coverage by a single oceanographic research cruise to that of an oceanographic program using several ships (Figure 2.4). This *instrumental space-time diagram* showed the capacity of a particular instrument, a research vessel. An instrumental ST diagram first appeared in terrestrial ecology 13 years later (Firbank, 1991).

ST diagrams are highly effective in comparing space and time scales of ecological questions to the capacity of research programs. In agricultural research, Firbank (1991) used an instrumental diagram to compare the space and time scales of experiments with those of surveys. In general, agricultural surveys extended beyond 10 km, with time scales of days to decades. Agricultural experiments were at the scale of a meter, with time scales ranging from one to 100 years. Other examples of ST diagrams compare phenomena to instrumental capacity for disturbance in arctic ecosystems (Walker and Walker, 1991) and detection of pollution effects by benthic experiments (Schneider et al., 1997).

A search of the literature at the end of the 20th century turned up over 60 diagrams with axes showing both space and time scales. The listing excluded reprints of earlier diagrams, which were common, but did include modified versions of earlier diagrams. These graphical expressions of the concept of scale increased exponentially from 1980 to 1998 (Figure 2.4).

ANOTHER LOOK AT SECTION 2.4

Sketch a conceptual space-time diagram for a problem of interest to you, then sketch an instrumental diagram.

2.5 Scale, Scope, and Power Laws

In an instrumental ST diagram, the distance between two points is the *scope*, or ratio of extent to resolution (Schneider, 1994b). In Figure 2.4 the spatial scope of an experimental study of fish on the Fladen Ground in the North Sea was roughly 100 km ÷ 0.8 days = 125, compared to a fish stock survey, with a spatial scope of roughly 1000 km ÷ 20 km = 50. The idea of scope is widely applicable. The scope of a meter stick with a resolution of 0.001 m is 1 m ÷ 0.001 m = 10^3. The scope of a survey of 200,00 km^2 sampled with plots of 50 m by 50 m is 200,000 km^2 ÷ 0.0025 km^2 = 8×10^7. A computational model representing population dynamics at half-day intervals for 10 years has a scope of 3652 days ÷ 0.5 days = 7304. Halving the time step doubles the scope to 14,600.

Scope as a quantitative concept leads naturally to *scaling relations* and *power laws*, both of which coordinate the scope of one quantity with that of another according to an exponent. For *isometric scaling*, the exponent is, by definition, unity. An example is the scaling relation of organism volume to mass.

$$\frac{Vol_{elephant}}{Vol_{mouse}} = \left(\frac{Mass_{elephant}}{Mass_{mouse}}\right)^1 \tag{2.1a}$$

That is, the volume of an elephant is to the volume of a mouse as the mass of an elephant is to a mouse. More generally, volume scales as mass.

$$Vol \cong Mass^{1} \tag{2.1b}$$

This scaling relation holds within groups of organisms that have the same specific gravity.

For Euclidean objects, the exponent is either an integer or the ratio of integers. An example is the *Euclidean scaling* of egg volume to egg surface area:

$$\frac{EggVol_{Albatross}}{EggVol_{Petrel}} = \left(\frac{EggSurface_{Albatross}}{EggSurface_{Petrel}}\right)^{3/2} \tag{2.2a}$$

More generally, egg volume scales as surface area to the 3/2 power.

$$EggVol \cong EggSurface^{3/2} \tag{2.2b}$$

In this case the exponent is a fraction rather than unity and hence is said to be allometric rather than isometric. In addition to fractions based on ratios of integers, the term *allometric scaling* also refers to nonintegral exponents that cannot be obtained from Euclidean geometry.

Organisms are traditionally treated as Euclidean objects having two-dimensional surfaces, living in two- or three-dimensional environments. But in fact organisms have surfaces that are convoluted or fractal (Mandelbrot, 1977), and they inhabit environments that are fractal (Burrough, 1981; Rodriguez-Iturbe and Rinaldo, 1997). For *fractal scalings,* the exponent relating the scope of a quantity to instrumental scope is a noninteger value lying somewhere between the Euclidean landmarks of 0 (a point), 1 (a line), 2 (a plane), or 3 (a volume). An example is the surface of a lung, for which the area scales in a fractal fashion with box area (resolution) of the grid used to measure it.

$$\left(\frac{LungArea_{coarse}}{LungArea_{fine}}\right) = \left(\frac{BoxArea_{coarse}}{BoxArea_{fine}}\right)^{-0.17} \tag{2.3a}$$

More generally, lung surface area scales as box area to the –0.17 power.

$$LungArea \cong BoxArea^{-0.17} \tag{2.3b}$$

Intuitively, the area of a lung has a fixed value, no matter how we measure it. But in practice, measured area will depend in a systematic way on resolution; hence the most informative way to quantify area is to use a power law to describe the rate of change with change in the size of the boxes used to make the measurement. The exponent is estimated by measuring lung area at several different box widths, then regressing area against box width on a log-log scale. Doubling the box size will decrease the measurement of lung area by $2^{-0.17} = 89\%$ because detail disappears as the box size increases. The lung has a *fractal dimension* of $D_f = 2 - (-0.17) = 2.17$. It has a surface more convoluted than a flat surface ($D = 2$) but not so convoluted as to completely fill a volume ($D = 3$). Fractal dimensions lie between Euclidean dimensions.

These examples were drawn from body size scaling, but they also apply to spatial and temporal scaling. An example of isometric spatial scaling is the 1:1 scaling of biomass of marsh grass to marsh area (because marsh grass reaches a fairly constant height). An example of Euclidean scaling with an allometric exponent is the scaling of marsh grass volume to marsh grass area (because marsh grass reaches a fairly constant height at a fairly constant density). An example of fractal scaling is length of the meandering channels that alternately flood and drain a tidal marsh. An early example of fractal habitat structure comes from Pennycuick and Kline (1986), who used dividers set at several step lengths to measure the coastline length inhabited by nesting eagles on Amchitka Island, Alaska. The scaling relation is:

$$\left(\frac{CoastLength_{coarse}}{CoastLength_{fine}}\right) = \left(\frac{StepLength_{coarse}}{StepLength_{fine}}\right)^{-0.7} \tag{2.4a}$$

An equivalent expression is:

$$CoastLength \cong StepLength^{-0.7} \tag{2.4b}$$

The strongly convoluted coastline of Amchitka had a fractal dimension of $D_f = 1 - (-0.7) = 1.7$, which is high compared to the typical coastline dimension of $D_f = 1.3$ (Rodriguez-Iturbe and Rinaldo, 1997).

Equations 2.1 and 2.2 are based on different measurement operations than Equations 2.3 and 2.4. The former equations compare two objects: elephant versus mouse and albatross versus petrel. Equations 2.3 and 2.4 compare *iterative measurement* of the same object: successively coarser scale measurement relative to finer scale measurement. The distinction between the two types of measurement will be a recurrent theme throughout this book. Noniterative measurement is used to quantify similarity across objects. This form of measurement rests on *principle of similitude*, which refers to similarity across objects, as in Equations 2.1 and 2.2. Reasoning based on this principle has a long history in biology, going back to Thompson's 1917 treatise *On Growth and Form* (Thompson, 1961). It has an even longer history outside biology, going back to Galileo. Reasoning based on iterative measurement is more recent, coming into use after 1950. Iterative measurement quantifies *self-similarity*, which is the degree to which a small part of an object resembles a larger part. Iterative measurement is essential for describing complex shapes, from lung areas to coastlines.

Both iterative and noniterative scaling relations result in power laws. A *noniterative scaling relation* equates the scope of one quantity Q_i to that of another quantity Y:

$$\frac{Q_i}{Q_{ref}} = \left(\frac{Y_i}{Y_{ref}}\right)^{\beta} \tag{2.5a}$$

where scope is defined for any object (numerator values) relative to a reference object (denominator values). This scaling relation is rearranged to:

$$Q = \left(\frac{Q_{ref}}{Y_{ref}{}^{\beta}}\right) Y^{\beta} \tag{2.5b}$$

If we take the ratio of Q_{ref} to $Y_{ref}{}^{\beta}$ as a fixed value k, Equation 2.5b becomes a power law that can serve as a *scaling function*.

$$Q = k \cdot Y^{\beta} \tag{2.5c}$$

The scaling relation (Eq 2.5a) is sometimes abbreviated to:

$$Q(M) \cong M^{\beta} \tag{2.5d}$$

This expression implies the presence of the constant k, which as we have seen, contains the reference values appearing in the denominator on either side of Expression 2.5a. Expression 2.5d is more general than the power law, while the power law is more informative because it includes the constant k (Gould, 1971). Because it is more informative, the power law allows computation of Q from any value of Y.

Noniterative scaling includes the scaling of a quantity to a measurement operation, not to another quantity. A *noniterative measurement relation* is defined as one that equates the scope of a quantity to the scope of measurement operation, such as the use of different grid sizes. An example of noniterative measurements is counting the number of species (Q) in quadrats of different size A.

$$\frac{Q_i}{Q_{ref}} = \left(\frac{A_i}{A_{ref}}\right)^{\beta} \tag{2.6a}$$

The distinction between Equations 2.5a and 2.6a lies in the fact that A/A_{ref} is defined by a measurement operation: A_i is a set of fixed multiples of A_{ref}. The reason for this distinction is that area in Equation 2.6a is measured without error. In contrast, area in Equation 2.5a is measured with error. This distinction will reemerge when we use regression (Chapter 15) to estimate scaling exponents. The noniterative measurement relation can be modified to a scaling function.

$$Q = \left(\frac{Q_{ref}}{A_{ref}{}^{\beta}}\right) A^{\beta} \tag{2.6b}$$

If we take the ratio of Q_{ref} to $A_{ref}{}^{\beta}$ as a fixed value k, the measurement relation (Equation 2.6b) can be transformed into a power law scaling function:

$$Q = k \cdot A^{\beta} \tag{2.6c}$$

Box 2.4 demonstrates the first use of a noniterative measurement relation (Equation 2.6a) to estimate a power law (Equation 2.6c) in ecology. In this example the quantity of interest is *Nsp* the number of species. This was scaled to quadrat area A. Box 2.4 demonstrates an early use of a noniterative measurement relation to estimate a power law. In this example the quantity of interest *Nsp* was scaled to quadrat area A, for quadrats of different size. Each quadrat is of fixed size set by a defined protocol and each is measured once.

Box 2.4 First Use of a Scaling Relation and Power Law in Ecology

The measurement relation is

$$\frac{Nsp}{Nsp} = \left(\frac{A}{A_{ref}}\right)^{\beta}$$

For herb-*Pinus* wood (Arrhenius, 1921), a decimeter quadrat contained, on average, 4.8 species; a meter quadrat contained an average of 33 species. For these values, the scaling relation is

$$\frac{33 \cdot species}{4 \cdot species} = \left(\frac{1 \cdot m^2}{(0.1 \cdot m^2)}\right)^{\beta}$$

Taking the logarithm of both sides of the equation, we have:

$$\log(33/4) = \hat{\beta}\ \log(10^2)$$

And hence the estimate of β is:

$$\hat{\beta} = \log(33/4) \div \log(10^2) = 0.4582$$

The scaling parameter k (Equation 2.6c) is:

$$k = \mathrm{Q}_{ref} \cdot \mathrm{A}_{ref}{}^{-\beta}$$

The estimate of the parameter k is:

$$\hat{k} = (4 \cdot \text{species})(0.1^2\,\mathrm{m}^2)^{\beta} = 4 \cdot 10^{(-2)\cdot(-0.4186)} \cdot \text{species} \cdot \mathrm{m}^{2\beta}$$

$$\hat{k} = 33\ \text{species} \cdot \mathrm{m}^{2\beta}$$

The power law is:

$$Nsp = k \cdot A^{\beta}$$

For herb-*Pinus* wood, the power law is:

$$Nsp = 33\ A^{0.4582}$$

This power law cannot be extrapolated to areas beyond $1\,\mathrm{m}^2$ (Gleason, 1922).

Scaling relations can be iterative as well as noniterative. An *iterative scaling relation* is based on repeated measurement of the same object. Iterative scaling relations equate the scope of a quantity Q to the scope of another quantity Y, where Q and Y are both measured at a series of resolutions ranging from L_o to L.

$$\frac{Q(L)}{Q_o(L_o)} = \left(\frac{Y(L)}{Y_o(L_o)}\right)^{\beta} \tag{2.7a}$$

Here, $Q(L) \div Q_o(L_o)$ and $Y(L) \div Y_o(L_o)$ are the ratios of the measurements at several resolutions. An example is the scaling relation of river discharge $Q(L)$ to river length $Y(L)$ at varying step sizes L. This scaling relation is rearranged to:

$$Q(L) = \left(\frac{Q_o(L_o)}{Y_o(L_o)^{\beta}}\right) Y(L)^{\beta} \tag{2.7b}$$

Taking the ratio of $Q_o(L_o)$ to $Y_o(L_o)^{\beta}$ as a fixed value k, Equation 2.7b becomes a power law scaling function:

$$Q(L) = k \cdot Y(L)^{\beta} \tag{2.7c}$$

Iterative scaling includes the scaling of a quantity to a measurement operation, not to another quantity. *Iterative measurement relations* equate the scope of a measured quantity to a set of fixed multiples of a base unit, defined by a measurement protocol. An example of an iterative measurement relation is recording the perimeter of a lake, using different step lengths L.

$$\frac{Q}{Q_o} = \left(\frac{L}{L_o}\right)^{\beta} \tag{2.8a}$$

The scaling relation can be modified to a scaling function:

$$Q = \left(\frac{Q_o}{L_o{}^{\beta}}\right) L^{\beta} \tag{2.8b}$$

Taking the ratio in parentheses as a fixed value k, Equation 2.8b becomes a power law scaling function:

$$Q = k \cdot L^{\beta} \tag{2.8c}$$

Table 2.1 lays out the consequences of iterative and noniterative measurement for simple and complex phenomena. For simple phenomena such as dinosaur running speed in relation to leg length, we already know from living animals that the dynamics of running speed depend more on leg length than on the complexities of leg shape. Consequently, we measure the straight-line distance across the length of the bone. We can then relate this quantity to velocity compared across animals via a scaling relation (Equation 2.5a) or a scaling function (Equation 2.5c). Because we are using a scale based on straight-line distance, we can compute the length of the leg bone in any units we choose. Iterative measurement, such as measuring the bone length in units of mm, then in units of cm, is not needed.

Table 2.1 Relation of Iterative and Noniterative Measurement to Simple and to Complex Phenomena

	Phenomenon: **Example:** **Geometry** **Dynamics**	**Simple** **Dinosaur Leg Length:** **Euclidean** **Linear**	**Complex** **Coastline Length:** **Fractal** **Nonlinear**
Noniterative measurement		OK	Observers cannot agree
Iterative measurement		Not needed	OK

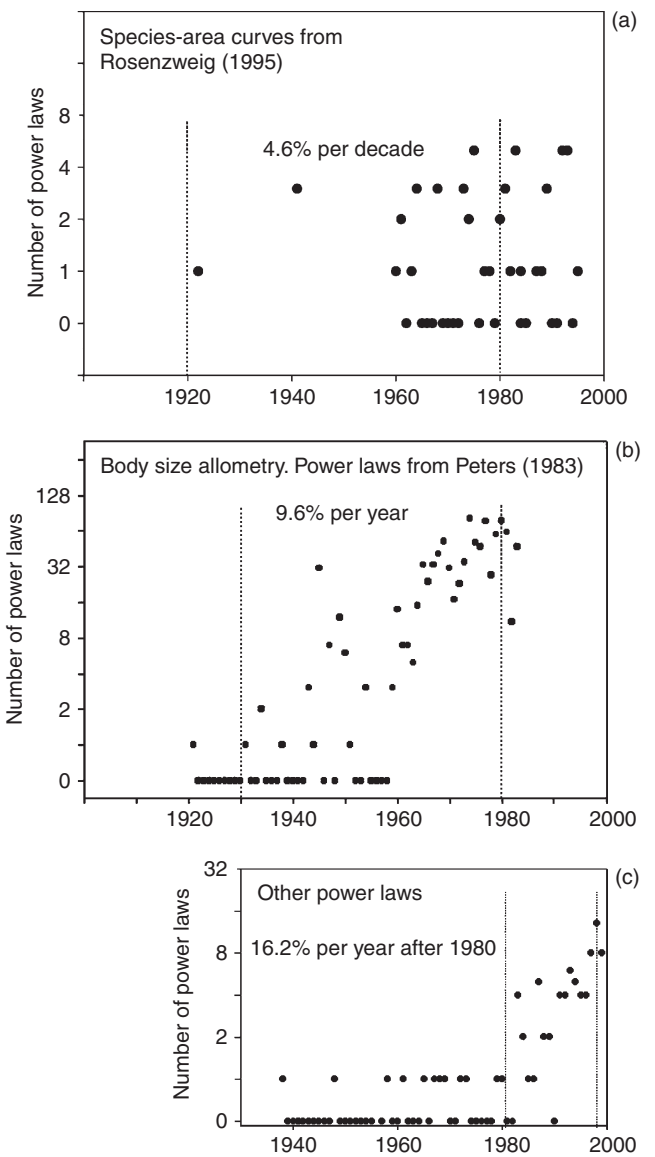


FIGURE 2.5 Number of Power Laws per Year. Redrawn from Schneider (2001b). (a) Species area curves; (b) Body size allometry; (c) Other scaling relations.

For fish habitat in a coastal nursery, we are interested in the complexity of coastline length because indented coastlines provide shelter from predators. If we use noniterative measurement, we will obtain different values of coastline length, depending on the unit size. We cannot compute the length at some scale of interest, given length at some other scale. To characterize coastline complexity, we need to use iterative measurement. This will allow us to quantify complexity as the rate of loss of detail with change in ruler size. This rate of loss, summarized as the exponent of a scaling relation (Equation 2.8a) or scaling function (Equation 2.8c), characterizes coastline complexity. It allows us to compute a complex quantity such as coastline length at any scale.

Until recently, power laws and scaling relations in ecology came mostly from the literature on body size allometry (based exclusively on noniterative measurement) and species area curves (based on either iterative or noniterative measurement). Rosenzweig (1995) discussed 48 articles containing 55 species-area curves for which the exponent can be determined. In this list the number of power laws grew exponentially from 1920 to 1980 (Figure 2.5a). In organismal biology, Peters (1983) listed 251 articles containing 1050 power laws that scale organism form or function to body size. Growth in the number of power laws (Figure 2.5b) was exponential from 1930 to 1980.

Power laws other than species-area curves or body size allometry appeared sporadically in ecology from 1940 to 1980, then grew exponentially at 16% yr^{-1} after 1980 (Figure 2.5c). Most of these power laws describe habitat complexity (O'Neill et al., 1983; Morse et al., 1985). Figure 2.5c shows only those power laws for which the constant k is reported or can be determined. The number of power laws would be much greater if it included scaling relations or measurement relations where only the exponent was reported.

Power laws may well be more common than is evident from the published literature. In normal practice, little or no effort goes toward identifying the underlying scaling or measurement relations (Equations 2.5a, 2.6a, and 2.7a). A linear scaling (exponent = 1) is often used in regression estimates of the scaling of one variable to another, even though a linear scaling may not be appropriate. Power law relations are easily overlooked because of overly facile application of linear regression. Are there guidelines for when to expect a linear versus a power law relation between two variables? In general, one expects a linear relation if the process linking them is additive (Table 2.2). That is, we expect a linear relation if an increment of change in one variable results in a constant increment to the other. For example, the biology of somatic growth leads us to expect each increment in mass of an organism to be matched by a

Table 2.2 Type of Linkage in a Functional Expression $y = f(x)$. Linkage depends on whether x and f(x) are additive (taking a difference) or multiplicative (taking a % or ratio)

x	***f(x)***	**Type of Linkage**	**Example**
Additive	Additive	Linear	Equation 2.8
Additive	%	Exponential	Equation 2.9
%	Additive	Logarithmic	—
%	%	Power	Equation 2.10

fixed increment in volume. The result is a linear relation between mass and volume and volume of an organism:

$$Mass = k \cdot Volume^{1} \tag{2.9}$$

We expect an exponential relation (Table 2.2) if an increment of change in an independent variable results in a fixed percentage increment in the dependent variable. For example, the biology of unrestricted population growth leads us to expect each increment of time to result in a fixed percentage increase in population numbers. The result is an exponential relation between variables. For unrestricted population growth, the relation of mice numbers to time is:

$$Mice_t = Mice_{t=0} \cdot e^{k \cdot t} \tag{2.10}$$

We expect a logarithmic relation (Table 2.2) if an increment (as a percentage) in an independent variable leads to an additive increment in the dependent variable. Finally, we expect a power law relation (Table 2.2) if an increment (as a percentage) in one variable is matched by an increment (again as a percentage) in another variable. An example is allometric growth in length of a body part (leglength) relative to the size (volume) of an organism. If the percentage change per unit time of a particular body part is fixed but does not match the (fixed) percentage change in overall growth rate (increase in volume and hence mass), we expect a change in proportion of that part to the whole, as described by a power law.

$$Leglength = e^{k_{ref}} \cdot Volume^{\left(\frac{k_{leglength}}{k_{Vol}}\right)} \tag{2.11}$$

This expression says that if leg growth ($k_{Leglength}$) exceeds volume growth (k_{Vol}), the exponent of the resulting power law exceeds a value of one: the organism becomes more leggy as it grows. Conversely, if leg growth fails to keep pace with volume growth (exponent less than unity), an organism becomes less leggy over time.

Power law relations can arise in many ways, of which the body size allometry example is only one. It turns out that power laws result from antagonistic rates that act episodically at different frequencies or time scales. This view of the origin of power laws, called *complexity*, is well known in physics. Applications are now starting to appear in biology (Schneider, 2001b). Power laws that arise from the lurching dynamics that generate complex phenomena have the potential to become the theoretical basis of spatial and temporal scaling in ecology (Chapter 16).

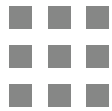
ANOTHER LOOK AT SECTION 2.5

Write a scaling relation (Equation 2.5a) for two quantities of interest to you. Is the exponent an integer, the ratio of integers, or a number estimated from data?

Defined Concepts and Terms for Review and Future Reference

Here is a list of key concepts that are explicitly defined in this chapter. For each concept, reflect briefly on its definition and context. You may wish to write down page numbers so you can use this list for later reference.

____ analysis, multiscale
____ complexity
____ diversity at four inventory levels: point, alpha, gamma, epsilon
____ diversity at three differentiation levels: pattern, beta, delta
____ measured variable, scale of
____ measured variable, spatial scale of
____ measured variable, temporal scale of
____ measurement, iterative
____ measurement relation, iterative
____ measurement relation, noniterative
____ measurement scale, type of
____ natural phenomena, scale of
____ patterns, scale dependent
____ processes, scale dependent
____ power laws
____ principle of similitude
____ space-time diagram, conceptual
____ space-time diagram, instrumental
____ scale, cartographic
____ scaling, allometric
____ scaling, ecological
____ scaling, Euclidean
____ scaling, fractal
____ scaling, isometric
____ scaling, noniterative
____ scaling function
____ scaling relation, iterative
____ scaling relation, noniterative
____ scope
____ self-similarity



PART

II

Scaled Quantities

3 Quantities

As with Aristotle, the Schoolmen considered things as more and less than each other, but not in terms of multiples of a definite quantity such as inches, degrees of arc, degrees of heat, and kilometers per hour. The Schoolmen, paradoxically, were mathematicians without being quantifiers.

—A. W. Crosby, *The Measure of Reality*, 1971

3.1 Synopsis

Many of the uses of scale in ecology are connected to the concept of scaled quantities, which link theory to measurement via scaled numbers. Ecologists, like all natural scientists, work with definable quantities, not with numbers divorced from units of measurement. A quantity consists of a name, symbol, procedural statement, numbers, and units of measurement. The rules of clear communication apply to the procedural statement, name, and symbol. The procedural statement should permit replication of the measurements. The name should convey a sense of the quantity. The symbol should be unique yet lend itself to easy visualization of the quantity for which it stands.

Units occur on four types of measurement scale: nominal, ordinal, interval, and ratio. The mathematical rules that apply to scaled quantities are more restrictive than those that apply to numbers.

Graphs must show the name and units of quantities, in addition to the plot of numbers. Showing the symbol connects the graph to procedural statements or equations in the surrounding text.

3.2 Definition of a Quantity

Like all natural scientists, ecologists work with definable quantities, not with numbers or mathematical abstractions divorced from measurement (Riggs, 1963). Ecologists work with quantities that have names and scaled values: a density *[N]* of 5000 animals per hectare, or an increase rate r of 4% per year, or a mutation rate μ of 10^{-6} per generation. Our interest is in physically or biologically interpretable quantities, not the mathematical manipulation of symbols. When told that dx/dt means

$$\lim_{\Delta t \to 0} \frac{\Delta x}{\Delta t},$$

doi: 10.1016/B978-0-12-627865-1.00003-9

the physicist Kelvin exclaimed, "Does nobody know that it represents a velocity?" (Hart, 1923).

Unfortunately, equations in ecological journals are too often presented as abstractions lacking any biological interpretation or statement of the units and dimensions of each symbol. As a result, the equation is incomprehensible; one quantity cannot be distinguished from another, and there is no way to know whether the results are comparable to previous results. If units and scale are absent from a theoretical report, there is no way to know the scale at which testing is appropriate. Numbers and symbols with no units force us to guess whether two studies are comparable or to guess what scale of measurement to use in testing a theoretical result.

A full definition of each quantity is important because the rules for working with quantities differ from the rules for working with numbers or algebraic symbols. One can take the logarithm of the number 4, but one cannot take the logarithm of 4 mosquitoes. Adding A to B makes sense if A and B are numbers, but adding $A = 4$ cabbages to $B = 8$ kingfishers makes no sense. Nobody would add 4 and 20 blackbirds to the number $\pi = 3.14$, but the expression $N + \pi$ makes it all too plausible unless quantities are defined and distinguished from numbers.

A fully defined quantity has five parts:

- A *name*
- A *procedural statement* that prescribes the conditions for measurement or calculation from measurements
- A set of *numbers* generated by the procedural statement
- *Units* on one of several types of measurement scale
- A *symbol* that stands for the set of scaled numbers

The units apply to all the numbers, so a convenient way of representing a quantity is to arrange the numbers into a *vector*, which is a sequence of numbers inside brackets. The symbol stands for the product of the units and the vector of numbers. Here is an example:

Procedural Statement	Name	Symbol	Numbers	Units
Gravimetric mass, at pupation	Pupal mass	PM=	$\begin{bmatrix}280\\250\\300\end{bmatrix}$	Milligrams

An adequate "Methods" section in a scientific report should contain these components. Unfortunately, practice is otherwise: Completely defined quantities are not the rule in the ecological literature. Symbols are often absent from experimental or field studies, whereas units are rarely used in theoretical journals (see Table 1.1 in Chapter 1). This contributes to the communication gap between theoreticians and field ecologists noted by Kareiva (1989). This gap could be bridged by better use of quantities, defined as symbols in reports of field research and treated as scaled quantities in theoretical reports. This gap is closing with the appearance of texts (e.g., Case, 2000) that make effective use of scaled quantities that underlie ecological theory.

3.3 Names and Symbols

Quantities should be read as *names* ("per capita birth rate"), not as symbols ($\dot{B}/N$), because a name conveys more meaning. Symbols appear in mathematical expressions for

the sake of clarity and in prose for the sake of preciseness, but when encountered symbols should still be read as names. Facility in reasoning with quantities comes in associating a name with a symbol, with a mental image of the biology, and with some typical values. For example, the quantity "per capita birth rate" is associated with a symbol $\dot{B}/N$ and with an image of the quantity, such as chicks jumping out of the nest of a pair of adult birds each year. Name, symbol, and image are further associated with a typical value obtained from calculation:

$$\dot{B}/N = \log_e(5 \text{ chicks}/2 \text{ parents})/\text{year} = 92\%/\text{year}$$

Skillful choice of *symbols* aids in understanding and reasoning with quantities. Mnemonic symbols are easier to remember and use than something arbitrary. A fisheries scientist, John Pope, has suggested that easily remembered icons (♆ = Number of trees) be used rather than letters (N = Number of trees). This reduces the burden of recalling the meaning of the symbol, but until recently it was impractical because of limits on typesetting of unusual symbols. The graphics capability of computer-based typesetting programs should make this increasingly practical. Coordination between symbols also aids recall. An example is x, y, and z for position in space relative to three axes. Another device that aids recall is to add a diacritical mark to familiar symbols rather than selecting a new symbol. An example is the use of the symbol $\overline{A}$ rather than m_A to designate the mean value of the surface area of lakes in a district. The symbol $\overline{A}$ emphasizes that the quantity is an area, whereas the symbol m_A obscures the sense of the quantity, which is area.

Another example of diacritical marks, common in physiology, is to place a dot over a quantity to represent the time rate of change in that quantity. The instantaneous time rate of change in the quantity Q is:

$$\dot{Q} \equiv dQ/dt \quad (\equiv \text{ means "equal by definition"})$$

This notation is due to Newton, who used a dot over a symbol to denote the time rate of change in the quantity represented by the symbol. The dot notation results in a simpler symbol, which is easier to read in an equation than the more complex symbol dQ/dt. In applied mathematics, Newton's compact notation works well because it draws attention to the quantity Q rather than the mathematical operator d/dt. In a similar way, compact notation for the spatial gradient ∇Q in the quantity Q works better in applied settings than the equivalent but more complex symbol dQ/dx, where x is location along a line.

In principle one can use any symbol for a quantity, but in practice conventional symbols are preferred because they allow rapid recognition of familiar quantities. A conventional symbol such as g for acceleration in the earth's gravitational field takes on meaning through frequent and consistent usage. Unfortunately, there is little consistency in use of symbols in ecology. There is little enough consistency within an area of ecology, even less consistency among areas (Krebs, 1972), and frequent conflict with conventional use outside ecology. For example, in demography the subscript x conventionally means "time since birth of a cohort"; this conflicts with the equally conventional use of x to mean "horizontal location in a three-dimensional xyz coordinate system relative to the earth." This conflict creates notational problems for spatially distributed population processes. When notation conflicts, precedence tends to go to the more widely used set of symbols.

The abstract language of symbols and scaled quantities is, at the outset, as incomprehensible as any new language. I am convinced that this abstract language is worth learning because it allows calculations about quantities of interest and importance, like the spread of Africanized bees or the productivity of the sea. This language is easier to learn than other languages because the vocabulary is smaller, with restricted definitions. Still, it is foreign (at first) and highly abstract, so we deserve a dictionary. With a dictionary it is only a little effort, rather than a lot of leafing through pages of text, to find the forgotten meaning of a symbol. If the dictionary is serving its purpose well, it becomes less necessary with time because the symbol becomes tied to a name and a concept. The reader is encouraged to write out symbols with the name and units, in a list at the end of the book. Readers fluent in more than one language may want to write the names in more than one language. Be sure to list diacritical marks separately, as these can be placed over any symbol.

ANOTHER LOOK AT SECTION 3.3

Define a measurable quantity of interest to you (name, symbol). Now define a new quantity by putting a dot over the symbol to represent the time rate of change in the quantity under the dot. State as concretely as possible how you visualize this new quantity.

3.4 Procedural Statement

The *procedural statement* must supply enough information so that another person could use it to obtain comparable numbers on the same scale. The statement should include the conditions for measurement. An example might be defining how we determined the end of a larval stage. This would be important in taxonomic groups such as fish, which do not end larval life with the dramatic pupation found in insects. The statement of measurement operations might be simple, referring only to standard units such as kilograms, meters, and seconds. The statement might include complex procedures, such as those of Winberg (1971) for calculating the production rate of a population.

Procedural statements are typically a mixture of measurement operations and calculations. Philosophical treatments of the topic of measurement (Campbell, 1942; Cushman, 1986) distinguish directly measured quantities from quantities derived by calculation from "laws." This distinction misses the practical impact of modern electronic instruments, which use empirical equations to report one quantity (such as salinity) calculated from direct measurement of a different quantity (such as electrical conductivity). Measurement devices with computer chips report scaled quantities that are a mixture of gauge readings and calculations. In light of this incorporation of calculations directly into the measurement device, it is especially important to report exactly how scaled numbers were obtained.

Another component of the procedural statement that will prove important in multiscale analysis is *recursive measurement*. To give an example, we can measure lake peripheries recursively at multiple scales of resolution. A recursive method begins with a single unit of measure, computes periphery at that scale, repeats this with a new unit of measure on the same lake, and continues at successive unit sizes to establish the rate of change in periphery with change in unit of measure. The mode of recursion matters. If we overlay an aerial photograph of a lake with grids of different resolution to obtain periphery at each scale of resolution, we will obtain a different measure than if we swing a pair of

dividers (set at different step lengths) along the periphery. If we swing a pair of dividers outward from the lake boundary at each step, we will obtain a different measurements than if we swing the dividers inward at each step. If measurement is recursive, based on repeated measurement of the same object, it is important to state the procedure.

ANOTHER LOOK AT SECTION 3.4

Write a procedural statement for a quantity of interest to you. Is it complete enough that someone else could use it to generate measurements comparable to yours?

3.5 Types of Measurement Scale

Stevens (1946) defined four types of measurement scale: nominal, ordinal, interval, and ratio scales. The outcome of a *nominal scale measurement* is a yes or no decision about whether an object belongs to a class. An example is whether a species occurs in an area. The outcome of *ordinal scale measurement* is a ranking: first, second, third, and so on. Comparison of objects produces a ranking, with no information about the magnitude of the difference between adjacent ranks. An example is the order of arrival of new species on a defaunated island. The outcome of *interval scale measurement* is the number of units that separate the objects of measurement from one another. An example is the body temperature of an animal in degrees Centigrade. There is no natural zero point, so the temperature of one animal cannot be said to be twice that of another on this scale. Clock time and degrees of longitude are also on interval scales. In contrast, *ratio scale measurements* have a natural zero point. The outcome of measurement on this scale is the number of units that separate the measurement from the zero point. The Kelvin temperature scale has a zero point (no thermal energy of molecules), and so on this scale the body temperature of an animal can be said to be 98% of that of another. Similarly, the number of organisms of one species in a quadrat can be said to be one-tenth that of another species, the length of one animal can be said to be three times that of another, and the intrinsic rate of increase of one population can be said to be 1.5 times that of another.

The procedural statement determines the type of measurement scale. For example, if the temperature in the nesting burrow of a shearwater is recorded in degrees Celsius, the result is an interval scale quantity with interval scale units of degrees °C. Measurements can be converted to a ratio scale, degrees Kelvin, by adding 273° to each reading.

A variable quantity typically consists of a set of values generated by the procedural statement. The numbers are gathered together inside brackets to form a vector. The same unit applies to all the numbers, so it is conveniently placed outside the brackets rather than being repeated inside the brackets. The vector of outcomes is rewritten as the product of a vector of numbers and a unit, such as °K:

Name	Symbol	Outcomes	=	Numbers	.	Units
Burrow temperature	$bT=$	$\begin{bmatrix}284.1\ ^\circ K\\283.8\ ^\circ K\\285.2\ ^\circ K\end{bmatrix}$	=	$\begin{bmatrix}284.1\\283.8\\285.2\end{bmatrix}$	·	°K

Here are the same measurements, this time on the original interval scale:

Name	Symbol	Outcomes = Numbers	.	Units
Burrow temperature	$bT=$	$\begin{bmatrix}11.1\ °C\\10.8\ °C\\12.2\ °C\end{bmatrix} = \begin{bmatrix}11.1\\12.8\\12.2\end{bmatrix}$	·	°C

An ordinal scale quantity results from ranking objects by direct comparison of objects or by comparison of more detailed measurements. An example is a sequence of five population counts that are thought to be accurate to rank, and no more:

Name	Symbol	Outcomes = Numbers	.	Units
Population size	$N_{ordinal} =$	$\begin{bmatrix}\text{third}\\\text{second}\\\text{first}\\\text{fourth}\\\text{fifth}\end{bmatrix} = \begin{bmatrix}3\\2\\1\\4\\5\end{bmatrix}$	·	Rank

The five outcomes have been ranked on the basis of more detailed counts, then gathered together in vector form, with the numbers inside the brackets, and a unit called a *rank* has been attached to the collection. Some might claim that a quantity must be on at least interval scale, if not on a ratio scale. But there is no logical justification for this claim (Russell, 1937, p. 183).

A nominal scale quantity results if only presence or absence is recorded. An example is the presence of insect larvae in three quadrats. The result of measurement is a quantity with nominal units (presence or not) on a nominal scale. Each outcome can be written as the product of a unit (presence) and a binary number: 1 for presence, 0 for presence. The measurement outcomes from the three quadrats are again gathered together in vector form inside brackets. This vector of outcomes can be rewritten as the product of a unit (presence) and binary numbers:

Name	Symbol	Outcomes = Numbers	.	Units
Larval presence	$N_{nominal} =$	$\begin{bmatrix}\text{present}\\\text{absent}\\\text{present}\end{bmatrix} = \begin{bmatrix}1\\0\\1\end{bmatrix}$	·	Presence

A quantity on a nominal scale can consist of several categories. The result is a multinomial rather than a binomial quantity. These multinomial quantities can be decomposed into binomial quantities. The outcomes can be written in logical categories: A + B for the presence of species A and B in the first quadrat, B + C for presence of species B and C in the second quadrat, and so on:

Name	Symbol	Outcomes = Numbers	.	Units
		Species **A B C**		
Larval presence	$N_{multinominal} =$	$\begin{bmatrix}A+B\\B+C\\A+C\end{bmatrix} = \begin{bmatrix}110\\011\\101\end{bmatrix}$	·	Presence

The vector of outcomes (A + B, B + C, A + C) has here been rewritten as the product of binary numbers and units (presences).

ANOTHER LOOK AT SECTION 3.5

For each of the four types of measurement scale, define (using name symbol and procedural statement) a quantity of interest to you.

3.6 Graphing Scaled Quantities

A graph of a scaled quantity must contain three of the five components: name, units, and numbers. A fourth component, the symbol, adds to the presentation by linking the graph to the text. A convenient way to include these components in a graph is to list the name of the quantity along the axis, then list the symbol and units, connected by an equality sign. Figure 3.1 shows this format, which links the quantity in the graph to a procedural statement in the text.

Listing the symbol is also a useful way of showing any rescalings of the axis. Logarithmic rescalings arise naturally in multiscale analysis because of the emphasis on proportional rather than additive changes. The three most common forms of logarithmic scaling are doublings (base 2), e-fold changes (base e), and tenfold changes (base 10). A logarithmic scale of numbers can be used as labels, as shown in Figure 3.2.

A logarithmic axis can also be labeled with the corresponding exponents, as shown in Figure 3.2. One of the common failings of graphical presentation of quantities on a logarithmic scale is that the base of logarithms is not reported. If only the exponent is reported, without the base, we have no way of knowing whether the exponent 3 represents 2^3, e^3, or 10^3. The solution to this common defect is to use the symbol to show exactly how the rescaling was done, as in the example of $\log_2(A/\text{cm}^2)$ in Figure 3.2. The advantage of this notation is that it allows us to stick with our intention of working with scaled quantities rather than numbers stripped of units. The quantity A has been divided by the base units, which reduce it to a unitless number. The logarithm of this ratio can then be taken (the operation of taking the logarithm of a unit such as cm^2 is

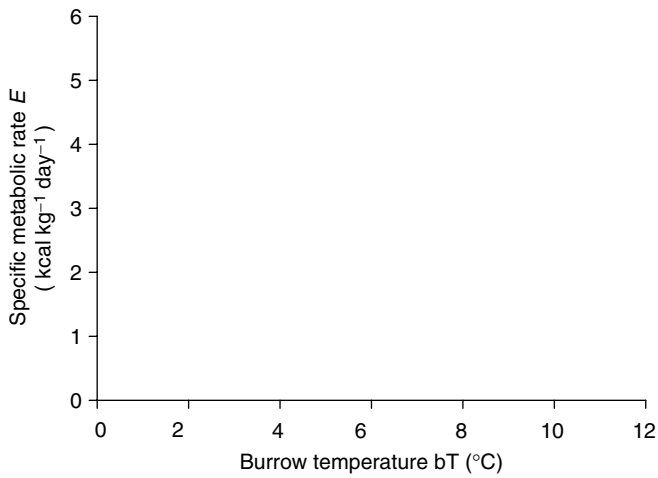


FIGURE 3.1 Fully Labeled Graph. The name of quantity, a symbol, and units are shown for both response (dependent) and explanatory (independent) variables.

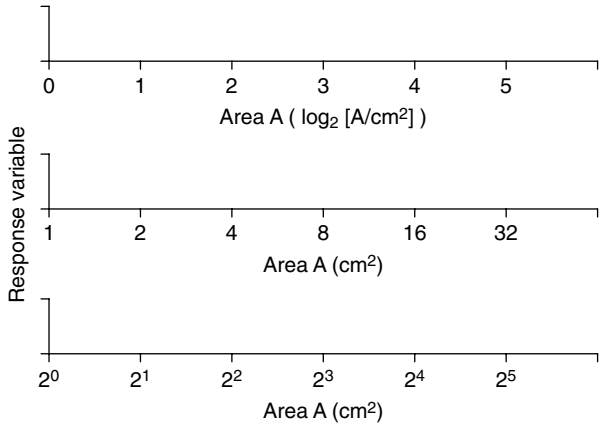


FIGURE 3.2 Three Different Ways to Label Logarithmic Axes Fully.

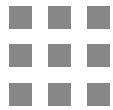
not defined). This illustrates an important principle described in the next chapter, which is that the rules for working with units and scaled quantities are not the same as the rules for working with numbers.

ANOTHER LOOK AT SECTION 3.6

Redo Figure 3.2 in base 10 rather than base 2, for a scaled quantity of interest to you.

Defined Concepts and Terms for Review and Future Reference

_____ fully defined quantity
_____ nominal, ordinal, interval, ratio scale measurements
_____ recursive measurement



4 Units

Let us consider the speed and momentum acquired by a body falling through the height, say, of a spear as a standard which we may use in the measurement of other speeds and momenta as occasion demands.

—Galileo, *Two New Sciences*, 1638

4.1 Synopsis

Units on a ratio type of scale are the basis for scaling. Multiscale analysis requires ratio scale units because these units can be used to represent the action of repeatedly halving or doubling a quantity. Zooming in on detail requires units that can be halved repeatedly, an operation that is natural with ratio scale units but not possible with nominal, ordinal, or interval scale units.

Standard units on a ratio scale are defined relative to seven base units in the SI system. Derived units are generated by taking the products and ratios of standard units. Derived and standard units come in standard multiples with standard prefixes: kilo, micro, and so on. Unconventional units such as Galileo's spear lengths are just as valid as standard units. In ecology, unconventional units may prove more useful than conventional units inherited from Euclidean geometry and mechanics, which omit much of the biology. An example is a fractal length to describe the convoluted paths of foraging animals, or a fractal area to describe the convoluted surfaces of the natural world, from leaves to trees to watersheds. Another example is the unit of an entity, which is far more useful than a mole (10^{23} entities) in analyzing population processes.

The rules that apply to ratio scale units differ from those that apply to numbers. The rules define the operations of addition, subtraction, multiplication, division, exponentiation, and the taking of absolute values. In applying these rules, it is important to distinguish similar units from dissimilar units. Similar units, such as units of mechanical and thermal energy, can be added together. Dissimilar units, such as units of time and distance, cannot be added together. The decision about which units are similar depends on biological reasoning, not on mathematical rules.

4.2 The Utility of Ratio Scale Units

Ratio scale units have several useful properties. One is that they allow unit conversion. An example is the calculation of energy flow through populations. Several units of energy occur in the literature: Joules, gram-calories, kilogram-calories, British thermal units. Still more units of energy per unit time (power) exist: kilocalories per day, liters of oxygen per hour, ergs, and Watts. Because these are ratio scale units, a table of conversion factors

doi: 10.1016/B978-0-12-627865-1.00004-0

can be used to make the conversions necessary to calculate energy flow through a system such as a freshwater spring (Odum, 1957) or a salt marsh (Teal, 1962).

The properties of ratio scale units make them useful in interpreting symbolic notation used in equations. The formidable symbol $e^{-\dot{D}\cdot t}$ takes on meaning by giving it a name, then undertaking calculations with scaled values. The symbol $e^{-\dot{D}\cdot t}$ represents the percentage of the population remaining after suffering a death rate of $\dot{D}$. Some sense of the symbol can be gained by substituting numbers ($t = 2$ and $\dot{D} = 0.01$), but a far better sense is gained by substituting scaled values such as $t = 2$ days and $\dot{D} = 0.01$ day^{-1}. Box 4.1 shows a series of calculations. Examining these calculations is no substitute for taking out a calculator right now to do the calculations.

Box 4.1 Interpretation of the Symbol $e^{-\dot{D}\cdot t}$ via Calculation with Scaled Values.

instantaneous mortality	time	% remaining
$\dot{D}$	t	$e^{-\dot{D}\cdot t}$
0.01 year^{-1}	1 year	99%
0.1 year^{-1}	1 year	90%
0.2 year^{-1}	1 year	82%
0.2 year^{-1}	2 years	67%
0.4 year^{-1}	1 year	67%

Ratio scale units have another useful quality, which is that they can be combined to make new units via multiplication and division. Multiplication of a unit by itself changes the exponent of the units. Such operations can be visualized. Think of sweeping sticks at right angles to make areas, or measuring velocity as the frequency (in units of time^{-1}) with which a unit of distance is traversed. This operation of changing exponents will become important in working with fractal objects, such as the convoluted structure of the stream beds inhabited by fish.

Yet another useful quality is that ratio scale units permit analysis at multiple scales of space and time. We can visualize multiscale analysis (Box 2.1) as the operation of zooming in toward greater detail, or conversely, expanding the scale to reveal larger-scale pattern and process. To represent this idea in formal terms so that calculations can be made, we need units that can be repeatedly doubled (or halved). Ratio scale units can be reduced by factors of 2, 10, or any other base. This increases the resolution, allowing us to zoom in on detail, an operation that we can represent mathematically by dividing units repeatedly to finer scales of resolution. Conversely, we can zoom back to frame larger-scale patterns. This concept is represented mathematically by expanding the range (or decreasing the resolution). These operations, which are natural with ratio scale units, cannot be carried out on ordinal or interval types of measurement scales.

Because ratio scale units have these useful properties, some people have taken the position that only ratio scale units are valid (e.g., Campbell, 1942). This narrow view does not stand up to logical analysis (Stevens, 1975; Luce and Narens, 1987). Thus in

defining a biological quantity it is more important to provide a clear statement of the type of units than to provide no definition because a ratio scale unit is not applicable.

ANOTHER LOOK AT SECTION 4.2

Ecologists often use transformations before undertaking inferential statistical analysis of data. Comment on what is lost by taking the square root of data on tree density per $1\,m^2$ area.

4.3 Standard Units

Standard units on a ratio scale are defined against a standard base so that anyone anywhere can obtain comparable results. Table 4.1 lists the seven *SI base units* in the International System of Units (Système Internationale, abbreviated SI). This system includes two supplementary units, one for plane angles and one for solid angles. These appear in definitions of angular velocity, acceleration, and momentum. They also appear in definitions of light flux and light exposure.

Combinations of the base units result in *derived units*. Some of these derived units have names, such as Watts for units of Joules per second. Table 4.2 lists derived units that commonly occur in ecology. A list of over 60 derived units can be found in Legendre and Legendre (1998). A collection of ratio scale constants and quantities used in marine ecology can be found in Mann and Lazier (1991).

The divisibility of ratio scale units into successively smaller fractions allows us to define a series of standard multiples. These and their abbreviations are shown in Table 4.3. These standard multiples, 10^{-1} = deci, 10^{-3} = milli, and 10^{-6} = micro, yield new units from a basic unit such as the Watt, a unit of energy. The *standard multiple units* listed in Table 4.3 are applicable to any ratio scale unit.

ANOTHER LOOK AT SECTION 4.3

1. Of the units in Table 4.2, how many have you used?
2. Of the multiples in Table 4.3, how many have you used?

Table 4.1 Base and Supplementary Units in the SI System

Quantity	Unit	Abbreviation
Length	Meter	m
Mass	Kilogram	kg
Time	Second	s
Thermodynamic temperature	Kelvin	K
Amount of substance	Mole	mol
Luminous intensity	Candela	cd
Electrical current	Ampere	A
Planar angle	Radian	rad
Solid angle	Steradian	sr

Table 4.2 Units That Commonly Occur in Ecology

Quantity	Units	Equivalent Unit (Name)
Acceleration		
Angular	$rad \cdot s^{-2}$	
Linear	$m \cdot s^{-2}$	
Area	m^2	
	$10^4 \cdot m^2$	ha (hectare)
Concentration	$mol \cdot m^{-3}$	
Energy (work)	$N \cdot m$	J (Joule)
	$4185 \cdot J$	Kcal (kilocalorie)
Energy flux	$J \cdot m^{-2} \cdot s^{-1}$	
Force	$kg \cdot m \cdot s^{-2}$	N (Newton)
Frequency	s^{-1}	Hz (Hertz)
Light		
Luminance	$cd \cdot m^{-2}$	
Luminous flux	$cd \cdot sr$	lm (lumen)
Illuminance	$lm \cdot m^{-2}$	lx (lux)
	$10.764 \cdot lx$	fc (footcandle)
Photon flux (*PAR)	$1 \cdot mole \cdot m^{-2} \cdot s^{-1}$	E (Einstein)$\cdot m^{-2} \cdot s^{-1}$
Mass density	$kg \cdot m^{-2}$	
Mass flow	$kg \cdot s^{-1}$	
Mass flux	$kg \cdot m^{-2} \cdot s^{-1}$	
Power	$J \cdot s^{-1}$	W (Watt)
Pressure (stress)	$N \cdot m^{-2}$	Pa (Pascal)
Surface tension	$N \cdot m^{-1}$	
Velocity		
Angular	$rad \cdot s^{-1}$	
Linear	$m \cdot s^{-1}$	
Viscosity		
Dynamic	$Pa \cdot s$	
Kinematic	$m^2 \cdot s^{-1}$	
Volume	m^3	(cubic meter)
	$10^{-3} m^3$	l (liter)
Volume flow rate	$m^3 \cdot s^{-1}$	
Wavelength	m	
Wavenumber	m^{-1}	

*PAR = Photosynthetically Active Radiation.

4.4 Unconventional Units

We do not need to restrict ourselves to standard units in reasoning about quantities. Unconventional units are as valid as standard units. Galileo used a spear length to reason quantitatively about velocities and momenta. A spear length is just as good for quantitative reasoning as the standard unit, a meter. But even though conclusions from quantitative reasoning are independent of the choice of base units, our ability to communicate a result does depend on our choice of unit. Measurements must be repeatable by others, which means either using a standard measure (meters) or using a nonstandard measure (spear lengths, feet, or inches) defined relative to a meter.

Table 4.3 Standard Multiples of Ratio Scale Units

Name	Multiple	Abbreviation	Example
Pico	10^{-12}	P	pW
Nano	10^{-9}	n	nW
Micro	10^{-6}	μ	μW
Milli	10^{-3}	m	mW
Centi	10^{-2}	c	cW
Deci	10^{-1}	d	dW
	10^{0}		W
Deca	10^{1}	da	daW
Hecto	10^{2}	h	hW
Kilo	10^{3}	k	kW
Mega	10^{6}	M	MW
Giga	10^{9}	G	GW

W = Watt.

Conventional multiples (Table 4.3) serve the useful purpose of creating units appropriate to the scale of a research question; meters are not particularly useful for measuring microbes. Even more can be accomplished by adopting unconventional multiples. If our interest were in the foraging ranges of owls, we might decide to define the range in biological terms, based on the minimum area (in standard units) required to meet daily energy requirements. If we define the minimum area as one unit, we can then examine the problem of foraging area needed by a pair of owls to successfully produce one chick, two chicks, and so on, relative to the number of minimum foraging units. To phrase this as a question, if one owl requires a certain area to meet its own energy needs, how many of these units will be needed by two owls to raise one chick? The answer, in minimum foraging units, leads to a better understanding of the biology than conventional units do (although both are correct). The advantage of unconventional multiples in quantitative reasoning is that they permit reference to biologically defined units.

With one exception, the base units in Table 4.1 are infinitely divisible. The exception is the mole, which is all but infinitely divisible because it is defined by an astronomical number of entities. In biology we are often interested in the dynamics of a small number of entities. A base unit that proves useful again and again in biology is the individual, or entity, for which a convenient symbol is the number sign (#). Examples of biological entities are individuals, cells, species, genes, attacks by a predator, or potential encounters. An *entity* is defined as a recognizable object belonging to a population of such objects. The entity is an unconventional or non-SI unit that is extremely useful in ecology and can be handled in a rigorous fashion (Stahl, 1962). The conventional SI unit is the *mole* (Table 4.1), which is equal to $6.022 \cdot 10^{23}$ entities. The mole is an appropriate unit for chemical entities such as atoms, ions, or molecules. It is far too large for ecological populations. Even the total population of the zooplankter *Calanus finmarchicus*, one of the most abundant species on the planet, does not amount to a picomole.

The philosophical objection to using counts of objects or events as a measurement scale (Ellis, 1966) can be easily met by insisting that this scale does not consist of numbers; it has units of entities (animals, genes, etc.) on a ratio scale of measurement. One distinctive feature of this unit is that we cannot halve it repeatedly in the same way that

we can halve the unit of a centimeter repeatedly. This does not prevent us from calculating expected values in fractions of entities. An example is average family size, which is expressed in fractions of individuals, even though any single family must have a discrete number of individuals.

Unconventional exponents are another source of useful units. For some problems, units of temporal frequency (e.g., sec^{-1}) or spatial frequency (e.g., km^{-1}) are more useful than units of time or distance. Fractal units such as $m^{1.8}$ are more appropriate than Euclidean lines (m^1), planes (m^2), and volumes (m^3) in describing a variety of ecological phenomena, including habitat structure (Pennycuick and Kline, 1986). These will be treated in more detail in Chapter 5.

ANOTHER LOOK AT SECTION 4.4

If there are about five generations per century, re-express the following measurements in generation times:

- One solar year
- One human lifetime of 75 years
- A 5% increase per year

4.5 Rules for Ratio Scale Units

The mathematical rules that apply to units on a ratio scale differ from those that apply to numbers. The rules for ratio scale quantities are few, but they are essential to accurate work. As will become apparent, these rules are also an important part of multiscale analysis. The rules define the operations of addition, subtraction, multiplication, division, exponentiation, and the taking of absolute values for ratio scale quantities. Illegal operations, such as the taking of logarithms, are also listed. Box 4.2 shows a series of example calculations for comparison with a verbal explanation of each rule. The rules are first applied to quantities with the same units, then to quantities with similar units (e.g., meters and centimeters), and finally to quantities with dissimilar units (e.g., days and degrees of temperature).

Rule 1 is that addition changes the number of units but not the unit itself. Rule 2 says that the same thing holds for subtraction. Similar or equal units on a ratio scale can be added (Rule 1) or subtracted (Rule 2). Dissimilar units cannot be added or subtracted. For example, apples and oranges are not similar units; we cannot add them. However, we can define a new unit, "fruit," that allows addition of one group of fruit (all apples) to another group (all oranges). One way of visualizing Rules 1 and 2 is that the same or similar units can be lined up, then counted, without having to omit items (we can count fruit but must omit oranges from a count of apples). Dissimilar units cannot be lined up and counted.

Rule 3 says that units can be multiplied, whether equal, similar, or dissimilar. This rule generates new units that sometimes have a name ($length^2$ = area) and sometimes not ($mass^2$ = ??). The product of dissimilar units is expressed as a hyphenated unit (e.g., degree-days, lizard-hours). One way of visualizing Rule 3 is that it sums one unit with respect to a second unit. For example, summation of a distance over a perpendicular

distance corresponds to an area, summation of exposure to heat over time results in degree-days, and summation of a mutation rate over time measures the total mutations. Many readers will have noticed the resemblance to integration.

Box 4.2 Calculations Based on Rules for Units.

Interpretations are shown to the right, in parentheses.
Same units. (5 rules apply):

1. $3 \cdot \text{cm} + 2 \cdot \text{cm} = (2 + 3) \cdot \text{cm} = 5 \cdot \text{cm}$
2. $7 \cdot \text{trees} - 2 \cdot \text{trees} = (7 - 2) \cdot \text{trees} = 5 \cdot \text{trees}$
3. $\text{cm} \cdot \text{cm}^2 = \text{cm}^3$ (volume)
   $\text{trees} \cdot \text{trees} = \text{trees}^2$ (tree pairs)
4. $\text{cm}/\text{cm} = 1$
5. $(\text{cm}^1)^2 = \text{cm}^2$ (area)

Similar units. (4 rules apply)
It helps to begin with Rule 4, to show the source of the conversion factor 100.

4. $\text{meter}/\text{cm} = 100$
1. $\text{meter} + \text{cm} = (100 + 1) \cdot \text{cm} = 101\,\text{cm}$
2. $\text{meter} - \text{cm} = (100 - 1) \cdot \text{cm} = 99\,\text{cm}$
3. $\text{meter} \cdot \text{cm} = 100\,\text{cm}^2$ (area)

Dissimilar units. (4 rules apply)

1. °K + day ILLEGAL
2. °K − day ILLEGAL
3. °K · day = degree-day (exposure to heating or cooling)
4. °K ÷ day = degrees per day (cooling or heating rate)

Signed units. (1 additional rule). An example is

$$\text{velocity east} = +2\,\text{m} \cdot \text{s}^{-1}$$
$$\text{velocity west} = -2\,\text{m} \cdot \text{s}^{-1}$$

6. $|-2\,\text{m} \cdot \text{s}^{-1}| = 2\,\text{m} \cdot \text{s}^{-1}$ (a speed, not a velocity)

Illegal operations

7. cos(tree) cm! 2^{day} $\log_2(\text{rabbit})$ $\log_{\text{rabbit}}(10)$

Rule 4 (the inverse of Rule 3) is that any unit can be divided by another unit. Rule 4 describes *scaling as a mathematical operation*, that is, taking the ratio of one quantity to another. A unit can be scaled to itself, to a similar unit, or to a dissimilar unit. A unit scaled to itself is equal to unity (1). A unit scaled to a similar unit is a number with no units. For example, a square kilometer is 100 times larger than a hectare, a kilometer is similar to a hectare, and hence a square kilometer scaled to a hectare is a ratio with no units:

$$\text{km}^2/\text{ha} = 100$$

According to rule 4, a quantity can be scaled to either standard or nonstandard units. For example, we can scale the area of an unconventional unit, a pine plantation, to a conventional unit, hectares. This results in a ratio:

$$\text{plantation/hectare} = 10{,}000$$

Or we can scale the area of the plantation to another unconventional unit, the territory defended by a nesting warbler:

$$\text{plantation/territory} = 40{,}000$$

The plantation can be measured by arbitrarily defined units of hectares, but it can also be measured relative to biologically defined units of nesting territories.

When applied to similar units, Rule 4 is the basis for the familiar operation of unit cancellation:

$$1\,\text{km}/10^3\,\text{m} = 1$$

The unitless ratio within the parentheses allows units to cancel out:

$$3000\,\text{m} \cdot 1 = 3000\,\cancel{\text{m}} \cdot \left(\frac{1\,\text{km}}{10^3\,\cancel{\text{m}}}\right) = 3\,\text{km}$$

Cancellation requires similar units.

Rule 4 also applies to dissimilar units. The operation now results in new units rather than in a unitless ratio. For example, the ratio of Joules to seconds is, by definition, a new unit:

$$\text{Joule/second} = \text{Watt}$$

A unit scaled to a dissimilar unit can be interpreted as the differencing of one unit with respect to another—the change in body mass with respect to change in time, for example. This scaling operation can be stated as a question: How much change occurs in the unit of interest, relative to another unit? The result is a new unit, which often has a name. For example, change in location with change in time is a velocity; change in population density with change in location is a density gradient; change in body mass with change in time is a growth rate. Some readers will again have noticed the resemblance to differentiation in calculus.

Rule 5 says that ratio scale units can be raised to any power, including fractional powers, described in the next chapter. This results in a new unit. Interpretable examples come primarily from geometry. For example, taking an area to the 3/2 power results in a volume.

Rule 6 applies to signed units. The most common examples are directions and quantities derived from directions such as velocities. If east is taken as positive, then west is negative. Another example is the accumulation of deficits or surpluses in quantities such as energy or money. To avoid confusion, it helps to assign different names to a signed quantity and the absolute value of the signed quantity. The most familiar example is the absolute value of a velocity, which is a speed.

Not all operations on numbers apply to units (Rule 7). Examples of illegal operations on units are taking logarithms, taking factorials, raising to powers, and applying

trigonometric functions. It is interesting to note that it is possible to develop rules for some of these operations. For example, taking of a factorial can be applied to a scaled quantity. This is a simple extension of multiplication.

$$(N \text{ trees})\,! = (N \text{ trees})(N \text{ trees} - 1 \text{ trees})\ldots = N!\ \text{trees}^{k}$$

This quantity, while definable, is of little utility.

The conventional rules for units do not recognize some operations. This raises a difficulty: it precludes taking the logarithm or powers of a scaled quantity. However, the ratio of similar units is a number (Rule 4) and so the resulting ratios can be taken to powers and expressed as logarithms. Thus, the logarithm of the ratio $(100\,m^2)/(1\,m^2)$ is 2 (see Figure 3.2). The reverse of taking a logarithm, taking a number to a power that is a unit, is also precluded: 2^{hour} has no meaning. But the product of quantities can be used in taking powers. Thus, the product of an instantaneous death rate ($\dot{D}$ = % year^{-1}) and a duration (t = years) results in a number that can be taken as a power (Box 4.1).

The rules for units can be written in the general form shown in Table 4.4 rather than in the specific form of the examples shown in Box 4.2. To write the rules in general form, we need a generic symbol for a unit, 1U. This is a single symbol, not a compound formed by multiplication of U by 1. Further, we need to distinguish between similar and dissimilar units. So, we will say that 1U and 1L are similar and that 1L and 1 M are not similar. Similar units can be added or subtracted, dissimilar units cannot be added or subtracted. For example, if 1U is a unit of mechanical energy and 1L is a unit of thermal

Table 4.4 Rules for Ratio Scale Units

Same Units		
1.	$k{\cdot}1U + n \cdot 1U = (k + n) \cdot 1U$	
2.	$k \cdot 1U - n \cdot 1U = (k - n) \cdot 1U$	
3.	$1U \cdot 1U = 1U^2$	
4.	$1U \div 1U = 1$	
5.	$(1U^{\beta})^{\alpha} = 1U^{\beta\alpha}$	
6.	$\lvert -1U\rvert = \lvert 1U\rvert = 1U$	
Similar Units		
1.	$1L + 1U = (k + 1) \cdot 1U$	
2.	$1L - 1U = (k - 1) \cdot 1U$	
3.	$1L \cdot 1U = k \cdot 1U^2$	
4.	$1L \div 1U = k$	
Dissimilar Units		
1.	$1L + 1M$	ILLEGAL
2.	$1L - 1M$	ILLEGAL
3.	$1L \cdot 1M = 1M \cdot 1L$	
4.	$1L \div 1M = 1L \cdot 1M^{-1}$	
Illegal Operations		
7.	$\cos(1U)$ $1U!$ k^{1U} $\log_k(1U)$ $\log_{1U}(k)$	

Note: See text for definition of symbols.

energy, we can add mechanical and thermal heat together to obtain the total energy, provided we know the conversion factor, which is called the *mechanical equivalent of heat*.

To state the rules in general form, we need some symbols for numbers without units: α, β, n, and k. The symbol k is the unitless ratio of two similar units 1U and 1L. An example is the mechanical equivalent of heat, described previously. With these symbols, the rules for working with ratio scale units can be stated in abstract form, applicable to any units.

Computer languages and commonly available packages typically leave out units. Fortran (or FORmula TRANslator) was one of the first languages for making computations from equations. Fortran, like most languages that followed it, could in principle include units. In practice this was rarely done, in part because of the effort required to format the program results. The rapid spread of graphical formats, where the computer screen displays words, icons, and pictures rather than lower-level programming code, now makes it practical to display units. At least one package (MathCad) displays units with calculations. This package was used to check the accuracy of the examples and computations on scaled quantities in this book. Such packages are an important route to learning how to use scaled quantities. Packages that display units and dimensions (see Chapter 6) are a tremendous aid in learning how to use scaled quantities, rather than numbers, to solve ecological problems.

Defined Concepts and Terms for Review and Future Reference

____ entity
____ derived units
____ scaling as a mathematical operation
____ SI base units
____ standard multiple units
____ standard units

5 Rescaling Quantities

"Oh, how I wish I could shut up like a telescope! I think I could, if I only knew how to begin." For, you see, so many out-of-the-way things had happened lately that Alice had begun to think that very few things indeed were really impossible.

There seemed to be no use in waiting by the little door, so she went back to the table, half hoping she might find another key on it, or at any rate a book of rules for shutting people up like telescopes.

—Lewis Carroll, *Alice's Adventures in Wonderland*, 1865

5.1 Synopsis

Quantities, unlike numbers or mathematical symbols, can be rescaled. This operation has many important uses in ecology. Among these are calibration of instruments, calculation of immeasurable or difficult-to-measure quantities, discovering relations between quantities, and statistical verification of suspected connections. Rescaling a quantity changes both its units and its numerical value, either by remeasurement or by algebraic operations that correspond to remeasurement. Algebraic operations on quantities follow a special set of rules because both units and numerical values change.

Rescaling is as diverse in procedure as it is useful in practice. Logical rescaling changes a quantity from one type of measurement scale to another. An example is categorization to a nominal scale of habitat types from measurements of species abundance on a ratio type of scale. Rescaling via normalization reduces a quantity to a unitless ratio called a scope, which can then be compared to other such ratios. The most common example of normalization is taking a percentage. Rigid rescaling replaces one unit (such as a yardstick) with another (such as a meterstick). There is no change in exponent. Elastic rescaling changes exponents and so in effect stretches or compresses units. Elastic rescaling arises in converting lines to areas, areas to volumes, and fractal units such as convoluted centimeters $cm^{1.3}$ to more familiar Euclidean units such as linear centimeters cm^1.

Rigid scaling factors have diverse sources. Some rigid factors arise by definition, some emerge from calibration, and others result from theory. Many factors in the literature are estimated from data and so are approximate rather than exact. Examples are efficiency of energy transfer between trophic levels (ca. 10%) and the mass-to-volume ratio of organisms, which is close to the density of water, $1000\,kg{\cdot}m^{-3}$.

Elastic factors arise by definition (relation of area to volume). Elastic factors also emerge by estimation from data. Examples are factors such as $meter^{0.3}$, which rescale from straight line units ($meter^1$) to fractal units ($m^{1.3}$). Another source of elastic factors,

doi: 10.1016/B978-0-12-627865-1.00005-2

yet to be tapped, is the theoretical derivation of fractal units, based on the processes that stretch a line, plane, or volume into a fractal over some scope of interest.

The material in this chapter is best learned with a pocket calculator at hand, beginning with the boxes in Section 5.3.

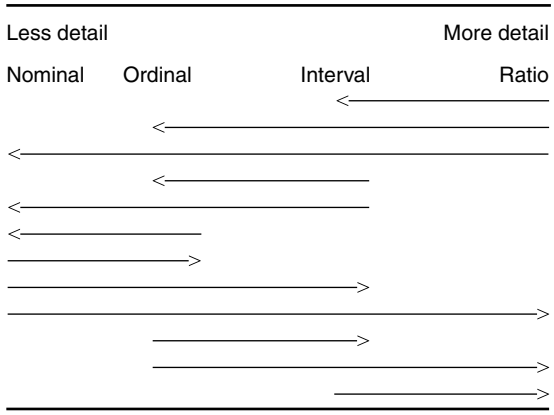
5.2 Logical Rescaling

Logical rescaling changes the type of measurement scale. There are 12 possible rescalings among the four types of measurement scale. Half are in the direction of a less detailed scale, shown as left-pointing arrows in Table 5.1. All these rescalings occur in the ecological literature. All can be executed with standard computer packages for data manipulation.

There are six possible rescalings in the direction of more informative scales, represented by right-pointing arrows in Table 5.1. These logical rescalings require that information be added, either by remeasurement or by combining several quantities to generate a more detailed scale. For example, an interval scale measurement of temperature in degrees Centigrade must be combined with a single-valued quantity, the freezing point of water in degrees Kelvin, to obtain temperature on a ratio scale. Another example is the combination of several nominal scale classifications of habitat (e.g., good/bad, wet/dry, sunny/shady) to produce a single ranking of habitat on a rank scale of, say, 1 to 5. Yet another example is taking the difference of two measurements on an interval scale (such as calendar date) to obtain a ratio scale measurement. Taking a difference guarantees a ratio scale quantity. Taking the ratio of two differences (as in the mathematical operation of differentiation) also guarantees a ratio scale quantity.

Logical rescaling to a less informative scale has many applications. For example, it may be necessary to recalibrate a quantity from a ratio to ordinal or nominal scale if data are uneven in quality. A series of annual observations that began as casual observations, then became more standardized to greater detail over the years, could all be converted to a nominal scale (presence or absence of a phenomenon) that would be consistent across the entire series. Another application of logical rescaling is exploratory data analysis to discover pattern. Rescaling to a less detailed quantity can make

Table 5.1 Logical Rescaling of Quantities



it easier to pick out pattern. For example, a series of satellite images can be remeasured to a nominal scale (presence or absence of weather fronts) to obtain a useful quantity for understanding the effects of weather systems on bird migration (Alerstam, 1990). Rescaling to a nominal scale is used in classification, including taxonomy. Clustering algorithms transform quantities measured on several types of scale (Jardine and Sibson, 1971) to a nominal scale quantity, the classification.

One common application of logical rescaling is the conversion of interval or ratio scale data to a rank type of scale, for statistical evaluation of outcomes via nonparametric methods. The advantage of this, before the common availability of computers, was that all possible outcomes could be tabulated, allowing an exact estimate of a Type I error, the error of accepting a difference that does not exist. Computers now make it possible to use randomization tests (Manly, 1991) to estimate Type I errors without rescaling quantities to ranks. These randomization tests have better discriminating capacity than tests that rescale the data to ranks. In statistical jargon, randomization tests have lower Type II errors than those based on rescaling to ranks. Despite the clear advantages of randomization tests over tests that reduce data to ranks, the rank-based relics have remained in use because they became fossilized in the repertoire of ecologists and they remain available in widely used statistical packages.

Rescaling to a more detailed scale is also useful. An example is *ordination*, which combines several quantities measured on any type of scale into one quantity measured on a ratio scale. The purpose of analysis may be to rank objects, but most ordination techniques produce interval or ratio scale quantities, not ordinal scale quantities. The literature on techniques is vast (Seal, 1964, Kershaw and Looney, 1985; McGarigal et al., 2000), but attention to type of measurement scale is rare (Gower, 1987).

ANOTHER LOOK AT SECTION 5.2

Above each of the 12 arrowheads in Table 5.1, place a check mark if you have used this form of logical rescaling.

5.3 Algebraic Operations on Quantities

Ratio and interval scale quantities are rescaled according to familiar algebraic rules. However, not all the operations possible for numbers apply to scaled quantities. Consequently, some care is needed in carrying out operations on scaled quantities, to avoiding multiplying cabbages by kingfishers, taking the logarithm of fungi, or adding parrots to percentages.

Box 5.1 shows examples of each of 11 rules for working with scaled quantities. In words, the rules are as follows: Quantities are unchanged by adding zero units (Rule 1) or multiplication by 1 (Rule 2). Quantities with the same units can be added or grouped in any way; quantities with unlike units cannot be added (Rules 3, 4, and 6). Quantities can be multiplied in any order to obtain a new quantity, whether or not they have the same units (Rules 5 and 7). The rules for taking powers of a quantity are the same as

those for numbers (Rules 8, 9, and 10). Physically interpretable examples of these last three rules are hard to find for non-integer powers. These three rules will become useful for fractal units, to be described later. The logarithm of a quantity cannot be taken, but the logarithm of a quantity scaled to its base units can be calculated (Rule 11).

Once the rules are grasped from the specific examples in Box 5.1, we can move on to a general statement of the rules. This is done by substituting algebraic symbols for each quantity used in Box 5.1. Thus, QX stands for 3 km, QY stands for 5 km, and so on. Table 5.2 shows the same rules as Box 5.1, this time in abstract notation. This makes the rules harder to grasp, but the notation is a necessary evil in order to state the rules in general form. With an example at hand (Box 5.1) and a list of rules in general form (Table 5.2), the rules can be applied to an unfamiliar situation.

Box 5.1 Computational Rules for Scaled Quantities, Applied to Units of Length

1. $3\,\text{km} + 0\,\text{km} = 3\,\text{km}$
   $3\,\text{km} + 0$ ILLEGAL
2. $3\,\text{km} \cdot 1 = 3\,\text{km}$
   $3\,\text{km} \cdot 1\,\text{km} \neq 3\,\text{km}$
3. $3\,\text{km} + 5\,\text{km} = 5\,\text{km} + 3\,\text{km} = 8\,\text{km}$
   $3\,\text{km} + 10\,\text{days}$ ILLEGAL
4. $3\,\text{km} + (5\,\text{km} + 6\,\text{km}) = (3\,\text{km} + 5\,\text{km}) + 6\,\text{km}$
   $= 14\,\text{km}$
5. $10\,\text{days} \cdot 3\,\text{km} = 3\,\text{km} \cdot 10\,\text{days} = 30\,\text{km-days}$
6. $10\,\text{days} \cdot (3\,\text{km} + 5\,\text{km})$
   $= 10\,\text{days} \cdot 3\,\text{km} + 10\,\text{days} \cdot 5\,\text{km} = 80\,\text{km-days}$
   $3\,\text{km} \cdot (10\,\text{days} + 5\,\text{km})$ ILLEGAL
7. $10\,\text{days} \cdot (3\,\text{km} \cdot 5\,\text{km}) = (10\,\text{days} \cdot 3\,\text{km}) \cdot 5\,\text{km}$
   $= 10\,\text{days} \cdot (15\,\text{km}^2) = 150\,\text{km}^2\text{-days}$
8. $(10\,\text{days} \cdot 3\,\text{km})^{-2} = 10^{-2}\,\text{days}^{-2} \cdot 3^{-2}\,\text{km}^{-2}$
   $= 900^{-1}\,\text{km}^{-2} \cdot \text{days}^{-2}$
9. $3\,\text{km}^1 \cdot 3\,\text{km}^2 = 9\,\text{km}^{1+2} = 9\,\text{km}^3$
10. $(3\,\text{km}^{1.5})^2 = 3^{1.5 \cdot 2}\,\text{km}^{1.5 \cdot 2} = 9\,\text{km}^3$
11. $\log(3\,\text{km}/\text{km}) = \log 3$
    $\log(3\,\text{km})$ ILLEGAL

The general rules in Table 5.2 lead to new ways of looking at familiar quantities. The rules concerning addition and subtraction force us to think about whether two quantities are similar. The rules concerning multiplication, division, and the taking of powers lead to composite quantities, some of which will be unfamiliar. Not all possible combinations of units are biologically interpretable. But often it is possible to interpret composite quantities, which provide a rich quantitative vocabulary for describing and understanding the natural world.

Table 5.2 Rules for Algebraic Operations on Ratio Scale Quantities

QX, QY, and QZ are symbols for quantities that have units 1U.
QT is a symbol for a quantity with units 1T.
1U and 1T are groups of dissimilar units.
α β and γ are symbols for numbers with no units.

$QX \equiv QZ \cdot 1U$.

1. $QX + 0 \cdot 1U = QX$
2. $QX \cdot 1 = QX$
   $QX + 1U = QX$
3. $QX + QY = QY + QX$
4. $QX + (QY + QZ) = (QX + QY) + QZ$
5. $QT \cdot QX = QX \cdot QT$
6. $QT \cdot (QX + QY) = QT \cdot QX + QT \cdot QY$
7. $QT \cdot (QX \cdot QY) = (QT \cdot QX) \cdot QY$
8. $(QT \cdot QX)^{\alpha} = QT^{\alpha} \cdot QX^{\alpha}$
9. $QX^{\beta} \cdot QX^{\gamma} = QX^{\beta+\gamma}$
10. $(QX^{\beta})^{\gamma} = QX^{\beta \cdot \gamma}$
11. $\log(QX/1U) = \log \gamma$

Illegal Operations

$QX + 0$
$QX + QT$
$\log(QX)$

5.4 Application: Individuals, Pairs, Residencies

To see what happens when the rules in Table 5.2 are applied to familiar quantities, let's apply the rules to units of entities, as defined in the fourth section of Chapter 4. Counts of organisms, genes, and other entities are central to population biology. Common quantities in population biology, such as population size, can be handled in a rigorous fashion using the units of entities (refer back to Section 4.4).

When we apply the rules for working with quantities (Table 5.2) to counts of organisms, we obtain a set of new quantities, shown in Box 5.2. All the quantities on the right side of each equation are logically correct because they were calculated according to general rules. Some of these new quantities are easy to interpret, some are harder to interpret, and others (e.g., deer^2) appear biologically uninterpretable but in fact are surprisingly useful.

The first four rules force us to consider whether quantities are similar or not. Rule 1 reminds us that we can add zero deer to 3 deer, but we cannot add the number 0 to 3 deer. Rule 2 reminds us that multiplication by 1 does not change the number of deer, but that multiplication of 3 deer by 1 deer does change the units. Rules 3 and 4 remind us that we can add two groups of deer together in any order, but we cannot add deer to days.

Box 5.2 Computational Rules for Scaled Quantities, Applied to Units of Time and Entities

Examples are for two kinds of units:

$1\mathrm{T}$ = time
$1\#$ = entities

$1\#^2$ = pairs, $1\#^3$ = triplets, $1\#^4$ = quadruplets, etc.
QX = 3deer, QY = 5deer, QZ = 6deer, QT = 10 days
$\alpha = 1$, $\beta = 2$, $\gamma = 3$

1. $3\text{deer} + 0\text{deer} = 3\text{deer}$
   $3\text{deer} + 0$, ILLEGAL
2. $3\text{deer} \cdot 1 = 3\text{deer}$
   $3\text{deer} \cdot 1\text{deer} \neq 3\text{deer}$
3. $3\text{deer} + 5\text{deer} = 5\text{deer} + 3\text{deer}$
   $3\text{deer} + 10\text{ days}$, ILLEGAL
4. $3\text{deer} + (5\text{deer} + 6\text{deer}) = (3\text{deer} + 5\text{deer}) + 6\text{deer}$
   $= 14\text{deer}$
5. $10\text{days} \cdot 3\text{deer} = 3\text{deer}\cdot 10\text{days} = 30$ deer-days
6. $10\text{days} \cdot (3\text{deer} + 5\text{deer}) = 10\text{days} \cdot 3\text{deer} + 10\text{days} \cdot 5\text{deer}$
   $3\text{deer} \cdot (10\text{days} + 5\text{deer})$, ILLEGAL
7. $10\text{days} \cdot (3\text{deer} \cdot 5\text{deer}) = (10\text{days} \cdot 3\text{deer}) \cdot 5\text{deer}$
   $= 10\text{days} \cdot (15\text{pairs}) = 150$ pair-days
8. $(10\text{days} \cdot 3\text{deer})^{-\alpha} = 10\text{days}^{-\alpha} \cdot 3\text{deer}^{-\alpha}$
   $= 30^{-1}\ \text{deer}^{-1} \cdot \text{days}^{-1}$
9. $3\text{deer}^{\beta} \cdot 3\text{deer}^{\gamma} = 3\text{deer}^{\beta+\gamma} = 243\text{quintets} = 243\text{pairs}^{2.5}$
10. $(3\text{deer}^{\beta})^{\gamma} = 3\text{deer}^{\beta \cdot \gamma} = 729\text{triplets}^2 = 729\text{sextets}$
11. $\log(3\text{deer}/\text{deer}) = \log 3$

The remaining rules result in new units, which we shall try to interpret. Rule 5 results in a new unit, the deer-day. This at first may seem unfamiliar, but it is readily interpreted by thinking about a group of deer occupying an area for some period of time. If 3 deer occupy an area for 10 days, then many of their activities, such as food consumption, will be the same as that of a group of 10 deer in an area for 3 days. The product of organism number and time is a measure of residence. If an activity, such as food consumption, depends on residence, we can measure this activity in deer-days. Rule 6 tells us that we can add units of deer-days, or residence, in any order.

Rule 7 results in a strange unit, deer^2-days, the product of a familiar unit and an apparently meaningless unit, deer^2. This unit seems to have no biological interpretation, because we are accustomed to thinking of squaring units as the operation of multiplying two lengths, at right angles to one another, to generate an area. In this context deer^2 makes no sense. However, we can make sense of this unit if we recall that the number of potential pairwise interactions in a group rises with the square of group size. The functional expression for calculating the potential number of pairs in a group of size N is:

$$\text{Duo}(N) \equiv \tfrac{1}{2} \cdot N \cdot (N - 1) \tag{5.1}$$

This is the conventional notation, which lacks units. The numeral 1 has no units, hence the symbol N must also be a number with no units. Consequently, $N(N-1)$ and Duo*(N)* are also numbers. We are working with deer, not numbers, so let's rewrite the formula for deer, with a group size of Q having units of entities #:

$$\text{Duo}(Q) \equiv \tfrac{1}{2} \cdot Q \cdot (Q - 1\#) \tag{5.2}$$

The function Duo*(Q)* shows how to calculate the number of potential pairs of deer in a group of size Q, which has units of deer. In this formula Q must have the same units as #, which in this case are deer. Duo*(Q)* will have units of deer^2. Deer^2 is thus a unit that measures the number of potential pairs. Similarly, if there are 5 alleles in a population, then there are a possible $5 \cdot (5-1)/2$ zygotes. Zygotes have units of alleles^2: (5 alleles · 4 alleles)/2 = 10 alleles^2, not 10 alleles.

A hand calculator helps in understanding this the concept of a Duo, defined as the number of potential pairs in a population. At this point, try calculating:

$$\text{Duo}(3 \text{ deer}) = \_\_\_\_\text{pairs} \qquad \text{Duo}(5 \text{ deer}) = \_\_\_\_\text{pairs}$$

Returning now to deer^2-days, we find that this strange unit can be interpreted as the potential number of pairs of deer that can form over a period of time. It is a unit that can be visualized and interpreted in terms of the behavior of deer.

Rule 8 results in another strange unit, the deer^{-1}-day^{-1}. To interpret this, we express it relative to something we know, deer-days:

$$\text{deer}^{-1}\text{-day}^{-1} = (\text{deer-day})^{-1}$$

This new unit is read "per deer-day." This unit is useful in working with quantities related to residence by a population. For example, we may be interested in the number of twigs browsed per deer-day by a resident group of deer.

Rule 9 results in another strange unit, deer^3. Again, we cannot interpret this unit relative to the usual geometric notion of cubes as the product of three lengths, all at right angles. We can interpret deer^3 as the number of potential triplets, in much the same way that deer^2 measures the number of potential pairs. The functional expression for the number of potential triplets in a group of Q organisms is:

$$\text{Trio}(Q) = \frac{1}{6} \cdot Q \cdot (Q - 1\#) \cdot (Q - 2\#) \tag{5.3}$$

Trios have units of entities^3. We can interpret a quantity having units of deer^3 as the number of potential deer trios. Trios may not be an important part of the biology of deer, but in other groups, such as colonial seabirds, trios are important during the breeding season. The number of new birds added to the population depends on the number of trios (two parents and an offspring). Duos (one parent and one offspring) usually fail to contribute to the next generation because of predation on unguarded chicks, inadequate food supply to chicks, or both.

Rule 10 results in still higher powers that have no obvious biological interpretation at the level of populations of individuals. Higher powers, such as entities^4, can take on meaning at the level of gene combinations in populations.

Rule 11 reminds us that we can take the logarithm of a unitless ratio but that the logarithm of a unit is not defined relative to a unitless base such as 2, e, or 10.

The examples in Box 5.2 show that higher powers of the unit entities # are just as interpretable as higher powers of a unit of length. Units such as $\#^2$ or $\#^3$ are initially strange but turn out to be visualizable as $\#^2$ = pairwise contacts, $\#^3$ = trios, $\#^4$ = quartets, and so on. The process of interpreting these new quantities involved assigning names, making calculations, and visualizing the new quantities relative to the biology of the component quantities. This application of the rules for operations on quantities showed that new ways of thinking result when scaled units, rather than just numbers, are used in quantitative ecology. This application of the rules for operations showed how the use of scaled quantities incorporates ecological reasoning, unlike quantitative ecology based on numbers devoid of units and scale.

ANOTHER LOOK AT SECTION 5.3

Secondary production is reported as grams of carbon fixed per unit area per year: g-C km^{-2} yr^{-1}. Typically this will be computed from densities: animals km^{-2}. What are the units of the quantity needed to compute secondary production from density?

5.5 Rescaling Via Normalization

We can reduce a scaled quantity to a ratio with no units by applying any of several forms of rescaling. *Simple rescaling* occurs when we divide a value by a reference value. The generic expression for simple rescaling is:

$$\left(\frac{Q}{Q_{ref}}\right)^{\beta}$$

The ratio has no units. The magnitude of a scaled quantity thus becomes independent of the units of measurement. Simple rescaling is based on comparison of objects with one another, rather than iterative measurement of the same object. It allows us to substitute one measurement unit for another, as in Galileo's use of spearlengths to measure velocity. It is the basis of classical dimensional analysis (Chapter 6).

A convenient reference quantity Q_{ref} is the largest observed or largest possible value, resulting in a variable that has been reduced to a range from 0 to 1. An example is running speed measured relative to the maximum for that species. Yet another useful reference quantity is Q_{min}, the minimum observed or possible value, which yields a reduced variable that ranges upward from 1. An example is metabolic rate as a multiple of the standard metabolic rate SMR, which is measured at rest and in the absence of absorptive activity by the gut. Reduction relative to Q_{min} expresses the quantity Q in steps that are relevant to that variable. In physiology the reduced quantity Q_{max}/Q_{min} is called a *scope*.

The reference quantity can be chosen on biological or physical grounds. A physical example is subsurface illumination in a lake relative to the illumination at the surface. A biological example is photosynthetic rate relative to the maximum rate. Reference quantities are

also chosen on statistical grounds. An example is ranging (Sneath and Sokal, 1973), which uses both the minimum and maximum value to reduce the quantity to the range 0 to 1.

$$Q' \equiv \frac{Q - Q_{\text{min}}}{Q_{\text{max}} - Q_{\text{min}}} \tag{5.4}$$

We can reduce a scaled quantity to a ratio with no units by *normalizing*. Normalizing occurs when we divide by a summary statistic, which introduces an iterative component to the operation. The generic expression for normalizing is:

$$\left(\frac{Q}{Q_o}\right)^{\alpha}$$

The most familiar example is taking a percentage: adding up the parts to compute the whole, then taking each part as a ratio relative to the whole. For a percentage, the reference quantity Q_o is the sum of all the values of Q. The exponent is $\alpha = 1$, resulting in dimensionless values that can range between 0 and 1.

An example of normalization according to a statistical criterion is the *normal score,* found in any book in statistical methods:

$$z' \equiv \frac{Q - \text{mean}(Q)}{\text{stdev}(Q)} \tag{5.5}$$

This is an example of normalization (in the scaling sense) because measurement is iterative: The mean and the standard deviation are computed from the same set of values. The normal score of a scaled quantity is unitless because both the mean and the standard deviation have the same units as the measurements used to compute these statistics. Normal scores permit comparison of quantities that differ in magnitude and variability. They reduce any quantity to a mean value of zero with a standard deviation of unity. Legendre and Legendre (1998) discuss applications and potential problems of this and other statistical reduction to dimensionless ratios.

The examples of normalization so far have been for each value of a variable. Normalization is also applied to measures of variability, resulting in a single ratio. The most familiar example is the coefficient of variation.

$$CV \equiv \frac{\text{stdev}(Q)}{\text{mean}(Q)} \tag{5.6}$$

The coefficient of variation is a unitless ratio that permits comparison of the variability of two scaled quantities, free of the effects of choice of measurement scale.

ANOTHER LOOK AT SECTION 5.4

List the several forms of normalization described in this section. For each, state whether you have used this form. List any additional normalization operations that you have used.

5.6 Rigid Rescaling

Rigid rescaling replaces one unit with another, either by remeasurement or by calculation based on calibration factors. Rescaling via remeasurement can be visualized as lining up small units into larger units or dividing large units into equal subunits. An example is using a 10 meter wire to mark off plots 100 m on a side, then using 1 meter paces to find locations along the perimeter of the plot. Rescaling from 1 decameter to 1 m units can be viewed as cutting a wire into smaller units. Another example is unit replacement of areas: An area of 1 hectare can be broken into exactly 100^2 squares, each of which is a meter on a side.

Rigid conversion factors consist of a fixed ratio between a large and a small unit. If the smaller unit occurs in the numerator, the factor represents the operation of breaking large units into smaller units. If the larger unit occurs in the numerator, the rigid factor represents the operation of aligning small units into a larger unit.

Here is a simple example of rigid rescaling. In this example, units "cancel out" because any unit scaled to itself is one: $\frac{\cancel{\text{m}}}{\cancel{\text{m}}} = 1$.

$$700\ \cancel{\text{yards}} \cdot \frac{0.9144\ \cancel{\text{m}}}{\cancel{\text{yards}}} \cdot \frac{1\text{m}}{1000\ \cancel{\text{m}}} \Rightarrow 0.64\ \text{km} \tag{5.7}$$

The symbol $\Rightarrow$ is read "calculated as". It indicates that the quantity Q_{final} at the end of the arrow is calculated from the quantity Q_{old} and two conversion factors at the origin of the arrow.

This procedure will be familiar to most readers. Table 5.3 lists a general recipe for rigid rescaling. The equation in Table 5.3 will not be familiar, so to explain the expression it has been aligned with a specific calculation in Box 5.3. The rigid conversion factors k_1 and k_2 rescale Q_{old} to a new quantity, Q_{final}. Conversion factors are listed in sequence so that the denominator of the first factor cancels the units of Q_{old} and the denominator of the next factor cancels the numerator of the preceding factor. Box 5.4 shows the derivation of the expression for rigid rescaling from the noniterative scaling relation introduced in Chapter 2.

Box 5.3 Rigid Rescaling of Quantities. Exponent = 1.

$$\begin{aligned} Q_{old} \quad & \cdot k_1 \quad & \cdot k_2 \quad & = Q_{final} \\ Q_{old} \quad & \cdot \frac{\text{k}_1\text{newunits}}{\text{old unit}} & \cdot \frac{\text{k}_2\text{finalunits}}{\text{newunits}} \quad & = Q_{final} \\ 700\ \text{yards} & \cdot \frac{0.9144\ \text{m}}{\text{yard}} & \cdot \frac{1\ \text{km}}{1000\ \text{m}} \quad & \Rightarrow 0.64\ \text{km} \end{aligned}$$

Table 5.3 Rigid Rescaling of Quantities

The sequence of steps in rigid rescaling is as follows:

1. Write the quantity $(Q_{old})^{\alpha}$ to be rescaled,
2. Apply rigid conversion factors k_1^{α}, k_2^{α}, k_3^{α}, etc. so that units cancel.
3. Complete the calculation of Q_{final}^{α}, with appropriate exponents.

The generic expression for rigid rescaling is:

$$Q_{old}^{\alpha} \cdot k_1^{\alpha} \cdot k_2^{\alpha} = Q_{final}^{\alpha}$$

Rigid rescaling does not change unit exponents.

Hence Q_{old}^{α} and Q_{final}^{α} must have the same exponent.

Box 5.4 Derivation of Generic Expression for Rigid Rescaling

Rigid rescaling arises from the noniterative measurement relation (Equation 2.6a) with exponent $\beta = 1$.

$$\frac{Q}{Q_{ref}} = \left(\frac{Unit}{Unit_{ref}}\right)^1$$

Equivalently,

$$Q = Q_{ref} \cdot \left(\frac{Unit}{Unit_{ref}}\right)^1$$

Taking Q as Q_{final}, Q_{ref} as Q_{old}, $(Unit/Unit_{ref}) = k$, and applying the exponent α, this becomes:

$$Q_{final}^{\alpha} = Q_{old}^{\alpha} \cdot k^{\alpha}$$

The reason for stating a generic expression is to show how to handle exponents other than $\alpha = 1$. Box 5.5 shows rigid rescaling with a familiar exponent of 2. Box 5.6 extends this to rigid rescaling for a fractal quantity, a length with nonintegral exponent ($km^{1.2}$).

Box 5.5 Rigid Rescaling of Quantities. Exponent = 2.

Exponents are applied to both units and numbers, not just to the numbers.

$$Q_{old}{}^2 = (700 \cdot \text{yard})^2 = 700^2 \text{yard}^2 \neq 700 \text{ yard}^2$$

Apply the exponent to obtain a conversion factor that will "cancel" units of $Q_{old}{}^2$:

$$k_1{}^2 = \left(\frac{0.9144 \text{ meter}}{\text{yard}}\right)^2 = \frac{0.9144^2 \text{m}^2}{\text{yard}^2}$$

Then apply exponents after lining up the conversion factors:

$$(700\text{yard})^2 \cdot \frac{0.9144 \text{ m}^2}{\text{yard}^2} \cdot \frac{\text{km}^2}{1000^2 \text{m}^2} \Rightarrow 0.41 \text{ km}$$

$$Q_{old}{}^2 \cdot \quad k_1{}^2 \quad \cdot \quad k_2{}^2 \quad = Q_{final}{}^2$$

Box 5.6 Rigid Rescaling of Quantities. Exponent = 1.2.

Noninteger exponents are handled the same way as integer exponents.

$$(700 \text{ yards})^{1.2} \cdot \frac{0.9144^{1.2} \text{ m}^{1.2}}{\text{yard}^2} \cdot \frac{\text{km}^{1.2}}{1000^2 \text{ m}^2} \Rightarrow 0.59 \text{ km}^{1.2}$$

$$Q_{old}{}^{1.2} \cdot \quad k_1{}^{1.2} \quad \cdot \quad k_2{}^{1.2} \quad = Q_{final}{}^{1.2}$$

Rigid factors are written as a symbol representing a ratio of units:

$$\frac{1 \text{ Joule}}{4.187 \text{ cal}} = k_{\text{Joule/cal}} \tag{5.8}$$

The reason for writing a rigid factor as a ratio is that units are converted by multiplication, not by substitution. It might seem that the relation of calories to Joules could be written 1 Joule = 4.187 cal, but this can only lead to error by encouraging substitution rather than multiplication to rescale quantities. The secret of success in rigid rescaling is to apply ratios that cancel units and to make sure that the exponents allow units to cancel.

Rigid scaling factors, which convert one quantity to another, have several sources. They often arise by defining a unit at one resolution as a multiple of a unit at a finer resolution. Thus a rigid factor $k_{g/Mg}$ is, by definition:

$$\frac{10^6 \text{ gram}}{1 \text{ Megagram}} \equiv k_{g/Mg} \tag{5.9}$$

Rigid factors also arise from definition of measurement units. The definition of a Watt is:

$$\text{Watt} \equiv \text{Joule} \cdot \text{s}^{-1} \tag{5.10a}$$

and consequently the rigid conversion factor is:

$$\frac{1 \text{ Joule} \cdot \text{s}^{-1}}{1 \text{ Watt}} = k_{\text{Joule/Watt-s}} \tag{5.10b}$$

Some rigid factors are precisely measured and considered to hold regardless of circumstance. For example, 1 unit of heat energy (1 cal) is equal to 4.187 units of mechanical energy, where each unit of mechanical energy is a Newton · m:

$$\frac{4.187 \text{ Newton} \cdot \text{m}}{1 \text{ cal}} \equiv k_{\text{work/heat}} \tag{5.11}$$

Many rigid factors are estimated from data. Because these factors are completely empirical, their applicability depends on circumstance. An example is calibrating a satellite image against a measure of vegetation cover. The calibration at a particular location could be used in similar circumstances but could not be applied to a satellite image from anywhere in the world.

In ecology it is useful to define rigid factors that are conditionally true rather than universally true. The symbol := is used to indicate an equality that is true under limited conditions. The symbol := is read "conditionally equal to". An example of a factor that holds conditionally is the density (mass per unit volume) of living organisms:

$$\frac{1 \text{ m}^3}{1000 \text{ kg}} := k_{\text{vol/mass}} \tag{5.12}$$

Most organisms have densities close to this value, even though many do not have exactly this value. There are exceptions, of course, such as benthic organisms with calcareous shells. Because the ratio is so useful, it is worth keeping the few exceptions in mind to be able to use it.

Another ratio that has a narrow enough scope to be worth using as a rigid scaling factor is the ratio of biomass consumed by a population $\dot{M}_{\text{in}}$ to the biomass transferred to higher trophic levels $\dot{M}_{\text{out}}$:

$$\frac{\dot{M}_{in}}{10\,\dot{M}_{out}} := k_{in/out} \tag{5.13}$$

The transfer efficiency between trophic levels is not fixed at 10%, but the scope of this ratio is small enough [(ca. 20%)/(ca. 5%) = 4] that it is useful in making order of magnitude calculations of production at one trophic level from measurements at another level.

ANOTHER LOOK AT SECTION 5.5

Write out the generic expression in Table 5.3 for rigid rescaling. Then write out the sequence of conversion factors required to compute the number of eagle nests per $km^{1.3}$ of coastline, if there are five nests per nautical $mile^{1.3}$. (A nautical mile = 1.8652 km)

5.7 Elastic Rescaling

Elastic rescaling changes the exponent of a ratio scale quantity. Both the unit and the numerical outcome acquire a new exponent. We can visualize this as a systematic stretching or compression of the unit of measurement. To illustrate the idea, let's look for an alternative to carrying around a rigid and bulky frame of 1 meter on each side to count plants in areas of fixed size. Instead, let's attach a stake on a swivel to one end of a meterstick, carry this tool to the study site, push the swivel into the ground at a point, then swing the stick in a circle around the swivel, counting plants as they pass under the meterstick. This sweeps an area, an activity that can be visualized as stretching a one-dimensional object (the meterstick) into two spatial dimensions (the circular area swept). Another way of stretching a meterstick out into an area is to set it down, then pull it at right angles to its length, to generate a rectangular area. The calculation that corresponds to stretching a meterstick over a distance of 2 m to generate an area is:

$$1\ \text{meter}^1 \cdot (2\ \text{meter})^{2-1} \Rightarrow 2\text{meter}^2$$

To make reliable calculations, we require a generic expression for elastic rescaling. The advantage of a generic expression is that it can be applied to any situation. The disadvantage is that it is hard to grasp on first encounter. So, once again we line it up with a calculation, to allow comparison with a known case:

$$\begin{aligned} Q^{old} \cdot k^{new-old} &= Q^{new} \\ 2\ \text{meter}^1 \cdot (1\,\text{meter})^{2-1} &\Rightarrow 2\,\text{meter}^2 \end{aligned} \tag{5.14}$$

The *elastic scaling factor* $k^{new-old}$ stretches or compresses units. It measures the degree of stretching of units with change in measurement frequency.

Table 5.4 lists the generic recipe for elastic rescaling of quantities. If the elastic scaling factor has the same units as Q^{old}, we can divide the equation in Table 5.4 by the units to obtain a version of the equation in unitless form. This form allows us to take logarithms (exponents), which are:

$$old + (new - old) = new \tag{5.15}$$

Elastic rescaling uses this relation to either "stretch" or "contract" old units into new units. The elastic factor $k^{new-old}$ expresses the degree of stretching (if $new > old$) or the degree of shrinking (if $new < old$). The generic recipe for elastic rescaling in Table 5.4 derives from the noniterative measurement relation introduced in Chapter 2. Box 5.7 shows the derivation.

Table 5.4 Elastic Rescaling of Quantities

The generic expression for elastic rescaling of the quantity Q_{old} by an elastic scaling factor $k^{new-old}$ is:

$$Q^{old} \cdot k^{new-old} = Q^{new}$$

The steps in elastic rescaling are:

1. Write the generic expression for elastic rescaling.
2. Substitute quantities and factors into the expression.
3. If quantities and factors do not have same units, use rigid rescaling on either quantities or factors.
4. Compute Q^{new}.

From the derivation (Box 5.7), it is clear that the elastic scaling factor k in Table 5.4 must have the same units as Q, the quantity to be rescaled. If the elastic scaling factor $k^{new-old}$ does not have the same units as Q, rigid conversion must be used to obtain one. To illustrate this idea, we place a 2 m long stick on the ground, then pull it sideways for 1 cm to generate a rectangular area. The distance pulled ($k = 1$ cm) has different units from the quantity being stretched (2 m long stick). The elastic factor is $(1\,\text{cm})^{2-1}$, which does not match the units of $Q^{old} = 2$ m. This is accomplished by rigid conversion *before* elastic rescaling. Box 5.8 shows the sequence of calculations for stretching a 2 m length sideways for 1 cm to generate a rectangular area.

Box 5.7 Computational Formula for Elastic Rescaling, Derived from the Noniterative Measurement Relation, Equation 2.6a

For heterogeneous exponents the measurement relation becomes:

$$\left(\frac{Q}{Q_{ref}}\right) = \left(\frac{L}{L_{ref}}\right)^{\alpha/\gamma}$$

where $\beta = \alpha/\gamma$.

This rearranges to:

$$Q^{\gamma} = {Q_{ref}}^{\gamma}\ {L_{ref}}^{-\alpha} L^{\alpha}$$

Where Q is measured in the same units as L, this becomes:

$$Q^{\gamma} = {Q_{ref}}^{\gamma} {Q_{ref}}^{-\alpha}\ Q^{\alpha}$$

Hence:

$$Q^{\gamma} = {Q_{ref}}^{\gamma-\alpha} Q^{\alpha}$$

Taking $\alpha = old$, $\gamma = new$, and fixing $Q_{ref} = k$:

$$Q^{new} = Q^{old} \cdot k^{new-old}$$

Box 5.8 Elastic Rescaling of $Q = 2$ Meters

Substitute the quantity to be rescaled and one or more elastic scaling factors into the generic expression for elastic rescaling:

$$Q^{old} \cdot k^{new-old} = Q^{new}$$

$$2 \text{ meters} \cdot (1\text{cm})^{2-1} = ?$$

Units do not match, so use rigid rescaling to obtain a new elastic scaling factor that has the same units:

$$1 \text{ cm} \cdot \left(\frac{0.01 \text{ meter}}{\text{cm}}\right)^{2-1} \Rightarrow (0.01 \text{ meter})^{2-1}$$

Calculate Q^{new}:

$$\begin{aligned} 2 \text{ m}^1 \cdot (1 \text{ cm})^{2-1} &= 2 \text{ m}^1 \cdot (0.01 \text{ m})^{2-1} \\ &= 2 \text{ m}^1 \cdot 0.01^{2-1} \cdot \text{m}^{2-1} \\ &\Rightarrow 0.02 \text{ m}^2 \end{aligned}$$

Altering the exponent of a rigid factor (e.g., $10^2 \text{cm} \cdot \text{m}^{-1}$) will not work, because this does not represent the operation of stretching a quantity:

$$2 \text{ meter}^1 \cdot \frac{(10^2 \text{ cm})^{2-1}}{\text{meter}^{2-1}} \Rightarrow 2 \text{ meter} \neq 0.02 \text{ meter}^2$$

The idea of stretching a unit seems unusual compared to familiar ways of visualizing areas relative to lines or visualizing accelerations relative to velocities. The idea of stretching units will aid considerably, however, in visualizing and working with fractal rather than integral exponents. Fractal exponents have proved to be a logically consistent way of measuring natural objects, from clouds and rivers to ecotones and blood vessels.

ANOTHER LOOK AT SECTION 5.6

Beginning with an area of $4\,\text{m}^2$, compress it by an elastic or reshaping factor of $(4\,\text{m})^{1-2}$ to obtain the line length after compression. Then compress the 4 m^2 area by an elastic factor of $(2\,\text{m})^{1-2}$ to obtain the line length after compression.

5.8 Elastic Rescaling with Fractal Exponents

Once we can visualize a familiar procedure (calculating an area) as the stretching of units, we can extend it to less familiar situations, such as nonintegral (i.e., fractal) stretching of units. The idea of stretching units will allow us to visualize and work accurately with $1\,\text{cm}^{1.4}$ as readily as $1\,\text{cm}^2$.

To illustrate elastic rescaling with nonintegral exponents, let's use a rubber band to measure the length of a natural (and hence typically crooked) object, a tree root. We pin each end of the rubber band to the extreme ends of the root to obtain a straight line distance. We then pin the rubber band against the root halfway between the first two pins. This stretches the rubber band as the measurement frequency changes from one measurement to 2. We again pin the rubber band against the root halfway between the existing pins, further stretching the rubber band as the measurement frequency increases to 4. We continue this procedure, keeping track of the amount of stretching with each change in measurement frequency. (This is a rubber band that changes color as it stretches.) We stop at a resolution (inner scale) set by the limits of the instrument—the pins will start interfering with the measurements at very close spacings. If we can escape instrumental limits on resolution, we will still meet a lower limit set by the size of cells; a root is no longer a root at the scale of a cell or less.

If we apply this unusual measurement instrument, a rubber band, to a linear object such as a board, no stretching will occur with successive pinning. This means that the length is a quantity with an exponent of unity. If we wrap the rubber band around a circle or around a regular polygon, no stretching will occur with successive pinning. Hence the perimeters of these objects are quantities with exponents of unity. If we measure a natural object, such as a tree root, we will find that the rubber band stretches in a regular way with each doubling of measurement frequency. Consequently, we can express the degree of stretching by the increase in the exponent of the quantity beyond unity. If the tree root is slightly crooked at all scales, the stretching will be slight at each pinning. Thus, the length is a quantity with an exponent slightly greater than unity. If the root is extremely crooked, the stretching will be considerable and the exponent of the quantity will increase toward 2. A root that is $4\,\text{cm}^{1.2}$ is slightly more crooked than any straight ruler we can choose. A root that is $4\,\text{cm}^{1.8}$ is far more crooked than a ruler of any length. But the exponent is still less than 2. It is not so crooked as to fill an area.

The idea of using rubber bands for measurement seems strange and of questionable value. But after a time it becomes familiar and indeed even seems the appropriate way of measuring crooked objects such as rivers, coastlines, territorial areas, and other convoluted features of landscapes, lakes, or seascapes. The concept of representing a crooked or convoluted object with a nonintegral *fractal exponent* allows us to compute the length of an object with respect to any unit of measure we choose. Let's begin with a moderately crooked reach of a river that is $2\,\text{km}^{1.3}$. How long is this reach, measured in convoluted meters $\text{m}^{1.3}$, rather than convoluted kilometers $\text{km}^{1.3}$? The reach does not have a length of $2000\,\text{m}^{1.3}$. The reach is:

$$2\,(1000\,\text{m})^{1.3} = 2 \cdot 1000^{1.3} \cdot \text{m}^{1.3} = 15887\,\text{m}^{1.3} \tag{5.16}$$

To make this calculation we use rigid rescaling, for there is no change in exponent: $m^{1.3}$ has the same exponent as $km^{1.3}$. A kilometer$^{1.3}$ is a convoluted length that was used to measure a convoluted object, and a meter$^{1.3}$ is an equally convoluted length for measuring the same object. The calculation according to the rules for rigid rescaling shows us that there are 15,887 convoluted meters, for which the symbol is $m^{1.3}$, in an equally convoluted kilometer, $km^{1.3}$.

Box 5.9 Elastic Rescaling of Quantities, Noninteger Exponent

Substitute quantities and elastic scaling factors into the generic expression for elastic rescaling:

$$\begin{aligned} Q^{old} \quad \cdot k^{new-old} &= Q^{new} \\ 3\ \mathrm{km}^{1.3} \cdot (1\mathrm{m})^{1-1.3} &= \underline{\quad ? \quad}\ \mathrm{km}^{1} \end{aligned}$$

Units do not match, so use rigid rescaling to obtain new elastic scaling factor:

$$1\ \mathrm{m}^{1-1.3} \cdot \left(\frac{0.001\ \mathrm{km}}{\mathrm{m}}\right)^{1-1.3} = (0.001\ \mathrm{km})^{1-1.3}$$

Calculate Q^{new}:

$$\begin{aligned} 3\ \mathrm{km}^{1.3} \cdot (0.001\ \mathrm{km})^{1-1.3} &= 3\mathrm{km}^{1.3} \cdot (0.001\ \mathrm{km})^{1-1.3} \\ &= 3\ \mathrm{km}^{1.3} \cdot 0.001^{-0.3} \cdot \mathrm{km}^{-0.3} \\ &= 3 \cdot 7.943 \cdot \mathrm{km}^{1} \\ &\Rightarrow 23829\ \mathrm{m}^{1} \end{aligned}$$

An elastic scaling factor of $k = \mathrm{km}^{1-1.3}$ will not work because this does not represent the operation of flattening a convoluted kilometer $km^{1.3}$ into flat meters m^1:

$$3\ \mathrm{km}^{1.3} \cdot \mathrm{km}^{1-1.3} \Rightarrow 3\ \mathrm{km}^{1} \neq 23.830\ \mathrm{km}^{1}$$

$km^{1-1.3}$ represents the operation of flattening a convoluted kilometer $km^{1.3}$ into a flat kilometer km^1.

How long is the reach of a river in straight meters m^1 rather than crooked meters $m^{1.3}$? To calculate this problem we use elastic rescaling because we are going to flatten or straighten the convoluted unit of a kilometer$^{1.3}$ into units of linear meters m^1.

The exponent changes and so elastic scaling is required. Box 5.9 shows computations for elastic rescaling of a convoluted reach of river into straight line measurements. The computations follow the steps in Table 5.4.

Until recently, the exponents for units in the natural sciences were boring. Mass units had exponents of one (kg^1 or kg^{-1}). Time units had exponents of one (day^1, day^{-1}) or two (e.g., acceleration of % day^{-2}). Length units had the greatest diversity: Exponents ranged from 3 (volumes) through 2 (areas) and 1 (lines) to −1 (per unit length), −2 (per unit area), and −3 (per unit volume). The concept of a *fractal* (Mandelbrot, 1977), based on Hausdorff's (1919) measure of complexity, allows unit exponents in the natural sciences to take on noninteger values, which are no less interpretable than integral exponents of 1, 2, or 3. Elastic rescaling provides a physical interpretation of these noninteger exponents. The elastic factor expresses the degree of stretching of a line into a convoluted boundaries such as a coastline or ecotone. The elastic factor expresses the degree of stretching of a two-dimensional measurement unit (a flat but elastic sheet) applied to a convoluted surface such as the surface area within a soil, the surface area of a cloud, or the surface area of a lung.

The fundamental notion underlying the measurement of complex phenomena is that a measured quantity Q scales as a power law function of the resolution (S = step size) of measurement:

$$Q = k \cdot S^{1-D} \tag{5.17}$$

Treatments of fractals use this expression or a variant (Mandelbrot, 1977; Hastings and Sugihara, 1993). Box 5.10 demonstrates that behind this expression lies iterative measurement, a concept introduced in Chapter 2:

$$\frac{Q}{Q_o} = \left(\frac{S}{S_o}\right)^{1-D} \tag{5.18}$$

In this expression, S is step size relative to some reference step size S_o. Q is the measured quantity at step size S, whereas Q_o is the measured quantity at the reference step size.

Expression 5.18 in turn rests on an *iterative counting relation:*

$$\frac{n}{n_o} = \left(\frac{S}{S_o}\right)^{-D} \tag{5.19}$$

In this expression, n is the count of units at step size S, whereas n_o is the number of units at the reference step size S_o. D is the number that renders the two ratios equal. Mandelbrot (1977) showed that a wide variety of natural phenomena are appropriately measured with exponents D that are not the familiar integers of Euclidean geometry.

Box 5.10 Complex Phenomena Are Described by a Power Law Based on Iterative Measurement

Iterative measurement is based on an iterative counting relation, which relates the number of units n to the scale of measurement via an exponent called the *dimension*;

$$\left(\frac{n}{n_o}\right) = \left(\frac{S}{S_o}\right)^{-D}$$

For the straight line distance between two points using two scales (e.g., S in meters and S_o in centimeters), the dimension is $D = 1$.

The reference scale S_o can be either an extent or a resolution. If S_o is an extent, then S represents increasingly smaller subsections of S_o. In this case S/S_o is a fraction, the frequency of measurement. If S_o is a resolution, then S/S_o represents increasing larger units or measurement as multiples of S_o.

Hausdorff (1919) expanded the definition of D to include complex objects where D exceeds the Euclidean measure of an object embedded in Euclidean space (line, grid, or volume).

A quantity consists of the count of units by unit size. The iterative scaling relation for the quantity is:

$$\frac{Q}{Q_o} = \left(\frac{n \cdot S}{n_o \cdot S_o}\right) = \left(\frac{S}{S_o}\right)^{1-D}$$

This rearranges to a scaling function:

$$Q = \left(Q_o \cdot S_o{}^{-1+D}\right) S^{1-D}$$

Taking the fixed ratio:

$$k = \left(Q_o \cdot S_o{}^{-1+D}\right)$$

results in a power law:

$$Q = k \cdot S^{1-D}$$

When Q and L are expressed in the same units, the factor k becomes the elastic scaling factor (refer back to Box 5.7).

A counting relation such as that in Equation 5.19 gives the number of steps n at one scale (step size S or ruler size L), compared to the number n_o at some reference scale (step size S_o or ruler size L_o). The exponent D quantifies the result of changing the unit of measurement from S_o to S. With Euclidean measurement protocols (i.e., count steps

on a straight line, count boxes on a Cartesian grid), the exponent D is unity and number of counts scales as the inverse of unit length. If we double unit length, we halve the number of steps. As a result, the scaled quantity of interest is independent of the units we use.

Not all counting relations have an exponent of unity. If we want to measure the length of a coastline, we will obtain different answers (ratio of n/n_o) depending on the ratio of unit size S/S_o. To obtain a consistent measurement relation, we need to use an exponent other than unity. The iterative counting relation for coastline length will typically have an exponent on the order of 1.3, as follows:

$$\frac{Nsteps}{Nsteps_o} = \left(\frac{L}{L_o}\right)^{-1.3} \tag{5.20a}$$

This relation quantifies the readily grasped concept that as we increase the size of our reference unit L/L_o the number of steps taken along the coast will decrease.

$$\frac{1\,\text{giantstep}}{400\,\text{babysteps}} = \left(\frac{10\text{ m}}{0.1\text{ m}}\right)^{-1.3} \tag{5.20b}$$

Units of time are appropriately interpreted with fractal exponents (Mandelbrot, 1977). Elastic rescaling provides a path to interpretation, as it did with fractal geometry. Beginning with integer exponents, a velocity (distance · time^{-1}) can be stretched into an acceleration (distance · time^{-2}). As with spatial units, the degree of stretching can be less than that expressed by an integral exponent. It can be partial. But how can a rate have a time exponent other than -1? This can happen if the rate tends to be more explosive at short than at longer time intervals. If the case rate of measles changes slowly at the time scale of decades, more rapidly at the time scale of two years, and still more rapidly on a seasonal and daily basis, we could use an elastic rescaling factor to express this idea. Elastic rescaling with this factor results in a new quantity with units of (% time^{-1})$^{\alpha}$. The exponent α is unity if the measles case rate neither accelerates nor decelerates at shorter time periods. If case rate changes more rapidly at short time periods, the exponent exceeds unity and time$^{-\alpha}$ becomes more negative, falling somewhere between a rate (time^{-1}) and an acceleration (time^{-2}). If a quantity such as population size changes less rapidly at short time periods than at longer periods, the exponent goes to less than unity; time$^{-\alpha}$ becomes less negative than a simple rate (time^{-1}). Fractal time units apply to the rate of change in population sizes of birds (Sugihara and May, 1990) and to case rates of measles (Sugihara, Grenfell, and May, 1990).

An increasing number of ecological studies use iterative measurement to obtain exponents that are interpretable as an elastic rescaling of lengths and areas. Elastic factors k (see Box 5.7) and exponents (see Box 5.10) have been estimated for structure of coral reefs (Bradbury, Reichelt, and Green, 1984), sponges (Kaandorp 1991), macroalgae (Gee and Warwick, 1994), eelgrass (Turner et al., 1999), plants (Morse et al., 1985; Davenport et al., 1996; Escos et al., 1997; Alados et al., 1998), landscapes (O'Neill et al., 1983; Krummel et al., 1987; Milne, 1992; Meltzer et al., 1992; Baudry, 1993; Vedyushkin, 1994; Otto, 1996; Bell, 1997; Garcia and Jorge, 1997; Miller et al., 1997; Cantero et al., 1998; Etzenhouser et al., 1998; Meyer et al., 1998; Nikora et al., 1999; Olff and Ritchie, 1999; Despland, 2000), the coastline habitat of eagles (Pennycuick and

Kline, 1986), intertidal habitat of gastropods (Beck, 1998), and the ice habitat of polar bears (Ferguson et al., 1998). Elastic factors and exponents have been estimated for animal trails (Dickie and Burrough, 1988; Johnson et al., 1992; With, 1994; Wiens et al., 1995; Erlandsson and Kostylev, 1995; Claussen et al., 1997; Bascompte and Vila, 1997; Anderson et al., 1997; Etzenhouser et al., 1998; Ferguson et al., 1998; With et al., 1999; Westcott and Graham, 2000). Many other studies report only the exponent, which provides less information than the elastic scaling factor k (see Box 5.10 and Equation 2.5c).

The rules for elastic rescaling in Table 5.4 assume that the elastic scaling factor is constant over the range that it is calculated (see derivation in Boxes 5.7 and 5.10). Fractal rescaling can be applied only between an inner and outer scale over which an elastic factor holds (Frontier, 1987). It is now clear that elastic scaling factors such as k apply over narrow scopes of measurement, that is, over only a few doublings in the frequency of measurement (Avnir et al., 1998). Multifractals (e.g., Pascual et al., 1995) allow the exponent and hence the elastic factor to change as a function of measurement resolution or measurement frequency.

Elastic scaling factors are obtained empirically as the average degree of stretching over an iterative series of measurements. Frontier (1987), Sugihara and May (1990), Williamson and Lawton (1991), Korvin (1992), and Hastings and Sugihara (1993) describe estimation techniques.

One of the exciting challenges in ecology is obtaining elastic rescaling factors from theory. For example, it should be possible to work out an elastic scaling for the path length of a predator foraging in a patchy environment. The long-known tendency of a predator to turn if successful or not turn if not successful (e.g., Baker, 1974) should permit calculation of an elastic scaling factor for a quantity, path length, that is related to the cost of foraging. Similarly, it should be possible to work out the elastic scaling factor for predator speed, which decreases as the spatial measurement frequency decreases, because of the tendency of an animal to turn rather than move in straight lines. Other quantities for which theoretical factors seem possible include the flux of fixed energy (as carbon) laterally or vertically through ecosystems, the vertical or horizontal flux of nutrients, the expected value of the recombination rate of two alleles in a population, and the expected value of the encounter rate between prey and predator.

Natural objects often have a self-similar or fractal quality of convolutions within convolutions. This stands in contrast to fabricated objects with which we surround ourselves—boxes, plates, and tabletops. To paraphrase a nursery rhyme:

There was a fractal man and he walked a fractal mile,
He found a fractal sixpence against a fractal stile;
He bought a fractal cat, which caught a fractal mouse,
And they all lived together in a nonfractal house.

The stile in this verse is said to be the border between England and Scotland, a fractal rather than completely straight line. A sixpence coin is not fractal, but the "crooked sixpence" (Charles I of England, 1600–1649) tried to cross the "crooked stile" with an army that would have had a fractal perimeter—so the sixpence was in this sense fractal.

Defined Concepts and Terms for Review and Future Reference

____ dimension
____ duos
____ elastic rescaling
____ elastic scaling factor
____ fractal
____ fractal exponent
____ iterative counting relation
____ logical rescaling
____ normal score
____ normalizing
____ rigid rescaling
____ rigid scaling factors
____ simple rescaling
____ scope
____ trios

6 Dimensions

Using dimensional analysis, researchers have been able to obtain remarkably deep results that have sometimes changed entire branches of science. ... Many people therefore attacked what one would think were almost identical problems using the same simple dimensional analysis approach. Alas, they almost always failed. Dimensional analysis was cursed and reproached for being untrustworthy and unfounded, even mystical ...

It was like the old Deanna Durbin film: a girl with a small suitcase arrives in New York and, in no time, charms the son of a millionaire. Films like this are pleasant to watch. However, if they are treated as a guide to what provincial girls should do, disillusionment is inevitable.

—G. I. Barenblatt, *Scaling, Self-Similarity, and Intermediate Asymptotics*, 1996

6.1 Synopsis

Nearly 200 years ago Joseph Fourier invented a systematic method of reasoning about dissimilar and similar quantities, based on Newton's principle of similitude. He grouped similar units together under the name of a dimension to work out how heat flows through objects that differ in size and shape.

Reasoning with dimensional groupings can be as effective in ecology as it has been in working with mechanical and thermodynamics problems. The dimension of biological entities, rather than traditional dimensions inherited from physics and chemistry, is required for ecological problems that require population reasoning.

The rules for working with dimensions are based on the definitions for a measurement scale and for similarity. The uses of dimensional groups can be loosely grouped under two headings: simplification and model development. By simplifying many variables to a few groups, dimensions allow equations to be checked quickly. Dimensional groups facilitate physical and biological interpretation of parameters obtained from measured quantities. They reduce problems with a large numbers of variables to the minimum number of independent ratios; this reduction is a proven guide to the design of observational and manipulative experiments.

6.2 Physical and Chemical Dimensions

Ratio scale units that are similar (refer back to Box 4.2 and Table 4.4) are grouped together into a *dimension*. Thus quantities measured in centimeters have the same

doi: 10.1016/B978-0-12-627865-1.00006-4

dimension as quantities measured in km, cubits, spearlengths, or nautical miles. Some of these groupings will be related to one another by a change in exponent. For example, the group (centimeters, meters, yards) is related to the group (centimeters^2, hectares, acres) by an increase in the exponent from 1 to 2. Groups of units related by a change in exponent are said to belong to the same dimension. Thus the group (centimeters, yards) and the group (centimeters^2, acres) are assigned to a single dimension, called "length", symbolized by L or [L] in the original notation developed by Rayleigh. The word "dimension" has two somewhat different meanings. It is a name for a *similarity group*. In this sense, length with dimension [L], area with dimension [L^2], and volume with dimension [L^3] belong to the same group. Time with dimension [T] and frequency with dimension [$T]^{-1}$ belong to another group. In its more common meaning, *dimension* refers to geometrical arrangement, sometimes augmented by time. In this sense, distance east is one dimension, distance north is a second dimension, distance up is a third dimension, and [T] is sometimes taken as a fourth dimension.

The similarity groups most often encountered are the dimensions of length [L], mass [M], and time [T], one for each of the first three base units in the SI system (Table 4.1).

Length [L]	Mass [M]	Time [T]
cm, km	gram, kg	second, hour
inch, foot, yard, fathom	grain, scruple, dram	day, week
furlong	ounce, pound	month
rod, mile	quintal, draft	year, millennia

The units of length belonging to a single dimension [L] include standard units (cm, km, etc.) and a large variety of nonstandard measures, from the inch and its multiples to rods and miles. Units of time, from seconds to millennia, all belong to a single dimension, time, symbolized by [T]. Strictly speaking, the units of mass are the gram and its multiples, for which the dimension is [M]. To illustrate the idea of dimensions, several old-fashioned units equivalent to mass (a draft is 2 quintals, or 224 pounds) have been included. Strictly

Table 6.1 Composite Dimensions in the Mechanical System of Dimensions of [M], [L], and [T]

Area		
A = hectare = $(100\,\text{m})^2$	L·L	$= L^2$
$A = 15\,\text{cm}^2$	L·L	$= L^2$
Volume		
$V = 1\,\text{cm}^3 = 1\,\text{cc}$	L·L·L	$= L^3$
Velocity		
$\dot{x} = 15\,\text{cm/sec}$	L/T	$= L^1\,T^{-1}$
Respiration		
$\dot{V} = 15\,\text{cc}\,O_2\,/\,\text{sec}$	L^3/T	$= L^3\,T^{-1}$
Kinetic energy		
$E = 15\,\text{kg}\,(2\,\text{cm/sec})^2 = 60\,\text{kg cm}^2\text{sec}^{-2}$	M·(L/T)^2	$= M\,L^2\,T^{-2}$

Note: For clarity, dimensions are shown without brackets.

speaking, pounds and ounces are not units of mass, but because a constant can be used to convert them to kg, these nonstandard units can be treated as if they belong to the dimension [M].

Measured quantities often have units with *composite dimensions*. These are multiples of dimensions, as shown in Table 6.1. Legendre and Legendre (1998) provide a list of over 65 quantities that have composite dimensions.

Mass, length, and time are not the only dimensions for problems in mechanics. We could choose time T, area A, and mechanical energy E as our base dimensions. Within this system units of mass become a composite dimension of T^2E/A, whereas volume becomes a composite dimension of $A \cdot A^{1/2}$. This or any other dimensional system is valid as long as the units grouped into one dimension do not also belong to another dimension. The dimensional groupings that are used depend on the way quantities are defined and on which units are taken to be similar.

For problems outside mechanics, more dimensions are required. In thermodynamics and bioenergetics, the additional dimension is temperature, for which the symbol is [θ]. The standard unit is 1 degree on the Kelvin scale.

Temperature	[θ]
°Kelvin	K

Temperature is a measure of heat content, and strictly speaking, we already have a composite unit for this in the mechanical system.

$$E/M = M^1\ L^2\ T^{-2}\ M^{-1} = L^2\ T^{-2} = \text{velocity}^2 \tag{6.1}$$

We can interpret the temperature of an object as the square of the velocity of its particles (atoms, molecules) and dispense with temperature as a separate dimension. However, this is awkward and inconvenient, so we treat temperature as a separate dimension rather than as a squared velocity.

For electromagnetic quantities the standard unit is an ampere, a measure of current for which the dimensional symbol is [I].

Current	[I]
Ampere	A

or

Charge	[Q]
Coulomb	C

The ampere is a composite unit, 1 coulomb per second, where a coulomb is a mole of electrons. Consequently:

$$\text{ampere} \equiv \text{coulomb/second} = Q^1\ T^{-1} \tag{6.2}$$

For electromagnetic quantities, the dimension of charge is sometimes used instead of the dimension of current, even though current is the standard unit.

An additional dimension is added for chemistry and biochemistry. That dimension is the amount of substance [N], for which the SI unit is the mole, which is equal to $6.022 \cdot 10^{23}$ particles. Examples of chemical particles are atoms, ions, and molecules.

Chemical entities	[N]
mole	mol

Table 6.2 Dimensions for Standard Units in the SI System

Quantity	Unit	Abbreviation	Dimension
Length	Meter	m	[L]
Mass	Kilogram	kg	[M]
Time	Second	s	[T]
Thermodynamic temperature	Kelvin	K	[θ]
Amount of substance	Mole	mol	[N]
Luminous intensity	Candela	cd	[J]
Electrical current	Ampere	A	[I]

Yet another dimension is added for light. The dimension is luminous intensity [J], for which the base unit is the candela.

Table 6.2 shows the dimensions for each of the seven standard SI units of measurement. As we will see in the next section, standard SI units are inadequate for the biological sciences.

ANOTHER LOOK AT SECTION 6.2

A *gradient* is the change in some quantity with change in location. A *mass gradient* has dimensions of $M\ L^{-1}$. Work out the dimensions for a gradient in each of the quantities in Table 6.1.

6.3 Biological Dimensions

A dimension is needed for biological entities (Stahl, 1962). The additional dimension is the population [N], for which the base unit is an entity # rather than a mole. A *biological entity* is defined as a recognizable object belonging to a population of such objects. Examples of biological entities are individuals, species, cells, or nerve impulses.

Strictly speaking, moles and entities belong to the same dimension, for which the conventional symbol is [N]. However, there are good reasons for distinguishing the chemical dimension [N] for amount of substance measured in moles from the biological dimension [N] of a population composed of entities. The mole is far too large a unit for population biology. A mole or a small multiple might by useful for bacteria or pine needles in an ecosystem. But in general, an appropriately small unit (entities) will be more practical than an enormous unit borrowed from chemistry. Population interactions typically occur between a small number of neighboring individuals. The number of potential prey of a predator is small, for example, compared to the huge number of ions available for interaction with any given ion of opposite charge.

Biological entities	[N]
mol	
dozen, gross	
score	
count, kilocount, megacount	#

Table 6.3 Quantities Based on the Dimension Entities, Represented by #

Quantity	Dimensions	Examples
Loss or gain rate	$N\,T^{-1}$	Contact, mitotic, birth, or death rates
Entities per mass	$N\,M^{-1}$	Cells per gram
Entities per length	$N\,L^{-1}$	Animals per transect; Genes per micron of chromosome
Density	$N\,L^{-2}$	Organism density; Cell density
Concentration	$N\,L^{-3}$	Species or individuals per volume of water
Entity flux	$N\,L^{-2}\,T^{-1}$	Vertical flux of propagules
Entity movement	$N\,L^{-1}\,T^{-1}$	Migration out of a reserve
Energy efficiency	$N\,E^{-1}$	Mitoses per Joule; offspring per Joule of food
Occupancy	$N\,T\,L^{-2}$	Ant-hours foraging per m^2; Residence by migrants

Note: Modified from Stahl (1962). Dimensions are M = Mass, L = Length, T = Time, and E = Energy.

If chemical and biological entities need to be distinguished in the same analysis, the count symbol [#] can be used for the dimension of biological entities. Thus, a quantity such as migration across a boundary would have dimensions $[N]\,[L]^{-1}\,[T]^{-1}$ in conventional notation, or equivalently $[\#]\,[L]^{-1}\,[T]^{-1}$. In either case the dimensional notation would be read as entities per unit length per unit time.

There are no standard units for the dimension of biological entities. Possible units are a mole (e.g., 2 mol bacteria $= 2 \cdot 6.02 \cdot 10^{23}$ bacteria), a dozen (e.g., 2 dozen genes $= 2 \cdot 12$ genes), a score (e.g., 2 score cells $= 2 \cdot 20$ cells), or a count (e.g., 2 kilocount ants $= 2 \cdot 10^3$ ants). A mole is inconveniently large, since we are often interested in the interaction of a small numbers of entities. A dozen and a score are of useful magnitude, but are not base 10. A count, for which the symbol is #, is of the right magnitude and can be modified by standard prefixes: kilocounts, megacounts, etc. Standard prefixes are also readily added to named entities. For example, a protein might measure 120 kilobases long.

Units such as a kilocount of cells, a kilocount of predator attacks, or a megacount of potential encounters are not standard, but they are useful in biology and can be handled in a rigorous fashion (Stahl, 1962). The philosophical objection to using counts of objects or events as a measurement scale (Ellis, 1966) can be met by insisting that this scale does not consist of numbers; it has units of entities (animals, genes, etc.) on a ratio scale. This reasoning follows Kyburg (1984), who argues that all measurements must have units. Physical and chemical dimensions will not suffice, because they miss the biology. An example is the number of organisms per unit area, a quantity that often is of interest in ecology. If we try to force this quantity into the conventional dimensions of chemistry and physics, we lose the count and hence the biology. Dimensions are most effectively used in reasoning about ecological problems if the choice of dimensional groupings is based on biological similarities, not on standard groupings carried over from physics and chemistry.

The utility of this similarity group in biology is evident from Stahl's (1962) listing of 22 quantities based on this dimension. Stahl's list, with additions and modifications, is here grouped into quantities based on entities (Table 6.3), per capita quantities (Table 6.4), and interaction rates (Table 6.5).

Care is needed in defining which quantities can be grouped together into the dimension of entities. We could group otters and sea urchins together into a single

Table 6.4 Per Capita Quantities Based on the Dimension Entities, Represented by #

Quantity	Dimensions	Examples
Length	$L\,N^{-1}$	Spacing of plants
Area	$L^2\,N^{-1}$	Avian territory; crown area of a tree
Volume	$L^3\,N^{-1}$	Volume filtered per organism
Time	$T\,N^{-1}$	Time per mitosis, per decision
Mass	$M\,N^{-1}$	Mass of cell, organism
Energy	$E\,N^{-1}$	Caloric content

Note: Modified from Stahl (1962). Dimensional symbols as in Table 6.3.

Table 6.5 Interaction of Entities, Represented by # · # = #^{2}

Quantity	Dimensions	Examples
Potential interaction	N^2	Cell, molecular, genetic contacts; duos
Diversity	$N^2\,L^{-2}$	Diversity
Interaction frequency	$N^2\,T^{-1}$	Contact rate; "biological temperature"
Interaction time	$T\,N^{-2}$	Synaptic delay time; search time by predator
Energy exchange	$E\,N^{-2}$	Joules per capture
Energy exchange rate	$E\,N^{-2}\,T^{-1}$	Change in energy, as in learning
Ratio of entities	$N\,N^{-1}$	Active to inactive genes; selection coefficient
Complex interactions	N^3	Social, colonial activity
Interaction ratio	$N^2\,N^{-3}$	Social, colonial activity

Note: Modified from Stahl (1962). Dimensional symbols as in Table 6.3.

Table 6.6 Entities at Different Levels of Biological Organization

Biochemical entities	Ions, atoms, molecules (including proteins)
Genetic entities	Chromosomes, genes, alleles, mutations
Cellular entities	Nuclei, mitochondria, cells
Behavioral entities	Attempts, successes, modal action patterns (MAPs)
Population entities	Interacting species
Ecosystem entities	Number of taxa (species, order, etc.); number of trophic levels.

dimension of animal counts. Otters encounter urchins, and hence there is a biological basis for placing them in the same similarity group. We would not group species counts along with nerve impulse counts because species and nerve impulses are at different levels of biological organization. Thus, a good rule is to assign entities to the same similarity group if they are at the same level of biological organization.

Table 6.6 distinguishes biological entities at different levels of biological organization. At the biochemical level, counts are large enough that a mole is the appropriate unit of measurement. For genetic, cellular, and population entities, units of megacounts or kilocounts are more appropriate. For behavioral and ecosystem entities we would tend to use a count rather than one of its multiples.

Quantitative reasoning based on dimensions has traditionally not been important in population and community ecology. This is due in part to the use of the conventional mechanical dimensions of mass, length, and time. These leave out much of the biology of

populations, in which the dynamics depend on rates of contact between organisms. The dimensional scheme shown in Tables 6.3–6.5 includes the biology of contact-dependent processes by defining units that measure contact rates. With this scheme, quantitative reasoning based on dimensions can be applied to population and community ecology. For example, much of the work on evolution has treated gene flow as a number lacking units and dimensions. This impedes the analysis of gene flow as a function of spatial and temporal scale. Once gene flow is treated as a scaled quantity, the role of space and time scales in evolutionary biology can be analyzed quantitatively. Another example is species diversity, which is traditionally treated as a dimensionless quantity when in fact it is a derived quantity based on entities and areas. Once diversity is treated as a quantity with dimensions, the analysis of diversity as a function of spatial scale (e.g., MacArthur, 1969) can be carried out with the aid of this style of quantitative reasoning (Schneider, 2001a).

ANOTHER LOOK AT SECTION 6.3

Of the composite dimensions in Tables 6.3 and 6.4, how many have you encountered?

6.4 Measurement and Dimension

The seven base units in the SI system are routinely used to define dimensional groupings (Table 6.2). But these are not the only valid groupings. Because the central idea is that quantities are grouped according to some notion of similarity, we can use any similarity group that we like. If we were interested in economics, we could define a dimension called *cash* with units of pennies, nickels, dimes, dollars, and megabucks. This dimension includes any of the world currencies: pesos, yen, francs, and so on. All these units are similar in that they can be used to purchase goods and services.

Cash	[$]
dollar, dime, nickel	
peso	
franc, centime	
yen	

But though we are free to use dimensions other than those of chemistry and physics, we are not free of the mathematical logic that makes dimensions useful. Gold (1977) noted that the biologist has the burden of understanding, in greater detail than a physicist or chemist, the principles of measurement and dimension so that, when necessary, appropriate dimensions are coined. The principles are simple (Box 6.1). First, the action of measurement associates a set of numbers with a set of objects, defining a *measurement scale*. Next, two measurement scales belong to the same dimension if the measurement of two objects produces the same ratio on both scales. This establishes an *operational definition for a dimension*, consistent with the rules in Box 4.2 and Table 4.4. According to this operational definition, measurements in kg and pounds belong to the same dimension because there is a 1:1 scaling (here on the surface of the earth). Box 6.1 presents in more detail the definitions of measurement scale and dimension that underpin dimensional methods.

Box 6.1 Definition of Measurement Scale and Dimensions

Definition 1. A *measurement scale* is a rule of association between a set of physical objects and the set of real numbers. The association entails a physical operation, that of measurement. The rule specifies a unique value in the range of real numbers for each element of the domain set of physical objects (from Gold, 1977).

Definition 2. Any two measurement scales, say, *S'* and *S*, belong to the same dimension if they have the same domain set and if:

$$\frac{S'_{ObjectA}}{S_{ObjectA}} = \frac{S'_{ObjectB}}{S_{ObjectB}}$$

ObjectA and ObjectB are any pair of objects in the domain set (see Gold, 1977).

With some rearrangement, Definition 2 results in a noniterative scaling relation (Equation 2.5a), but with only two objects.

$$\frac{S_{ObjectA}}{S_{ObjectB}} = \left(\frac{S'_{ObjectA}}{S'_{ObjectB}}\right)^1$$

This is the scaling relation for completely similar objects having a length dimension of $D = 1$.

The corresponding scaling function (see Equation 2.5c) is:

$$S_A = k \cdot S' \cdot A$$

where:

$$k = S_B/S'_B$$

When there are multiple variables, reduction of variables to one or more dimensionless Π ratios will prove to be a useful device.

$$\Pi = \left(\frac{S_A}{S_B}\right)\left(\frac{S'_A}{S'_B}\right)^{-1}$$

Rule 1 in Box 6.1 establishes an operational definition for any quantity. To make a measurement, we need to define our objects and then assign a number to each object. The objects can be natural (lakes, islands, organisms) or artificial (clock ticks, ruler marks, quadrats). The numbers we assign can be on any of four types of scale: nominal, ordinal, interval, or ratio (see Section 3.5). Measurements on a ratio type of scale are special in that they allow us to calibrate one scale to another. The simplest form of calibration is a one-to-one scaling, which places two scales in the same dimension (Definition 2, Box 6.1). Two measurement scales (both must be ratio scales) belong to the same dimension if they produce the same ratio of measurements across objects. Measurement scales belonging to the same dimension can be substituted for each other directly. We can make measurements with one scale (e.g., inches) and then compute the measurement on any other scale belonging to the same dimension (e.g., centimeters or meters).

The definition of dimension in Box 6.1 implies comparison of two objects, as when we compare the ratio of lengths of a large to a small pendulum as measured in spearlengths to the ratio of lengths as measured in meters. This is a noniterative scaling relation (Equation 2.5a). The concept of dimension can also be derived from an *iterative counting relation*. An example is quadrat size, for which the iterative counting relation is:

$$\frac{n}{n_o} = \left(\frac{L}{L_o}\right)^{-2} \tag{6.3}$$

In this expression, L_o is the length along the side of a small tile (say, of unit size 10 cm), L is the length along the side of a larger tile (of, say, 1 m). n is the number of large tiles to fill the quadrat, whereas n_o is the number of small tiles to fill the same quadrat. The dimension is defined by the exponent, which in this case is 2. This exponent connects the operation (L/L_o) on the right side of the equation to the numbers (n/n_o) that result from measurement. The area of a quadrat is the product of tile size and the number of tiles:

$$\frac{A}{A_o} = \frac{n \cdot L^2}{n_o \cdot L_o{}^2} \tag{6.4}$$

Note that this relation applies to iterative measurement of a single object. If we multiply both sides of Equation 6.3 by $(L/L_o)^2$, we obtain:

$$\frac{n \cdot L^2}{n_o \cdot L_o{}^2} = \left(\frac{L}{L_o}\right)^{-2}\left(\frac{L}{L_o}\right)^{2} \tag{6.5a}$$

and then rearranging, we have:

$$\frac{A}{A_o} = \left(\frac{L}{L_o}\right)^{0} \tag{6.5b}$$

The exponent is zero, which means that the area of the object (the quadrat) is independent of the units we use to measure the object. We take this for granted, but as we will see, this independence applies only to Euclidean objects (lines, regular polygons, regular solids) or to objects that we treat as Euclidean, ignoring any complexity of shape or structure.

Having defined a dimension via an iterative counting relation, we move next to Definition 2 (Box 6.1), which compares objects. For 2 quadrats, one small and the other big, the iterative counting relations are as follows:

$$\frac{nsmall}{nsmall_o} = \left(\frac{L}{L_o}\right)^{-2} \tag{6.6a}$$

$$\frac{nbig}{nbig_o} = \left(\frac{L}{L_o}\right)^{-2} \tag{6.6b}$$

As a consequence, the ratio of counts will be equal:

$$\frac{nsmall}{nsmall_o} = \left(\frac{L}{L_o}\right)^{-2} = \frac{nbig}{nbig_o} \tag{6.7a}$$

This is a reformulation of Definition 2 in Box 6.1. The reader can verify this relation by drawing two quadrats: a small one with 4 units on a side, and a bigger one with 6 units on a side. Draw in the tiles $nsmall_o = 16$ tiles and $nbig_o = 36$ tiles. Then group the tiles into squares consisting of 4 tiles. At this new scale the two quadrats will measure, respectively, 2 units and 3 units on a side. The number of larger tiles in the two quadrats will be $nsmall = 4$ tiles and $nbig = 9$ tiles. Equation 6.7b expresses this particular rescaling, shown in abstract form in Equation 6.7a:

$$\frac{4}{16} = \left(\frac{2}{1}\right)^{-2} = \frac{9}{36} \tag{6.7b}$$

To complete our examination of comparative scaling, we multiply Equation 6.7a by the scope of measurement $(L/L_o)^2$:

$$\frac{nsmall}{nsmall_o}\left(\frac{L}{L_o}\right)^2 = \left(\frac{L}{L_o}\right)^{-2}\left(\frac{L}{L_o}\right)^2 = \frac{nbig}{nbig_o}\left(\frac{L}{L_o}\right)^2 \tag{6.7c}$$

Area is the product of tile size and tile count (Equation 6.5a), and so Equation 6.7c is reexpressed as follows:

$$\frac{Asmall}{Asmall_o} = \left(\frac{L}{L_o}\right)^0 = \frac{Abig}{Abig_o} \tag{6.7d}$$

Next, we move to comparative scaling of several quadrats. We begin by rearranging Equation 6.7d. Multiplying both sides by the same factor $(Asmall_o/Abig)$, then applying an exponent of -1, we have:

$$\frac{Abig}{Asmall} = \frac{Abig_o}{Asmall_o} \tag{6.8}$$

This says that the ratio of areas of a big and a small quadrat will be the same, regardless of whether we use large tiles L^2 or small tiles L_o^2. Consequently, we can scale any quadrat to the smallest quadrat:

$$\frac{A}{Asmall} = \frac{A_o}{Asmall_o} \tag{6.9}$$

Even more generally, we can rescale any quadrat to a reference quadrat A_{ref}:

$$\frac{A}{A_{ref}} = \frac{A_o}{A_{ref_o}} \tag{6.10}$$

This allows us to establish measurement relations (Equation 2.8a) in which we scale a quantity to units defined by a measurement protocol. For example, we can scale species number to quadrat size, as in Box 2.3:

$$\frac{Nsp(A)}{Nsp(A_{ref})} = \left(\frac{A}{A_{ref}}\right)^{\beta} \tag{6.11}$$

This scaling relation is based on a protocol whereby a quadrat with area A has the same shape as the reference quadrat A_{ref}. Another characteristic to note is that Equation 6.11 is based on comparing objects. This notation captures this distinction. The ratio A/A_{ref} identifies that noniterative scaling is across objects, each measured with respect to a single unit of measurement. In contrast, the ratio A/A_o identifies an iterative scaling (two or more units of measurement applied to a single object).

In Equation 6.11 we define the units on the right side as a series of quadrats that have the same shape, leaving the results of the measurement operation to appear on the left side. Thus, this is a noniterative measurement relation. Not all power laws are based on measurement relations. Many are based on the scaling of one quantity to another, where both quantities are defined relative to some measurement operation. An example of a scaling relation that is not a measurement relation is species number in relation to island size:

$$\frac{Nsp(A)}{Nsp(A_{ref\ Island})} = \left(\frac{A}{A_{ref\ Island}}\right)^{\beta} \tag{6.12}$$

In this example species number is defined relative to a noniterative counting relation, whereas area is defined relative to a different measurement protocol (also noniterative). Under what conditions can we use this scaling relation? Returning to the iterative definition of dimension (Equation 6.3) that underlies the comparative definition in Box 6.1, we have an expression for the number of square tiles (with side L and side L_o) that are required to cover an island:

$$\frac{n}{n_o} = \left(\frac{L}{L_o}\right)^{-D} \tag{6.13}$$

The exponent D will be close to a value of 2 for islands with simple shapes, such as the nearly round islands created by a single volcano. The exponent will rise above 2 as the island shape becomes increasingly complex due to a highly indented coastline. The exponent will be greater for an archipelago than for a single island. The area of an island at two different tile sizes (either L or L_o on a side) will then be:

$$\frac{A}{A_o} = \frac{n}{n_o}\left(\frac{L}{L_o}\right) = \left(\frac{L}{L_o}\right)^{-D+1} \tag{6.14}$$

The area of the reference island at two different tile sizes will be:

$$\frac{A_{ref}}{A_{ref_o}} = \frac{n}{n_o}\left(\frac{L}{L_o}\right) = \left(\frac{L}{L_o}\right)^{-Dref+1} \tag{6.15}$$

The area of any island relative to the reference island will be:

$$\frac{A}{A_{ref}} = \frac{A_o}{A_{ref_o}} \left(\frac{L}{L_o}\right)^{-D+Dref} \tag{6.16}$$

Comparing this equation with Equation 6.10, we now see a correction factor that depends on the difference in complexity between an island and the reference island:

$$\frac{A}{A_{ref}} = \frac{A_o}{A_{ref_o}}(Correction) \tag{6.17a}$$

$$Correction = \left(\frac{L}{L_o}\right)^{-D+Dref} \tag{6.17b}$$

Thus, the scaling of species number to island size is:

$$\frac{Nsp(A)}{Nsp(A_{ref})} = \left(\frac{A}{A_{ref}}\right)^{\beta} = \left(\frac{A_o}{A_{ref_o}} \left(\frac{L}{L_o}\right)^{-D+Dref}\right)^{\beta} \tag{6.18}$$

If we use quadrats of fixed shape, then $D = D_{ref}$ and there is no correction because $(L/L_o)^0 = 1$. If we use Euclidean shapes to approximate island area, the difference is zero and there is no correction. However, if we suspect that species number depends on island shape and we know our islands differ in complexity of shape, we can introduce a correction, as in Equation 6.18.

This analysis demonstrates the difference between a scaling relation based on two quantities and a scaling relation based on a quantity and a measurement protocol (refer back to Equation 2.5a versus Equation 2.6a). If we ignore the distinction, we are open to hidden uncertainty and to values of scaling exponents that are less certain than they appear. This casts doubt on the literature in which scaling arguments are developed by combining two or more scaling functions. The literature on species-area relations is plagued by this confusion. This analysis also demonstrates the difference between iterative and noniterative scaling. If we ignore the difference, we are open to making an erroneous conclusion based on comparing exponents for which there is no basis of comparison. The literature on species area curves is handicapped by unfounded comparison of scaling exponents. The distinction between noniterative scaling among systems (Equations 2.1, 2.2, 2.5, 2.6, 6.11) and iterative scaling within a system (Equations 2.3, 2.4, 2.7, 2.8, 6.3) will prove useful in making sense of the literature on scaling. Iterative scaling is the basis for the measurement of structural complexity and for quantifying complex phenomena (Boxes 2.1 and 5.10). Iterative scaling is relatively new, appearing in the last two decades of the 20th century. Noniterative scaling relations have a much longer history in biology, going back several centuries to body size allometry (scaling across organisms) and back more than a century to species area relations (scaling across quadrats or islands).

ANOTHER LOOK AT SECTION 6.4

Find a power law of interest to you. Re-express it as an equivalence between two scopes, with the appropriate exponent. Is the relation iterative or noniterative? Is it a scaling relation (that is, a noniterative observationally based scaling relation) or a measurement relation?

6.5 Dimensional Analysis in Biology

Dimensional analysis applies to both iteratively and noniteratively measured variables (Barenblatt, 1996). However, most presentations are based on measurement relations that are noniterative, as in Equation 2.6a. This classical form of dimensional analysis employs the comparative definition of dimension in Box 6.1. Books on the topic of dimensional analysis (with the exception of Barenblatt, 1996) present Euclidean spatial dimensions and a Newtonian time dimension $[T]^1$, even though fractal exponents are possible based on the definition of a measurement relation and dimension (Boxes 5.10 and 6.1).

Dimensional analysis in this classical form is limited compared to what can be accomplished from iteratively defined measurement relations (e.g., Equation 6.4). Nevertheless, classical dimensional analysis remains useful in simplifying a problem and guiding quantitative reasoning.

To illustrate the advantage of simplification, we begin with an informal dimensional analysis of species diversity in relation to habitat diversity on islands that differ in their distance from the mainland. With just four variables (island area, island distance, species diversity, habitat diversity), there are six pairwise graphs. In one of these graphs (species diversity versus habitat diversity), we expect a strong relation due to island area. Instead of plotting all six graphs or undertaking a multiple regression analysis, we can simplify the problem with dimensional analysis of species and habitat diversity in relation to island area. We begin by listing the four variables in terms of measurement operations (Box 6.2). In all four, the measurement operation is across units. The operations consist of defining the diameter of an island, then measuring the radius as a straight line distance (r/r_{ref}); defining habitat categories, then counting them ($\#/\#_{ref}$); defining species and counting them ($\#/\#_{ref}$); defining what is meant by distance to mainland and then measuring the straight line distance (L/L_{ref}). It is worth noting that these operations will not be automatically defined in the same way by all investigators. For example, island diameter might be defined as the maximum diameter, the minimum diameter, the average of the maximum and minimum, or in some other way. Island radius and distance from mainland belong to the same dimension (length), so we can form a dimensionless ratio, $IslandA/Distance^2$. The two remaining variables can also be taken as a dimensionless ratio, *Nsp/Nhab*, based on the shared dimension. The first ratio measures room for dispersion within an island relative to intensity of dispersion from the mainland. That is, we expect dispersion from the mainland to become important when the square of the distance from the mainland is less than the area of the island. The second ratio measures species diversity relative to habitat diversity. Plotting these two ratios against each other (a single graph) will summarize all the information we have in the most economical way possible.

Box 6.2 Noniterative Scaling Relations for Analysis of Species Diversity in Relation to Habitat Diversity on Islands That Differ in Their Distance from the Mainland

Island area:

$$\frac{IslandA}{IslandA_{ref}} = \left(\frac{\pi \cdot r^2}{\pi \cdot (r_{ref})^2}\right) = \left(\frac{r}{r_{ref}}\right)^2$$

Species number:

$$\frac{Nsp}{Nsp_{ref}} = \left(\frac{\#}{\#_{ref}}\right)$$

Number of habitats:

$$\frac{Nhab}{Nhab_{ref}} = \left(\frac{\#}{\#_{ref}}\right)$$

Separation from mainland:

$$\frac{Distance}{Distance_{ref}} = \left(\frac{L}{L_{ref}}\right)$$

Over the last two centuries, several techniques have been developed to aid dimensional analysis. One is the *dimensional matrix*, invented by Joseph Fourier (1822). The matrix displays quantities relative to dimensions. Setting up a dimensional matrix forces us to think about a problem in terms of scaled quantities. Once the matrix is formed, it can be revised based on biological knowledge. Revision of the matrix leads to new and efficient ways of looking at a problem.

To work with dimensions, we will need a list of algebraic rules. Table 6.7 shows the list developed by Gold (1977). These rules follow from the definitions in Box 6.1; they are consistent with the rules for working with ratio scale units listed in Table 4.4. Riggs (1963) developed a shorter list of rules for checking the dimensional consistency of equations (see Chapter 13).

In addition to algebraic rules, we will need *Buckingham's theorem,* which tells us that the number of independent dimensionless ratios (called *Π ratios*) that can be formed will be equal to the number of variables, minus the number of dimensions. As we will see, Π ratios have many uses, from efficient summarization of system characteristics to the development of scaling functions. There are several prescriptions for obtaining Π ratios, once variables and dimensions are defined. One prescription, called sequential elimination, works by rearranging the dimensional matrix. This prescription is more

easily learned than the other two, Rayleigh's method and Buckingham's method. Case Study 6.5.1 shows all three techniques for a single example.

Dimensional analysis is a way of simplifying a problem based on measurement or scaling relations grouped into dimensions. Table 6.8 shows a generic approach to dimensional

Table 6.7 Rules for Working with Dimensions

1. When quantities are multiplied, the result is the product of the dimensions (and units) of the two factors.
   Example: velocity · time = distance $[L/T]\,[T] = [L]$
2. When one quantity is divided by another, the result is the ratio of the dimensions and units of the two quantities.
   Example: production / biomass = % / time $[M/T] / [M] = T^{-1}$
3. Dimensionless ratios result from division of two quantities with the same dimensions.
   Example: (velocity · time) / distance = R $([L/T]\,[T]) / L = L^0\,T^0 = 1$
4. Exponents and logarithms have no units or dimensions.
   Example: $\log_{10}(10\,\text{km} / \text{km}) = 1$
   Note that a quantity must be divided by its units to meet this condition.
5. Conversion factors are dimensionless ratios with different units.
   Example: 35.274 ounces / 1 kg = 1 $[M] / [M] = [M^0] = 1$
   Conversion factors leave dimensions unchanged.
6. Addition and subtraction apply only to quantities with the same units.
7. Quantities set equal to each other must have the same dimensions.
   Example: $Nsp = k{\cdot}A^z$ $[N] = [N \cdot L^{-z}]\,[L^z]$
   The quantity $k{\cdot}A^z$ has dimensions #.
   As a result, this composite quantity can be set equal to species number.
8. Dimensions are unchanged by magnitude, including the operation of taking a limit.
   Example for N = number of ants, A = area, t = time:
   $(dN/dA)\,(dA/dt) = dN/dt$ $[N\,L^{-2}]\,[L^2\,T^{-1}] = [N\,T^{-1}]$
   $A^{-1}\int N\,dA = \text{mean}(N)$ $[L^{-2}]\,[N]\,[L^2] = N$
9. Probabilities are dimensionless ratios. This follows from the definition of a probability as the limit of an expression of the type:
   (Number of occurrences / Number of possible occurrences)
   For completeness, an additional rule is needed.
10. Pure numbers (e, π) have no dimensions (Riggs 1963).

From Gold (1977).

Table 6.8 Use of the Dimensional Matrix to Obtain Dimensionless Ratios, Scaling (or Measurement) Relations, and Scaling Functions

The dimensional matrix displays the relation of quantities that apply to a problem. The steps in applying it are:

1. List all quantities, with symbols and units.
2. State what is known about the relation of these quantities.
3. Define each quantity in terms of measurement operations (dimensions).
4. List quantities as rows, dimensions as columns.
5. Fill out the resulting dimensional matrix.
6. Construct and interpret dimensionless ratios.
7. As needed, revise the list of quantities or dimensions.
8. Fill out a new matrix, construct new ratios, and interpret.
9. Rewrite dimensionless ratios to scaling (or measurement) relation and scaling functions.

analysis, with emphasis on the dimensional matrix. The generic approach in Table 6.8 is illustrated by a series of case studies that demonstrate the utility of dimensional analysis in ecology.

6.5.1 Case Study 1: Dinosaur Running Speed

To illustrate the dimensional matrix and its role in developing scaling functions, we begin with a classic example: How fast could dinosaurs run? Alexander (1989) used dimensional methods to analyze the problem. Beginning with Step 1 in Table 6.8, the quantities are leg length, which can be measured on fossils, and running speed. Convenient symbols and appropriate units are:

$$v = \text{straight line running speed in units of meter} \cdot \text{sec}^{-1} \tag{6.19a}$$

$$legL = \text{leg length in units of meters} \tag{6.19b}$$

The suspected relation (Step 2), based on experience with running, is that running speed increases as a function of leg length:

$$v = \text{f}\,(legL) \tag{6.20}$$

The dimensions for each of the quantities are:

$$v = \text{m} \cdot \text{sec}^{-1} = [\text{L}]\,[\text{T}]^{-1} \tag{6.21a}$$

$$legL = \text{meter} = [\text{L}] \tag{6.21b}$$

When quantities are listed as rows (Step 3) and dimensional groups as columns (Step 4), the dimensional matrix (Step 5) is:

	[L]	[T]	$n$rows − $n$columns = 0
v	1	−1	
$legL$	1		

According to the Π theorem of Buckingham (1914), the number of independent dimensionless ratios will be the number of variables (rows) minus the number of dimensions (columns). It is evident that we have too few rows (too little information) with our first formulation. We need another quantity. As will soon become apparent, we need a quantity that includes the dimension of time. Because dinosaurs are extinct, we cannot obtain measured variables that have a time component, such as steps per minute. However, we do have a constant that relates distance to time in the Triassic era as well as now, which is acceleration due to gravity:

$$g = 9.8 \text{ meter} \cdot \text{sec}^{-2} = [\text{L}]\,[\text{T}]^{-2} \tag{6.22}$$

The form of the relation becomes:

$$v = \text{f}\,(legL, accel) \tag{6.23}$$

The revised dimensional matrix is:

	[L]	[T]	*n*rows − *n*columns = 1
v	1	−1	
legL	1		
g	1	−2	

The next step is to construct dimensionless ratios. Massey (1986) lists three approaches: Rayleigh's method, Buckingham's method, and sequential elimination within the dimensional matrix. The third method is often the easiest to use. This method most easily combines biological concepts and judgement with the logic of similarity groups. The third method also makes it evident, more than the other two, that dimensionless ratios can be formed in any of several ways. The method produces dimensionless ratios by sequential combination of rows (quantities) so as to reduce the dimensions (columns) to zero. Rows are combined by multiplying quantities (adding exponents), by taking the ratio of quantities (subtracting exponents), and by raising quantities to powers (multiplication of exponents by a constant) as needed to reduce dimensions to zero. This continues until rows are reduced to the number of dimensionless ratios set by Buckingham's theorem. Box 6.3 demonstrates the method. The reduction to a dimensionless ratio is carried out twice to show two different solutions for the same ratio.

Box 6.3 Dimensionless Ratios Formed by Sequential Elimination

The dimensional matrix has three rows and two columns, hence one dimensionless ratio:

	[L]	[T]
v	1	−1
legL	1	
g	1	−2

To reduce the time dimension, square the velocity, then divide by acceleration. The new matrix is:

	[L]	[T]
v^2/g	1	0
legL	1	0

To reduce the length dimension, subtract the second row from the top row.

	[L]	[T]
$v^2/(g \cdot legL)$	0	0

Here is sequential elimination in a different sequence.

Start with 3 by 2 dimensional matrix. Take the product of acceleration and leg length, raised to the ½ power:

	[L]	[T]
v	1	−1
$(legL \cdot g)^{1/2}$	1	−1

Subtract the second row from the top row.

	[L]	[T]
$v/(g \cdot legL)^{1/2}$	0	0

The dimensionless ratio summarizes the measurement and scaling relation among the variables. The ratio, which we will label with the Greek letter *pi* for "product", is:

$$\Pi = v \cdot (g \cdot legL)^{-1/2} \tag{6.24a}$$

Equivalently:

$$\Pi = v^2/(g \cdot legL) \tag{6.24b}$$

The Π ratio in Equation 6.24b is the square of that in Equation 6.24a. Because Π is a dimensionless number, there is no need to distinguish the Π ratio for velocity (Equation 6.24a) from the Π ratio related to mass-specific energy (Equation 6.24b).

Once a dimensionless ratio has been obtained (Step 6, Table 6.8), it is interpreted and revised as needed (Step 7). For Equation 6.24b, kinetic energy of running is proportional to the square of running speed v^2, whereas the potential energy of an animal balanced on its leg is the product of leg length and gravitational acceleration. Thus, the Π ratio is the ratio of kinetic to potential energy, which is held in balance during the coordinated activity of running.

Scaling relations (Equation 2.5a) and scaling functions (see Equation 2.5c) are obtained from the Π ratio by algebraic reorganization (Box 6.1 and Step 9 in Table 6.8). For the dinosaur example, acceleration g is a constant. Π is always a dimensionless constant. Hence the scaling relation is:

$$\left(\frac{v}{v_{ref}}\right) = \left(\frac{legL}{legL_{ref}}\right)^{1/2} \tag{6.25a}$$

From this, we expect that a dinosaur with a leg length 1.5 times that of a horse could run at a speed of $1.5^{1/2} = 1.2$ times that of a horse. If we want to make specific predictions, we can rewrite the scaling relation as a scaling function:

$$V = k \cdot legL^{1/2} \tag{6.25b}$$

where $k = v_{ref} \cdot legL_{ref}{}^{-1/2}$. To verify the scaling function, we can plot the running speed against the square root of leg length of animals of different sizes. Once we verify that the relation is a straight line on this plot, we can estimate k, then compute the expected running speed of a dinosaur of known leg length.

Sequential elimination is not the only way to form dimensionless ratios. Rayleigh's method uses algebra. An equation is written for each dimensional group, then the scaling function is obtained by solving the system of equations. The Π ratio can be written from the scaling function. Box 6.4 shows the procedure.

Box 6.4 Formation of Dimensionless Ratios by Rayleigh's Method

$v = \mathrm{f}(\ legL, g\ )$	The functional expression
$v = k \cdot legL^{\alpha} \cdot g^{\beta}$	Expressed as a series of products
$\mathrm{L}^1 = \mathrm{L}^0 \cdot \mathrm{L}^{1\alpha} \cdot \mathrm{L}^{1\beta}$	Equation for length dimension
$1 = 0 + \alpha + 1\beta$	Same equation for exponents of length
$\mathrm{T}^{-1} = \mathrm{T}^0 \cdot \mathrm{T}^0 \cdot \mathrm{L}^{-2\beta}$	Equation for time dimension
$-1 = 0 + 0 - 2\beta$	Same equation for exponents of time

There are two equations with two unknowns (α and β). The solution is $\alpha = ½$ and $\beta = ½$, hence:

$$v = k \cdot legL^{1/2} \cdot g^{1/2}$$
$$\Pi = v^{-1} \cdot legL^{1/2} \cdot g^{1/2}$$

This method, unlike sequential elimination, leads directly to scaling functions. However, the focus on the algebraic manipulation tends to push aside the use of biological concepts and judgment in carrying through the analysis.

The third approach, Buckingham's method, produces Π ratios based on structured inclusion of knowledge of the system. An important step is to identify "nonrecurring" variables. These are the variables that, based on our knowledge, we want to convert to dimensionless form. Table 6.9 lists the steps for Buckingham's method. Box 6.5 shows the procedure.

Table 6.9 Buckingham's Method, Based on the Π Theorem

1. Form the dimensional matrix, consisting of a row for each variable and a column for each dimension. Fill in the matrix with the exponents for each variable in terms of dimensions.
2. Calculate the number of Π ratios as *nratios = nvariables - ndimensions*.
3. From the list of variables, identify variables considered of interest based on knowledge of the system. These are called *nonrecurring variables*. Choose as many nonrecurring variables as there are dimensionless ratios. The remaining variables are called *recurring variables*.
4. Make sure that as a set the recurring variables include all the dimensions listed as columns. Make sure that no two recurring variables form a dimensionless ratio. If the recurring variables do not meet these conditions, revise the choice of nonrecurring variables or, if necessary, revise the dimensional matrix.
5. Partition the dimensional matrix so that the recurring variables are listed together, usually at the bottom.
6. Use the set of recurring variables to form a nondimensional Π ratio for each nonrecurring variable. This can be done by inspection or by Rayleigh's method.

Box 6.5 Dimensionless Ratios Formed by Buckingham's Method

	[L]	[T]
$legL$	1	
v	1	−1
g	1	−2

1. The dimensional matrix has three rows and two columns.
2. $Nrows - ncolumns = 1$ ratio.
3. Velocity depends on leg length (not *vice versa*), so velocity is chosen as the nonrecurring variable.
4. This will leave leg length and gravitational acceleration as the recurring variables (see Table 6.9). They are acceptable because the group includes all dimensions and no pair in the group forms a dimensionless ratio.
5. The matrix is partitioned into a recurrent group (below the line) and nonrecurrent variables (above the line).

	[L]	[T]
v	1	−1
$legL$	1	
g	1	−2

6. The recurrent group is reorganized so that it will form a dimensionless ratio relative to the nonrecurrent group:

	[L]	[T]
v	1	−1
$(legL \cdot g)^{1/2}$	1	−1

7. Form a Π ratio for the nonrecurrent variable.

	[L]	[T]
$v/(g \cdot legL)^{1/2}$	0	0

The dinosaur example was deliberately simple, to show how the three different methods work, to demonstrate that they all produce equivalent results, and to display the relation of these standard methods to the scaling relations introduced in Chapter 2 (compare Equation 6.25a to Equation 2.5a). The next example is more complex, with several variables and more than one dimensionless ratio. This example demonstrates the utility of dimensional reasoning about a problem before undertaking field research. It also demonstrates a neglected step: reorganization of the dimensional matrix in light of knowledge of the system.

6.5.2 Case Study 2: Otter Monitoring

Moving from body size scaling (as an accessible example) to spatial scaling, we look at whether changes in geographic range can be used to monitor changes in sea otter

populations. Reasoning with dimensions is used to identify assumptions and isolate the quantities most in need of measurement, before committing resources to a field survey.

Referring to Table 6.8, the first step is to list the quantities that apply to the problem. For the problem of sea otter range, we start with a guess about the factors that determine rate of range expansion. Reasonable guesses are birth rates, death rates, and individual spacing. The quantities are:

_____ Population size at time t
_____ Birth rate
_____ Death rate
_____ Individual spacing
_____ Length of coastline occupied

A coordinated set of easily remembered symbols would be nice, so the symbol N will stand for population numbers, $\dot{D}$ will stand for the crude death rate, and $\dot{B}/N$ will stand for per capita birth rate. A dot over a symbol signifies the time rate of change. This notation uses three conventional mnemonic symbols (N, $\dot{B}$, and $\dot{D}$) together with a diacritical mark, to distinguish counts from rates.

To complete the first step, here is the list of quantities with symbols and units:

Symbol	Name	Units
N_t	Population size at time t	Otters
$\dot{D}$	Death rate	Deaths/year
$\dot{B}/N$	Per capita birth rate	Pups/otter pair in 1 year
A	Individual spacing	m^2/otter
cL	Length of coastline occupied	km

The next step is to state what is known about the relation of variables. The relation of interest is coastline length as a function of population size and the other variables:

$$cL = \mathrm{f}\,(N_t, \dot{D}, \dot{B}/N, A) \tag{6.26}$$

It is suspected that otters space themselves into feeding territories along the coast, and hence that spacing A is a constant. To work out dimensions (Step 3 in Table 6.8), we analyze each unit into components. These are:

otters, pups, otter pairs, deaths, years, m^2, and km

The component units are related as follows:

pups = otter after leaving mother
otter pair = 2 otters
death = 1 otter
$\sqrt{m^2} = 0.001$ km

Next, the components are placed into the conventional similarity groups found in textbooks: [L], [T], and [M].

[L]	Length	(km)	[T]	Time	(year)
$[L]^2$	Length^2	(m^2)	[M]	Mass	—

The three dimensions are listed as column headings, the five quantities are listed as rows, and the exponents are placed into the two-way table:

Name of Quantity			Unit Groupings		
	Symbol	**Units**	**[L]**	**[T]**	**[M]**
Population size	N_t	Otters	0	0	0
Death rate	$\dot{D}$	Otter year^{-1}	0	−1	0
Birth rate	$\dot{B}/N$	% year^{-1}	0	−1	0
Spacing	A	m^2 otter^{-1}	2	0	0
Coastline occupied	cL	km	1	0	0

Our first try is unsatisfactory on several counts: N_t has no dimensions, mass [M] has no quantities listed in its column, and $\dot{D}$ and $\dot{B}/N$ are assigned the same dimension, yet they are not similar: addition of %/time to numbers/time cannot be correct. The conventional mechanical dimensions miss much of the biology. The biology of birth and death can be forced into the conventional scheme using body mass as a dimension, but this will introduce a complicating factor, the growth rate of individuals, into a problem where we had no reason to think this matters. Of course, it may well be that individual growth rate *is* important. But at the outset, why complicate the problem unless there is a compelling biological reason to do so? It is simpler to introduce the dimension of entities, which has units of otters for the problem at hand. It replaces the dimension mass. This substitution is truer to the biology of the problem and the initial conjecture, which was that the contraction and expansion of the geographic range depend on births, deaths, and behavioral interactions that determine spacing.

Returning to Step 3 (Table 6.8), the new grouping of components is:

Length	(km)
Length^2	(m^2)
Time	(year)
Entities	(otters, otter pairs, juvenile recruits, deaths)

This grouping is based on the idea that juvenile recruits become equivalent to other otters as soon as they leave their mothers. It suggests that a fledging rate (pup departures/year) would be more relevant to the problem than number of pups per pair in a year. Replacing the quantity $\dot{B}/N$ (pups/pair in a year) with a new quantity $\dot{N}pup$ (pups departing per year) results in a new grouping of component units:

Length	(km)
Length^2	(m^2)
Time	(year)
Entities	(otters, pup departures, deaths)

The revised dimensional matrix is:

Name of Quantity			Unit Groupings		
	Symbol	**Units**	**[L]**	**[T]**	**[#]**
Population size	N_t	Otters	0	0	1
Death rate	$\dot{D}$	Otter year^{-1}	0	−1	1
Pup departure rate	$\dot{N}pup$	Pup year^{-1}	0	−1	1
Spacing	A	m^2 otter^{-1}	2	0	−1
Coastline occupied	cL	km	1	0	0

The number of independent dimensionless ratios will be $5 - 3 = 2$, which are readily obtained (Step 6, Table 6.8) by sequential elimination. Or with Buckingham's method (Table 6.9), we make a preliminary choice of population size and coastline occupied as the variables of interest. This leaves death rate, pup departure, and spacing as the recurring group. However, death rate and pup departure form a dimensionless ratio, so another choice of nonrecurrent variables will be necessary. The logic of variables in relation to measurement dimensions will force us to consider a demographic variable (either death rate or pup departure) as a nonrecurring variable. Choosing departure rate, the matrix can now be partitioned (Step 5, Table 6.9) so that the recurring group is in the denominator of the Π ratio. The dimensionless departure rate is:

$$\Pi_1 = \frac{\dot{N}pup}{N^{\alpha}\dot{D}^{\beta}A^{\gamma}} \tag{6.27a}$$

Examining the dimensional matrix, we can see that the only way to make Π_1 dimensionless is to set the exponents in the denominator of Equation 6.27a to $\alpha = 0$ and $\gamma = 0$, leaving $\beta = 1$:

$$\Pi_1 = \frac{\dot{N}pup}{N^0\dot{D}^1A^0} = \frac{\dot{N}pup}{\dot{D}} \tag{6.27b}$$

The other ratio will be:

$$\Pi_2 = \frac{cL}{N^{1/2}\dot{D}^0A^{1/2}} = cL \cdot N^{-1/2} \cdot A^{-1/2} \tag{6.28}$$

One group is a function of the other:

$$\Pi_2 = \text{function}\,(\Pi_1) \tag{6.29}$$

Interpreting the two ratios (Step 6, Table 6.8), we find that they represent population dynamics (Π_1) and spatial configuration (Π_2). Thus, distribution is an unknown function of population dynamics. It is now evident that in setting up the problem we did not include information on the relation of distribution to population dynamics. If we want to include this relation, we need to introduce a quantity that links numbers to distribution (such as a dispersal rate). With such information we could revise the list of quantities, then fill out a new matrix leading to a better set of dimensionless ratios (Step 8, Table 6.8).

For the problem of population monitoring, the ratio of interest is Π_2 because it contains both coastline length and population size. The scaling relation (Step 9) is:

$$\left(\frac{cL}{cL_{ref}}\right) = \left(\frac{N}{N_{ref}}\right)^{1/2} \cdot A^{1/2} \cdot \Pi_2 \tag{6.30}$$

Equation 6.30 says that occupied coastline increases with the square root of population size. Is this the case for otters? It would be if otters spread in all directions. But if otters occupy adjacent territories along the coast rather than spreading away from the

coast, spacing would need to be measured in units of coastline segment rather than in unit of area per otter. Thus, in designing a monitoring program, we need to determine whether otters spread outward from the coast and whether spacing remains constant. In retrospect it might seem obvious that spacing needs to be monitored, but it is not obvious that a pilot study is needed to determine whether otters occupy areas away from the coast or only along the coast. These conclusions would not have emerged from the more typical approach to the problem, which would have been correlation or regression analysis of the initial set of variables.

To summarize, length of occupied coastline is easier to measure than population size. A monitoring program that tracks change in population size by tracking length of occupied coastline would be inexpensive relative to a survey to estimate population size. Analysis of dimensional groups improved our definition of the set of variables. It also showed that a pilot study of spacing is needed. If we want to monitor otter population size by tracking the occupied coastline, we need to keep track of whether spacing remains constant. Dimensional analysis was thus a guide to the design of a scientifically based monitoring program, one based on testable predictions backed by appropriate pilot studies.

6.5.3 Case Study 3: Patch Size of Phytoplankton

The third case study is a classic in the literature on dimensional methods in ecology. Skellam (1951) and Kierstead and Slobodkin (1953) worked out the patch size at which phytoplankton growth rate (tending to maintain the patch) just balances diffusive motions of the water (tending to erode the patch). The patch diameter at which the opposing processes are equal in magnitude is the critical patch size. Patches larger than this are possible, but smaller patches are not. Platt (1981) showed that the critical diameter could be obtained by dimensional methods without solving an advection-diffusion equation. Platt (1981) and Legendre and Legendre (1998) both applied Buckingham's method to the advection-diffusion equation to obtain the Π ratio that gives the critical patch diameter. With the method of sequential elimination, the advection-diffusion equation is not needed at all.

Here are the quantities of interest, with typical values from O'Brien and Wroblewski (1973):

Symbol	Name	Typical Value	
K_H	Eddy diffusivity	$0.5 \cdot 10^8 \text{cm}^2\text{s}^{-1}$	Sargasso Sea
		$1 \cdot 10^8 \text{cm}^2\text{s}^{-1}$	Peru Upwelling
r	Growth rate	$10\%\ \text{day}^{-1}$	Sargasso Sea
		$5\%\ \text{day}^{-1}$	Peru Upwelling
d	Patch diameter	Km?	

By inspection, eddy diffusivity has dimensions of $[\text{L}]^2\ [\text{T}]^{-1}$ and growth rate has dimensions of $[\text{T}]^{-1}$. Patch diameter has dimensions of [L] by definition. The quantities are listed in the dimensional matrix by row, the dimensions are listed by column, the cells are filled in, and then the matrix is reduced by sequential elimination, as shown in Box 6.6,

following the procedure in Box 6.3. The scaling function and scaling relation are worked out from the dimensionless ratio obtained via sequential elimination. These three expressions have different uses. The dimensionless ratio can be used to compute, in any given case, whether diffusive forces as measured by K_H prevail over patch generation rate $d^2\, r$, which is the product of growth rate r and d^2 (proportional to area). The scaling relation can be used to compute the increase in patch size, given a doubling in the ratio of eddy diffusivity to growth rate. The scaling function can be used to compute the *critical scale*, in this case the critical size of a patch, at which growth rate is adequate to maintain the patch relative to the mixing rate measured as K_H.

Box 6.6 The Critical Patch Scale of Phytoplankton

Sequential elimination will be used to obtain the dimensionless Π ratio, the scaling relation, and the scaling function.

The dimensional matrix has three rows and two columns, hence one dimensionless ratio.

	[L]	[T]
K_H	2	−1
r		−1
d	1	

To reduce the time dimension, take the ratio $K\_H\,/r$.
The new matrix is:

	[L]	[T]
K_H/r	2	0
d	1	0

To reduce the length dimension, divide the second row by the square root of the first row.

	[L]	[T]
$d^1\ (K\_H\,/r)^{-1/2}$	0	0

The dimensionless ratio is: $\Pi = d^1\ (K\_H/r)^{-1/2}$

The scaling function is: $d = \Pi \cdot (K\_H/r)^{1/2}$

Diffusive loss balances growth when $\Pi = 1$.

This gives the critical scale: $d_{crit} = (K\_H/r)^{1/2}$

The scaling relation is: $$\frac{d}{d_{ref}} = \left(\frac{K\_H/r}{K\_H_{ref}/r_{ref}}\right)^{1/2}$$

ANOTHER LOOK AT SECTION 6.5

State the difference among a scaling function, a scaling relation, and a dimensionless ratio. Which (if any) of these have you encountered before?

6.6 Applying the Logic of Dimensions

Text examples of dimensional analysis emphasize the machinery shown in Boxes 6.3, 6.4, and 6.5. However, the logic of dimensions is far more flexible than the machinery of forming independent ratios. Kline (1976) emphasizes that reduction to the minimum number of dimensionless ratios will often be too limiting in situations where little is known about the problem or about the component variables. Instead of the focus on the minimum number of independent ratios, Kline recommends choosing from three approaches, which he calls the *method of similitude, the method of Buckingham's Π theorem*, and *the method of governing equations*. These are described in Table 6.10.

Text examples convey the impression that the goal of analysis is a unique set of dimensionless ratios that are independent of one another. The disillusionment in the quote at the beginning of this chapter results when a unique set of dimensionless ratios fails to appear. In simple text problems there is often only one ratio (as in the dinosaur and phytoplankton examples), resulting in a unique solution. For simple problems such as the otter example, the only possible ratios are rearrangements of the minimum number of ratios. In contrast, many problems in ecology (e.g., Miller et al., 1984; Hatcher and Frith, 1985; Yager et al., 1993) begin with a large number of variables that do not fall into convenient groups. The method of similitude (Table 6.10) is most useful in these situations, when there are many variables and no unique set of independent ratios. This method is the most flexible, relying on past knowledge and judgment.

The method of similitude (Kline, 1976) begins with a list of relevant quantities, takes all possible dimensionless ratios, and then relies on judgment and knowledge to choose the appropriate scaling relations. An example comes from Hatcher and Frith (1985), who were interested in the nitrogen dynamics of coral reefs. Hatcher and Frith listed six variables that affect the total nitrogen accumulation by a reef (Box 6.6). Relative to dimensions of mass [M], length [L], and time [T], this results in three indepen-

Table 6.10 Three Strategies for Applying the Logic of Dimensions to a Problem

Method of similitude:

1. List all possible dimensionless ratios.
2. Choose among these, based on knowledge of the problem.

Method of the Π theorem:

1. Reduce variables to the minimum number of Π ratios.
2. If there is only one ratio, apply this (either as a scaling function or a scaling relation).
3. If there is more than one ratio, compute these from data, then plot the Π ratios against one another.

Method of governing equations:

1. State the conservation equation.
2. Use this rather than the dimensional matrix to obtain dimensionless ratios.

Adapted from Kline (1976).

dent Π ratios. Inspection of the dimensional matrix reveals that there are a large number of ways of forming the three ratios. Two of the variables have dimensions of length but not mass or time. This suggests a "natural" Π ratio, the area in which nitrogen is fixed to the area from which nitrogen is lost by denitrification, $\Pi_1 = A_{fix}/A_{denitr}$. This leaves four variables, all fluxes with the same units, to be combined into two additional ratios. There are six simple flux ratios. This means there are three possible pairs of ratios Π_2 and Π_3 (Box 6.7), not just one pair. Based on knowledge of reef systems, Hatcher and Frith (1985) defined an internal fixation ratio, the nitrogen fixation time scale (inverse of fixation rate) relative to horizontal mixing time scale (inverse of mixing rate). Hence $\Pi_2 = [\dot{N}_{mix}]/[\dot{N}_{fix}]$. This leaves one independent ratio to be formed from the two remaining variables: $\Pi_3 = [\dot{N}_{denitr}]/[\dot{N}_{adv}]$. This ratio is that of total loss to advective resupply.

Rather than relying on the textbook machinery to produce Π_2 and Π_3, Hatcher and Frith (1985) used knowledge of coral reef biology to form two dimensionless ratios, an internal fixation number Π_{fix} and an internal denitrification number Π_{denitr}. These ratios scale denitrification and nitrogen fixation to the mixing loss $[\dot{N}_{mix}]$. Box 6.7 shows the

Box 6.7 Dimensional Analysis of Nitrogen Dynamics of a Coral Reef

DIN = dissolved inorganic nitrogen
$A \equiv$ Area (m^{-2}) $\quad [\dot{N}] \equiv$ nitrogen flux (μmoles DIN m^{-2} hr^{-1})

$[\dot{N}_{tot}] = f([\dot{N}_{fix}], A_{fix}, [\dot{N}_{denit}], A_{loss}, [\dot{N}_{mix}], [\dot{N}_{adv}])$

		[M]	[L]	[T]
$[\dot{N}_{tot}]$	Nitrogen accumulation	1	−2	−1
$[\dot{N}_{fix}]$	Nitrogen fixation rate	1	−2	−1
A_{fix}	Area over which fixation occurs		2	
$[\dot{N}_{denitr}]$	Fixed nitrogen loss rate (denitrification)	1	−2	−1
A_{denitr}	Area of denitrification losses to atmosphere		2	
$[\dot{N}_{mix}]$	Nitrogen loss due to horizontal mixing	1	−2	−1
$[\dot{N}_{adv}]$	Advective input	1	−2	−1

Variables − dimensions = 3 independent Π ratios.
Reduction by inspection (Box 6.2):

$$\Pi_1 = A_{denitr}/A_{fix} \qquad \Pi_2 = [\dot{N}_{mix}]/[\dot{N}_{fix}] \qquad \Pi_3 = [\dot{N}_{denitr}]/[\dot{N}_{adv}].$$

Reduction based on coral reef biology (Hatcher and Frith, 1985):

$$\Pi_{fix} = ([\dot{N}_{mix}]/[\dot{N}_{fix}])\cdot((A_{denitr} + A_{fix}) / A_{fix})$$
$$\Pi_{denitr} = ([\dot{N}_{mix}]/[\dot{N}_{denitr}])\cdot(A_{denitr} + A_{fix}) / A_{fix})$$

Π_{fix} can be rewritten in terms Π_1 and Π_2:

$$(A_{denitr} + A_{fix})/A_{fix} = \Pi_1 + 1$$
$$\Pi_{fix} = \Pi_2 (\Pi_1 + 1)$$

Π_{denitr} cannot be rewritten in terms of the three independent ratios, hence Π_{denitr} and Π_{fix} are not independent.

From Hatcher and Frith (1985).

relation of these two Π ratios to three independent Π ratios produced by the machinery (Box 6.3) of sequential elimination. Both biologically based ratios contain $[\dot{N}_{mix}]$ and hence are not independent. Due to the built-in relation, little is gained by plotting the ratios against each other. But the ratios did prove useful in making order of magnitude comparisons of denitrification versus fixation rate, both taken relative to mixing loss. The example, which uses Table 6.10 as the starting point, exhibits the interplay of biological reasoning with dimensional analysis.

The method of the Π theorem (Tables 6.9 and 6.10) emphasizes the number of independent ratios (based on Buckingham's theorem). The advantage of this method is that a plot of independent ratios will provide a more compact summary than several plots of variables one against another. Plots of ratios will be free of spurious associations found in plots of variables (an example is the association of percent change with initial density). To illustrate the advantage of plotting ratios rather than variables, the method of the Π theorem is applied to a well-known empirical relation (Ryder, 1965), that of annual fish catch from lakes as a power function of the morphoedaphic index, the ratio of total dissolved solids to lake depth (Figure 6.1). This empirical relation summarizes the tendency of large, clear lakes to be unproductive relative to small or shallow lakes. What is the contribution of lake shape and clarity to fish production? To undertake dimensional analysis of the question, we begin by disaggregating Ryder's (1965) derived variable, H = kg ha^{-1}yr^{-1} into its component measurements, annual fish catch $\dot{M}$ = kg yr^{-1} and area A = hectares. Box 6.8 shows the dimensional analysis.

Box 6.8 Dimensional Analysis of Fish Catch Relative to the Morphoedaphic Index of Ryder (1965)

$H = k \cdot MEI^{0.4461}$

H = fish catch per unit area (kg ha^{-1}yr^{-1})
TDS = total dissolved solids (ppm)
z = lake depth (meters)
MEI = TDS/z = morphoedaphic index of Ryder (1965)

The functional expression for the measured variables:

$\dot{M} = f(A, TDS, z)$

	[L]
A = lake area (ha)	2
TDS = total dissolved solids (ppm)	0
z = lake depth (meters)	1

Variables − dimensions = 2 independent Π ratios
Reduction by inspection (Box 6.3)

$\Pi_1 = A^{1/2}/z \qquad \Pi_2 = TDS$
$\dot{M} = f(\Pi_1, \Pi_2)$

Evaluate by plotting Π_1 versus Π_2 then $\dot{M}$ against Π_1 and Π_2.

Once the Π ratios are formed, they are plotted against each other. Such a plot shows whether variables are related, without spurious relations forced by the presence of a measurement unit common to both axes. In Figure 6.1, for example, the relation of fish catch per unit area might be driven in part by the relation of lake area (which appears on the left side of the equation) with lake depth (which appears on the right). Figure 6.2 shows that lake shape Π_1 and water clarity Π_2 are independent.

Because Π_1 and Π_2 were independent, it was then of interest to plot fish catch against each ratio. Figure 6.3 shows that annual catch is related to lake shape, described by the dimensionless ratio of area$^{1/2}$/depth. Figure 6.4 shows that annual catch is unrelated to total dissolved solids. The relation of fish catch per unit area to the morphoedaphic index (Figure 6.1) is driven by the lake shape. Based on this analysis, a simplification of Ryder's (1965) model was developed (see Chapter 13) and tested.

Plots such as those in Figure 6.2 are easy to construct and effective in disentangling the relations among multiple variables. Plots based on the method of Π ratios have been used to examine the form and function of dinosaurs (Alexander, 1989), the growth dynamics of phytoplankton (Lewis et al., 1984; Gallegos and Platt, 1985), the mixing characteristics in experimental mesocosms (Crawford and Sanford, 2001), and local food supply to benthic marine organisms (Miller et al., 1984; Yager et al., 1993).

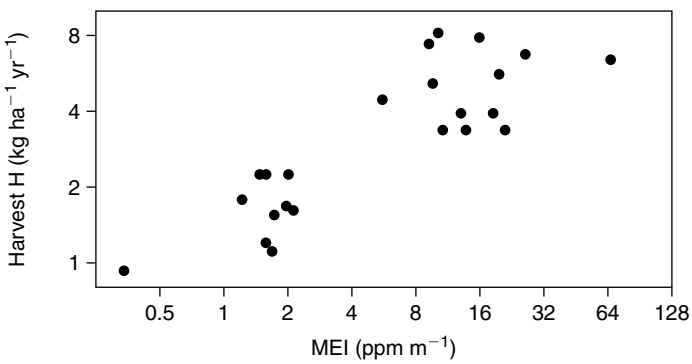


FIGURE 6.1 Relation of Fish Catch per Unit Area *H* to Ryder's Morphoedaphic Index *MEI*. Data from Ryder (1965).

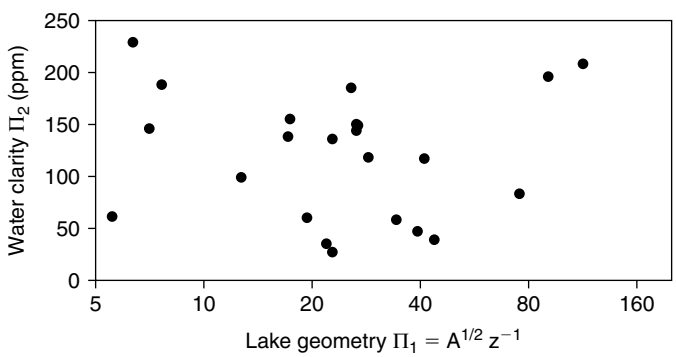


FIGURE 6.2 Dimensional Analysis (Box 6.8) of Lake Shapte (Π_1) versus Water Clarity (Π_2).

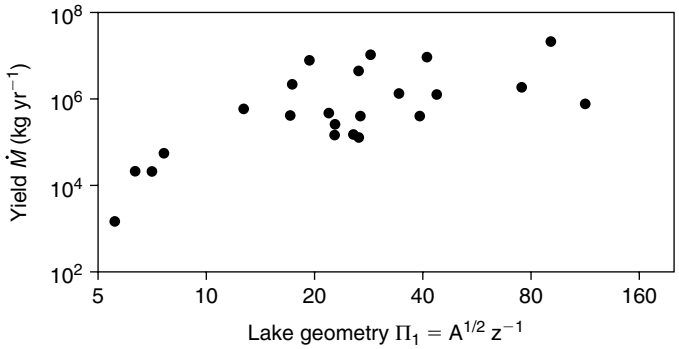


FIGURE 6.3 Fish Yield versus Lake Shape (Π_1) from Dimensional Analysis (Box 6.8).

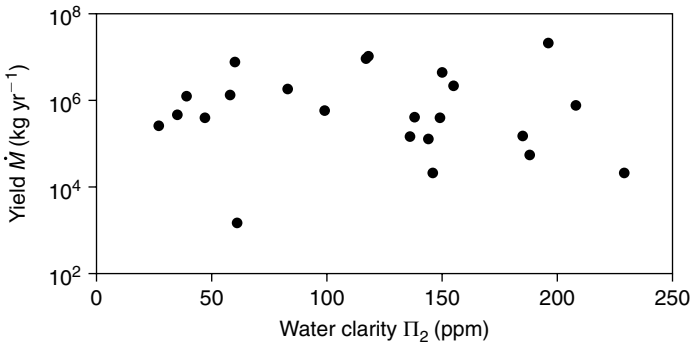


FIGURE 6.4 Fish Yield versus Water Clarity (Π_2) from Dimensional Analysis (Box 6.8).

The method of governing equations (Table 6.10) represents a third and different use of the logic of dimensions. This method begins with an equation rather than proceeding toward the development of a scaling function. As with the method of Π ratios, this method is used to develop or modify models and so will be covered in Chapter 13.

ANOTHER LOOK AT SECTION 6.6

For a problem of interest to you, write a functional expression $Y = f(X1, X2,....)$, form a dimensional matrix, and determine the number of dimensionless ratios. Then determine whether there is a unique set of Π ratios or many possible Π ratios. Comment on the applicability of the method of the Π theorem (Tables 6.9 and 6.10) to your problem.

6.7 Use and Limitations of Dimensions

Similarity groups, or dimensions, are an important part of quantitative reasoning with scaled quantities. Similarity groups have a variety of uses. Lists by Platt (1981) and Peterson and Hastings (2001) overlap Kline's (1976) list of 10 uses grouped into three categories: simplification of a problem, checking equations, and developing equations. Simplification includes the use of dimensions to aid comprehension of abstract quantities,

the standardization of disparate measurements to the same scale, and guidance of further research. The dinosaur and otter examples demonstrate the utility of simplification guided by the logic of similarity groups. Chapter 13 covers checking and developing equations.

Dimensional groupings aid in visualizing the physical or biological content of any equation and its constituent symbols. For example, the Coriolis parameter f occurs often in descriptions of atmospheric and ocean circulation. The physical content of the symbol becomes clearer from its dimensional symbol, which is T^{-1}. The dimension shows that the abstract symbol f has the dimensions of frequency. It is the frequency or angular component of acceleration acquired by an air or water parcel once it has been set in motion and so moves either closer or farther from the earth's axis of rotation. The same approach works for equations: one writes out the dimensional symbols immediately above or below each symbol in the expression. This will make the idea behind the expression more comprehensible by relating it to familiar ideas of mass, time, distance, energy, or contacts. My copy of *Circulation in the Coastal Ocean* (Csanady, 1982) had two blank pages at the end that are now filled with a list of symbols for physical quantities, their names, their dimensions, and the page numbers where they were first defined. Dimensional grouping was key to understanding the physical content of Csanady's equations for calculating flows in coastal waters.

Most of the applications of the logic of dimensions in the ecological literature are for organism form and function in relation to body size (Hill, 1950; Gunther, 1975; Schmidt-Nielsen, 1984; Alexander, 1989; Pennycuick, 1992). An example is the case study of dinosaur running speed in the previous section. Several studies have applied the logic of dimensions to ethology (Dunbar and Stephens, 1993; Dugatkin and Mesterton-Gibbons, 1995). Applications at the population or ecosystem level, such as the otter monitoring and phytoplankton growth case studies, are less common. Most of the published applications are from aquatic habitats. Many begin with a conservation equation rather than a set of variables.

A few published applications begin with a set of variables rather than an equation. These use the logic of dimensions to interpret a body of literature (Vogel, 1981; Frith and Hatcher, 1985; Duffy and Schneider, 1994), guide experimentation (Miller et al., 1984; Uhlmann, 1985; Yager et al., 1993; Englund and Peterson, 2005), or synthesize information from disparate sources (Neis et al., 1999). Vogel (1981) uses the many dimensionless ratios known from fluid mechanics to interpret a wide variety of biological phenomena. Vogel discusses Reynolds, Prandtl, Nusselt, Grasshof, and Froude numbers, their variants, and another nine lesser used numbers. A striking application is the dynamic similarity that underlies a surprising resemblance at disparate scales: fiber pattern inside plant cell walls, compared to pattern in glaciers surrounding mountain peaks. Vogel shows that both phenomena operate at low Reynolds numbers, where viscous forces prevail over forces associated with the movement of objects.

Hatcher and Frith (1985) used dimensional analysis to reduce six variables that affect the nitrogen economy of coral reefs to two dimensionless ratios, which they used to interpret disparate conclusions about whether coral reefs depend on external supplies of nitrogen. Duffy and Schneider (1994) devised a sequence of five dimensionless ratios to evaluate whether fishing vessels compete with seabirds to the detriment of seabird reproductive success (Figure 6.5). The ratios were listed in order of cost to obtain data, starting with the least expensive ratio, for which data are often available. Uhlmann (1985) and Schneider (2001a) described several applications of the logic of dimensions to biological production by enclosed bodies of water, from chemostats and laboratory

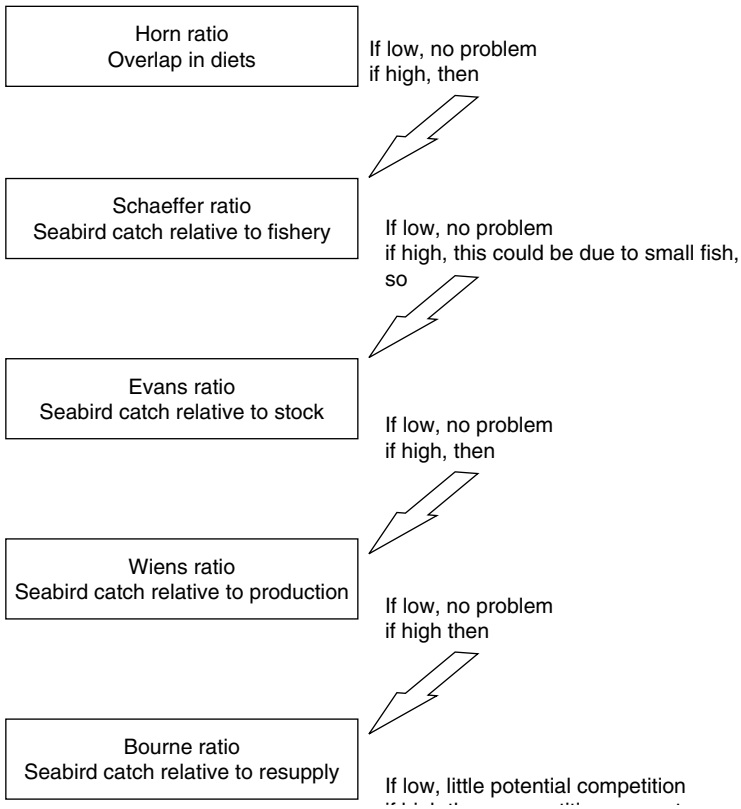


FIGURE 6.5 Sequence of Dimensionless Ratios to Evaluate the Potential for Competitive Interactions between Seabirds and Fisheries. Redrawn from Duffy and Schneider (1994).

tanks to sewage lagoons and lakes. In marine habitats, Miller et al. (1984) listed eight variables that characterize feeding dynamics of benthic animals. This reduced to five dimensionless ratios. Three of these brought coherence to previous results. The dimensionless ratios permit previous models of sediment ingestion to be matched to appropriate sedimentary regimes. Thus the ratios can serve to guide the design of experiments on foraging behavior in relation to modes and rates of sediment transport.

Yager et al. (1993) used dimensionless ratios to guide design of experiments to uncover the conditions leading to enhanced food supply to deposit feeders inhabiting sedimentary pits. They looked at trapping rate of particles by pits as a function of 10 variables, which they reduced to seven dimensionless ratios. In their experiments, they held three of these ratios constant, then varied either the pit-aspect ratio or one of three ratios summarizing the effects of flow turbulence. Simplification via the logic of dimensions allowed Yager et al. (1993) to describe the ways in which food supply depends on flow regime. Schneider (2001a) and Peterson and Englund (2005) used dimensional analysis to evaluate alternative explanations for experimental results for mesocosms (tanks) of different size and shape.

These examples show the flexibility of the logic of dimensions, beyond the text examples that use conventional dimensions and the machinery for obtaining the minimum number of ratios. Instead of conventional dimensions, Miller et al. (1984) used dimensions of particles [P], animals [M], and time [T]. These dimensions were truer to

the biology of the research question than the conventional dimensions of mass, length, and time. Duffy and Schneider (1994) developed a sequence of dimensionless ratios relevant to a stated problem rather than reducing a set of potentially measurable variables to a small number of dimensionless ratios. Hatcher and Frith (1985) chose dimensionless ratios that they considered relevant to the lagoon nitrogen dynamics rather than choosing ratios produced by a standard dimensional analysis.

In addition to its utility in specific applications, the method of dimensional grouping is an important avenue for interdisciplinary understanding of ecological processes. One of the advantages in using dimensional groupings to reason about ecological problems is that this method is routinely used in reasoning about physiological processes that connect organisms to their environment. The method is used by physiologists working at space and time scales relevant to the cell or the individual. It is also used by physical scientists working at global time and space scales. So this method should be of considerable use to ecologists who work in between, where physiological performance (growth rate, birth rate) interacts with the dynamics of the physical environment.

Similarity groups have played a central role in the analysis of the form and function of organisms, beginning with Thompson's landmark book on form and function (1917) and continuing into the present. Similarity groups are essential in *biofluid dynamics* (Vogel, 1981), which, broadly defined, includes any application of fluid mechanics to biology. Stephens and Dunbar (1993) and Mesterton-Gibbons and Dugatkin (1995) use dimensional methods to develop models of behavior independent of species and the particularities of place. Hastings (1997) showed that stability of population models depends on parameters reduced to dimensionless form. Similarity groups have made important contributions to biological oceanography (Lewis et al., 1984; Miller et al., 1984; Gallegos and Platt, 1985; Yager et al., 1993) and sensory ecology in fluid environments (Weissburg and Zimmer-Faust, 1993; Nevitt, 2000; Weissberg, 2000). More recently, dimensional methods and similarity groups have put mesocosm experiments on a sound footing by quantifying the degree to which conditions in tanks do and do not resemble field conditions (Sanford, 1997; Porter et al., 2000; Crawford and Sanford, 2001; Englund and Peterson, 2005).

However, like the girl in the quote at the beginning of the chapter, too much can be expected of classical dimensional methods. Texts (including this chapter) illustrate the method with examples that work. Texts typically use integer dimensions based on noniterative measurement protocols that use *complete similarity* (Barenblatt, 1996), where a few ratios completely govern the dynamics. Accessible texts typically do not treat nonintegral dimensions, iteratively measured quantities, and what Barenblatt (1996) calls *incomplete similarity*, where the governing effect of ratios is asymptotic, not complete. In biology, the Euclidean geometry of lines, planes, and volumes (spherical chickens) is at best a convenient approximation. The geometry of organisms and their environment is that of Hausdorf (1919) and Mandelbrot (1977), with paths more crooked than Euclidean lines or surfaces more rugged than Euclidean planes. Noninteger exponents emerge when more than one similarity criterion applies or when opposing rates are usually far from equilibrium. The failure of physical and biological systems to follow the scaling relations obtained by classical dimensional methods has long been recognized. It is attributed to mixed scaling regimes (Gunther, 1975; Uhlmann, 1985; Barenblatt, 1996) instead of a single regime that can be analyzed by the classical methods presented in this chapter. In mixed regimes the exponents are not integers or ratios of integers, and the classical approach to dimensional analysis described in this chapter will fail (as with the cautionary quote that leads this chapter). A different approach is needed. This approach, called *renormalization*, is beyond the scope of this book.

ANOTHER LOOK AT SECTION 6.7

At what point in your life did you encounter dimensions?

- High school courses
- University physics or chemistry; other university courses
- After university
- Never

Do you think your experience was typical?

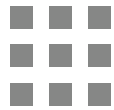
6.8 Further Reading on Dimensions and Measurement Theory

Reasoning with dimensional quantities can be traced to Galileo (1638) and Newton (1686). Unit grouping in a dimensional matrix was invented by Fourier (1822, Chapter 2, Section 9) to analyze the geometry of heat flow. Buckingham (1914) stated the Π theorem about the minimum number of independent dimensional groupings in response to the claim (Tolman, 1914) that if all measurements were doubled, no one would notice. Whitney (1968) presents an exposition of the mathematical basis of dimensional groupings. Older texts (Bridgman, 1922; Langhaar, 1951; Taylor, 1974) emphasize equations and the minimum number of independent ratios. The texts by Kline (1976) and Massey (1986) provide more flexible treatments of the logic of dimensions. Stahl (1961, 1962) extended rigorous treatment of dimensional groupings to biology. General treatment of dimensional analysis in biology can be found in Gold (1977), Platt (1981), Legendre and Legendre (1998), Pennycuick (1992), and Petersen and Hastings (2001). An advanced treatment, including fractals and a few biological examples, appears in Barenblatt (1996).

A general account of measurement theory can be found in Krantz et al. (1971). Philosophical accounts of measurement theory can be found in Ellis (1966), who argues that counts cannot be assigned dimensions, and Kyburg (1984), who argues that all measurements must be assigned dimensions. Falconer (1985) develops fractal concepts from measure theory.

Defined Concepts and Terms for Review and Future Reference

____ biological entity
____ Buckingham's theorem
____ complete vs incomplete similarity
____ composite dimensions
____ critical scale
____ dimension
____ dimensional analysis
____ dimensional matrix
____ iterative counting relation
____ measurement scale
____ operational definition for a dimension
____ Π ratio
____ similarity group



7 The Geography and Chronology of Quantities

When the comet crossed the orbit of the moon it was moving at a velocity of 30 kilometres per second and the end of the Cretaceous was three hours away.

—D. A. Russell, *An Odyssey in Time: The Dinosaurs of North America, 1989, p. 205*

7.1 Synopsis

Measured quantities have temporal attributes: their chronology, resolution, and duration. Temporal attributes are usually expressed on a ratio type of scale. But in the absence of detailed information, it is useful to express these attributes on ordinal or even nominal scales. Attributes on a ratio type of scale readily serve as a vector address for each value of the quantity. The alternative to vector representation of a quantity is to adopt an index to match each value of a quantity to its temporal attributes, which are now treated as quantities themselves.

The spatial attributes of quantities (position, resolution, and extent) are of interest to any environmental scientist, whether a biologist, chemist, economist, engineer, geographer, geologist, meteorologist, or physicist. Geographic attributes are expressed on measurement scales ranging from nominal to ratio, depending on the level of available detail and the purposes of an investigation. Stating the type of geographic measurement scale aids considerably in making sense of the confusing array of geographically explicit techniques now available.

Computers make it possible to analyze a quantity relative to all its spatial attributes, not just position. Scanning through a sequence of positions, together with zooming in on detail at finer resolution or zooming back on larger-scale pattern, contribute to understanding of populations and communities in relation to their environment.

Clear and consistent notation is an important part of quantitative biology. Attention to the principles of good notation can greatly improve quantitative practice in ecology, as it has in physiology and oceanography. The notation used for spatially explicit quantities in oceanography and meteorology is appropriate for any geographic topic, whether rocks, life, or economics.

The spatial and temporal attributes of a quantity are conveniently and efficiently summarized in the compact form of vector notation. In this notation, the symbol $Q_{\boldsymbol{x}\,\boldsymbol{t}}$ stands for the spatial attributes $\boldsymbol{x}$ and temporal attributes $\boldsymbol{t}$ of a quantity. This notation lets us describe and quantify the spatial and temporal scale of any quantity.

doi: 10.1016/B978-0-12-627865-1.00007-6

7.2 Temporal Attributes

Measured quantities have several *temporal attributes*: the duration of each measurement, the time between successive measurements, and the time required to complete the set of measurements. These characteristics determine the time scale of a measured quantity. The duration of a single measurement sets the resolution, whereas the time between first and last measurement sets the range or extent. In principle any measurement can be indexed as to time and duration. In practice the art lies in judging whether temporal characteristics matter. For example, in the experiments summarized by Downing et al. (1999), the time taken to measure nutrient concentration does not matter, as long as it is short relative to the rate of change in concentration. The duration of each experiment does matter because of lag effects on nutrient uptake at short time scales and because of grazing effects at longer time scales (Downing et al., 1999). If a nutrient uptake experiment is repeated over several days, the time between experiments should be irrelevant to the outcome; a thorough experimentalist would, of course, check to make sure that this was true.

Similar considerations apply to the temporal attributes of survey variables. In a survey of the number of shorebirds using a beach, the temporal attributes are the count duration (e.g., 1 hour), count frequency (e.g., daily), and duration of the series (e.g., three months). The temporal scope of this survey 90:

$$Scope = \frac{range}{resolution} = \frac{3\,\text{months}}{1\,\text{day}} \cdot \frac{30\,\text{days}}{\text{month}} = 90$$

The survey cannot resolve dynamics at time scales of less than a day, such as diurnal or tidal cycle change.

There are several ways of representing the temporal attributes of a quantity. One is to use subscripts that have the values of the attribute. With this notation, the sequence of counts of one species of migratory shorebird, the Willet, *Catoptrophorus semipalmatus,* along a 5.7 km stretch of beach along the Gulf of Mexico on four successive days at the beginning of the period of northward migration is represented as:

$$N_t = [150\ 145\ 82\ 111] \cdot 5.7^{-1} \cdot \text{km}^{-1}$$
$$t = 17\ \text{April 86}\ldots 20\ \text{April 86, 1 day}$$

The subscript t ranges from April 17, 1986, to April 20, 1986, with a resolution of 1 day. This increases the information about the quantity N. Without the temporal attribute, all one can say is that the Willet count ranged from 150 to 82, with an average value of 122 per 5.7 km. Given the temporal attribute, one can say that the Willet count decreased suddenly on April 19.

On April 17–19, a high-pressure weather system passed through the area, resulting in winds favorable for northward migration after passage of the system. This suggests the hypothesis that a rise in air pressure stirs migratory restlessness, with departure occurring as soon as the barometer stops rising. The hypothesis that high-pressure systems trigger migratory departure could be tested by continuing the observations to determine whether the passage of a subsequent system again resulted in a drop in numbers of Willets.

The use of the attribute as a subscript can handle missing values, as in the following example:

$$N_t = [182 \text{ nc } 150\ 145\ 82\ 111\ ] \cdot \text{Willets } 5.7^{-1} \cdot \text{km}^{-1}$$
$$t = 15 \text{ April } 86, 20 \text{ April } 86\ldots 1 \text{ day}$$

The subscript ranges from April 15 to 20, with a resolution of 1 day. The first count was $N(15 \text{ April } 1986) = 182$ Willets 5.7^{-1}km^{-1} between 8:00 and 10:00 A.M.; there was no count on April 16.

This notation becomes cumbersome for sparsely populated time series, which have a large number of missing values within the range of the temporal attribute. The notation cannot be used if the resolution changes or if temporal attributes are on a nominal scale (before or after a certain date) or on an ordinal type of scale. The solution is to list the attribute as a separate quantity, then employ a sequence of integer numbers ($i = 1\ldots n, 1$) to address both the quantity and its attributes. The index i has no units; it is here displayed as a range (1 to n) at a resolution of 1. Alternatively, the index is displayed as ($i = 1, 2, 3\ldots n$). Here is an of example a sequence of 5 measurements of biomass (M_t in units of grams), on days 2, 6, 7, and 14:

$$M_t = [15.0_{\text{day 2}}\ 18.1_{\text{day 6}}\ 18.1_{\text{day 7}}\ 20.2_{\text{day 14}}\ 20.1_{\text{day 14}}] \cdot \text{grams}$$

The temporal attribute (day of measurement) will no longer serve as an address, because there are two values on day 14. The quantity M is reorganized as two quantities coordinated by an index i:

$$\begin{array}{lcccccc} i & = & 1 & 2 & 3 & 4 & 5 \\ M_i & = & [15.0 & 18.1 & 18.1 & 20.2 & 20.1\ ] \cdot \text{grams} \\ t_i & = & [\ 2 & 6 & 7 & 14 & 14\ ] \cdot \text{days} \end{array}$$

The index i serves as a bookkeeping device to keep track of a sequence of addresses. With this notation, missing values are not displayed and all measurements of mass have unique addresses, even if they have the same temporal attribute.

The same reorganization to indirect addressing works if the quantity of interest is expressed on an ordinal scale:

$$\begin{array}{lcccccc} i & = & 1 & 2 & 3 & 4 & 5 \\ M_i & = & [\text{ First} & \text{Second} & \text{Second} & \text{Fourth} & \text{Third}\ ] \\ t_i & = & [\ 2 & 6 & 7 & 14 & 14\ ] \cdot \text{days} \end{array}$$

This notation becomes necessary if the temporal attribute is on a rank (ordinal) scale:

$$\begin{array}{lcccccc} i & = & 1 & 2 & 3 & 4 & 5 \\ M_i & = & [\ 15.0 & 18.1 & 18.1 & 20.2 & 20.1\ ] \cdot \text{grams} \\ t_i & = & [\ \text{First} & \text{Second} & \text{Third} & \text{Fourth} & \text{Fourth}\ ] \cdot \text{day} \end{array}$$

Indirect addressing will also become necessary if the temporal characteristics are expressed on a nominal scale. To test the hypothesis of meteorological triggering of Willet migration, I continued the counts at the same beach until another low-pressure system arrived. The temporal characteristic of interest was whether the count occurred before the passage of a system ($t = 0$) or during and after ($t = 1$). The temporal attributes are expressed on a binomial scale, represented by t_i. The temporal attribute t_i has

only two values so it cannot be used as an index. The index i is needed to coordinate the quantity N with its temporal attributes, either before ($t_i = 0$) or after ($t_i = 1$) the passage of the high-pressure system.

$$\begin{array}{lcl} i & = & [\ 1 \quad 2 \quad 3 \quad 4 \quad 5\ ] \\ N_i & = & [\ 88 \quad 104 \quad 130 \quad 78 \quad 58\ ] \cdot \text{Willets } 5.7^{-1} \cdot \text{km}^{-1} \\ t_i & = & [\ 0 \quad 0 \quad 0 \quad 1 \quad 1\ ] \cdot \text{after front} \end{array}$$

When a temporal attribute is on a ratio type of scale, the one-to-one correspondence of each value of a quantity with a value of an attribute allows construction of scaling and measurement relations, as described in Section 7.6 of this chapter.

ANOTHER LOOK AT SECTION 7.2

Name a quantity of interest to you, measured at irregular times. Write out four typical values in a vector with units; then, beneath each value, list a typical temporal resolution, time since previous measurement, and time since initial measurement. State the range, resolution, and scope.

7.3 Geographic Attributes in One Dimension

Measured quantities have *geographic attributes*—their resolution, position, separation, and spatial extent. A measurement occupies a finite distance, area, or volume, just as a measurement occurs within a finite length of time. This is the *spatial resolution*, or *grain* (Wiens, 1989; Legendre and Legendre, 1998). A second geographic attribute is the distance between any two measurements. For contiguous measurements, the separation (taken as the distance between midpoints of each measurement) will equal the resolution. The range or extent (Legendre and Legendre, 1998) is the greatest separation between two measurements in a set of measurements. The spatial scope is the ratio of the range to the resolution.

Any measurement can be assigned spatial attributes; these will matter in some but not all situations. In a laboratory experiment, for example, one goal is to obtain results that are independent of location. An experimentalist wants measurement outcomes that are repeatable by somebody else, regardless of where the lab happens to be located. The goal of repeatability applies to field studies, which also seek results that are independent of location. In field studies, however, geographic variability is often substantial and usually cannot be controlled by experimental manipulation. Results often depend on location and spatial scale. When this happens we look to spatial attributes to understand the observed variability.

As with temporal attributes, we have two modes for recording spatial attributes—either in vector form or by conversion to a quantity with shared index. Here is a series of Willet counts made along a sequence of 1 km stretches of beach at increasing distance from a pass, or break, between two barrier islands:

$$N_x = [\ 5\ 23\ 12\ 11\ 9\ 0\ 1\ ] \cdot \text{Willets km}^{-1}$$

Here the count at address $x = 0$ is $N(0) = 5$ Willets km^{-1} at 7:00 A.M. on April 26, 1987, at latitude 26° 28.96' N, longitude 82° 11.00' W. The counts are shown as an ordered sequence, where the subscript x represents the geographic ordering, or address, along the beach. In this case the value of x is also the temporal ordering of the data in 10-minute intervals.

Willets vary considerably in numbers along the beach. Some variability is obviously due to social factors because Willets often forage in loose groups. Variability also results from response to environmental factors such as prey abundance, wave action, and disturbance by people. To understand the variability, we need to examine counts in light of their spatial attributes, notably resolution (1 km segments) and extent (7 km stretch of beach). The spatial attribute x suggests that there is a spatial trend or gradient in density from km 2 to km 7. Without the spatial data, all one could say is that Willet density ranged from 0 to 23 km^{-1}, with a median value of 8.7 km^{-1}.

Here is another example of geographic variation in Willet numbers, but this time with uneven spacing and resolution. The counts were made along a 20 km stretch of beach on April 27, 1986, using sections of irregular but known length. The index i becomes necessary to coordinate several quantities: the count N_i, the distance x_i from the start point (latitude 26° 28.96' N, longitude 82° 11.00' W), and the extent L_i of each segment.

i	=	1	2	3	4	5	
N_i	= [	7.4	2.3	0.7	2.1	0.7	] · Willets km^{-1}
x_i	= [	3.2	8.1	11.7	13.7	17.6	] · km
L_i	= [	6.4	3.5	2.9	1.9	5.8	] · km

Two spatial attributes are displayed. L_i is the length of each of five successive stretches of beach. x_i is the distance of the midpoint of each successive stretch of beach from the north end of a barrier island. The spatial range of these counts is 20 km, the entire length of a barrier island, Sanibel, on the west coast of Florida, in the Gulf of Mexico. The range is calculated as the sum of the stretches L_i, for which the symbol is ΣL_i. The spatial units were contiguous, and so separation between midpoints also varied (e.g., $(6.4 + 3.5)/2 = 4.95$ km from the first to second unit). The average resolution was Mean$(L_i) = 4.0$ km. This is a somewhat inconvenient scale because of the variable resolution, but it is by no means uncommon. At this spatial scale (scope = 20 km/4 km = 5), there were distinct differences in Willet numbers between the northwest and southeast ends of the island. This notation correctly and compactly represents counts with irregular spatial attributes. It would take some effort to represent these counts and their spatial attributes accurately with a geographic information system in which the default is a regular spatial grid.

The spatial attributes of a quantity can be expressed on any of the four types of measurement scale. Experimental or process-oriented studies often result in geographic attributes on a nominal scale. For example, measurements from a reciprocal transplant experiment will have spatial attributes in four categories: sample from location A moved to A, sample from A moved to B, sample from B moved to B, and sample from B moved to A. Surveys produce measurements with geographic attributes on any of the four types of measurement scale. An example in one spatial dimension is measurement along a transect up a mountain. The positional attribute of a measurement can be expressed on a ratio scale, the distance from a zero point such as the base of the mountain. We could express

the same attribute on an interval scale, as when we use geographic coordinates. A ranking of sites with respect to elevation can be used. In studying the effects of altitude, it may be sufficient merely to rank the sites by height. We may find it sufficient to define mountain habitat in just three classes: above, at, or below the tree line.

The unit of spatial resolution of a measurement can be expressed in lengths, areas, or volumes. The size of the unit will depend on the measurement procedure and the quantity of interest. For the Willets, the resolution was conveniently expressed as segments of beach on the order of a km. For the mole crab *Emerita analoga* on which the Willets feed, the resolution along the same beach would usually be expressed in numbers per area of sand sampled. For the particulate matter on which the mole crabs feed, measurement might well be a count per unit of volume of water washing over the beach.

As with temporal units, the one-to-one correspondence of each value of a quantity with a value of a spatial attribute on a ratio scale allows construction of spatial scaling functions (as described in Section 7.6 of this chapter).

ANOTHER LOOK AT SECTION 7.3

To measure nematode abundance, 10 samples of soil, each with a volume of 10 cm^3, are collected at random along a 10 m transect. State the units of the quantity and then describe the separation between samples and the extent of the survey.

7.3.1 Application: Evaluation of Geographically Explicit Studies

The distinction between nominal, ordinal, and ratio types of geographic scales helps in comparing and evaluating published studies, whether theoretical or empirical. Geographically explicit studies have become increasingly important in process-oriented research (Bell, McCoy, and Mushinsky, 1991) and have always been important in applied ecological research in wildlife and fisheries, where spatially extensive surveys are common. Geographic attributes vary in the type of scale used, depending on the availability of data, on the computational resources at hand, and on the level of detail required to obtain a useful calculation. In some studies the geography is defined very simply, on a coarse scale, as irregularly shaped polygonal regions. This often occurs in experimental or process-oriented studies because it is sufficient to the purpose and it is easy to assign a measurement to a crudely defined area. For example, if we are interested in plant communities relative to rainfall patterns on the continent of South America, a series of relatively large polygonal regions (one for the western coastal desert, another for the Andes mountains, another for the Amazon basin, and another for midlatitude regions south of the Amazon basin) may suffice for the purpose. They are relatively simple, yet they capture substantial amounts of variation in habitat. But if we want to know the effects of forest removal on the rate at which carbon is fixed by plants in the Amazon basin, we might need more resolution to capture spatial variability in production versus cutting. For this purpose it is likely that we would construct a Cartesian grid of contiguous blocks, with quantities having geographic attributes of position and resolution on a ratio type of scale.

Geographically explicit techniques have proliferated in the last decades, but it is often difficult to compare results among studies. One way of making sense of the sometimes confusing array of techniques is to examine the way that geographic attributes are

Table 7.1 Comparative Analysis of Geographically Explicit Studies

1. What quantities are being used ?
2. Are the quantities measured or computed ?
3. What type of scale is the quantity on ?
4. What are the temporal attributes of each quantity ?
   - Chronology ?
   - Duration of each measurement ?
   - Time between measurements ?
   - Time from start to end of set of measurements ?
5. What are the spatial attributes of each quantity ?
   - Location ?
   - Area of each measurement?
   - Minimum separation between nearest measurements?
   - Maximum separation between measurements?
6. On what type of scale is each attribute expressed ?

expressed. Table 7.1 lists a sequence of questions for comparative examination and analysis of geographically explicit studies. This list of questions, used as a mental checklist in examining a study, frequently brings out similarities among studies that appear to differ. An example is the relation of spatial autocorrelation (Cliff and Ord, 1974) to metapopulation analysis. The two methods appear to have little in common, yet metapopulation studies often use nearly the same definition of spatial variables as in Cliff and Ord (1974). Consequently, the statistical techniques of Cliff and Ord are potentially applicable in evaluating metapopulation models against data. Of course, applicability also depends on the question at hand.

7.4 Geographic Attributes in Two and Three Dimensions

For objects that move in two or three dimensions, we have two ways of recording position. *LaGrangian data* consist of a sequence of values at points occupied by the object, whether an animal, a seed, or a parcel of water. Examples are data from activity recorders strapped to caribou or to drifters set loose at sea. The devices record quantities such as temperature at a series of points occupied by the drifter, or by the caribou. *Eulerian data* consist of a values at fixed points on a grid, regardless of path. An example is the count of Willets between fixed points along a beach.

There are two conventions for expressing the position and extent of quantities in more than one spatial dimension. One convention uses a Cartesian or square grid; the other uses a polar or circular grid. The latter is sometimes encountered in ecology or related fields of environmental research. In this system $\boldsymbol{r}$ represents a position as a displacement away from a zero point. This is a vector, hence the boldfaced symbol $\boldsymbol{r}$. It is composed of a unit vector (step size in a stated direction) and a scalar (number of steps). The scalar number r is the number of steps, each of unit size, away from the zero point. θ represents the angle to the right (clockwise) of $\boldsymbol{r}$ in a plane. Positions in a volume can also be stated in polar coordinates.

The other convention is the familiar Cartesian system in which x stands for displacement in a fixed direction, y for displacement in a direction 90° to the left of x, and z for displacement in a third direction, perpendicular to both x and y. In geographic applications of this system, z is aligned with the earth's gravitational field, either positive upward from the center of the earth or positive downward from the sea surface, as in oceanography. The x direction usually runs zonally (west to east), leaving y to run meridionally toward one of the poles of rotation (north-south). In this system x can be positive (say, east) or negative (west), y can be positive (say, north) or negative (south), and z can be positive (say, up) or negative (down). A more compact notation compresses all this information into a single symbol, boldface $\boldsymbol{x}$, which designates position in three dimensions:

$$\boldsymbol{x} \equiv [x\boldsymbol{i} + y\boldsymbol{j} + z\boldsymbol{k}]$$

Boldface type is the conventional way of distinguishing a *vector quantity* (which has a direction) from a *scalar quantity* (which has no direction).

This symbolic expression is sufficiently abstract that an interpretation of each component is in order, beginning on the left side of the expression. Imagine that you are responsible for mapping the location of a species of orchid in a rainforest and that you have a well-defined zero point at the center of a nature reserve. For each orchid in the reserve, boldface $\boldsymbol{x}$ designates the position as a displacement in three dimensions from the zero point. The symbol $\boldsymbol{i}$ is called the *unit vector* in the x direction. It is the size of the step used to measure distance from the zero point. A convenient unit is the kilometer. x is the number of steps, each a kilometer in length, that a particular orchid is located to the north ($\boldsymbol{i}$ positive) or to the south ($\boldsymbol{i}$ negative) of the zero point. So, $x\boldsymbol{i}$ is the total displacement of this orchid from the zero point, in the x direction. The same goes for east and west. $\boldsymbol{j}$ is the step size (again, in kilometers), y is the number of steps, and $y\boldsymbol{j}$ is the east-west displacement of this orchid from the zero point, in kilometers. The zero point is on the ground, and the orchid is up in a tree, so a complete statement of position requires both $\boldsymbol{k}$ (the step size in the vertical) and z (the number of steps in the vertical). The product $z\boldsymbol{k}$ is the height of the orchid above the ground, or the vertical displacement from the zero point. The vertical displacement might be anywhere from a fraction of a meter to several hundred meters above the ground.

The three unit vectors $\boldsymbol{i}$, $\boldsymbol{j}$, and $\boldsymbol{k}$ define the spatial resolution of the orchid position $\boldsymbol{x}$, whereas the products $x\boldsymbol{i}$, $y\boldsymbol{j}$, and $z\boldsymbol{k}$ define the range of the orchid position $\boldsymbol{x}$. The unit vectors $\boldsymbol{i}$, $\boldsymbol{j}$, and $\boldsymbol{k}$ are fixed in a particular study, so these are often dropped from the notation:

$$\boldsymbol{x} = [x\ y\ z] \cdot \text{units}$$

If $\boldsymbol{x}$ is measured in kilometers km, then $x\ y\ z$ are the numbers of steps, each a kilometer in size. Position is sometimes represented in still more abbreviated form:

$$\boldsymbol{x} = [x\ y\ z]$$

This notation assumes that $\boldsymbol{x}$ stands for $x\boldsymbol{i}$, and so on. This notation suffices if we are interested in scanning across a sequence of measured values with respect to position in one, two, or three dimensions. The notation is inadequate for scaling functions, which require units.

The vector notation described here leads to two forms of measurement relation. A quantity can be scaled to displacements $x\boldsymbol{i}$, $y\boldsymbol{j}$, and $z\boldsymbol{k}$. Taking just one of these

dimensions, if we have counts of orchids on transects of length ***xi***, we can scale the count to the transect length, in a fashion similar to the Willet example:

$$\frac{N_A}{N_B} = \left(\frac{x_A i}{x_B i}\right)^{\beta} \tag{7.1a}$$

Holding ***i*** constant, we have:

$$\frac{N_A}{N_B} = \left(\frac{x_A}{x_B}\right)^{\beta} \tag{7.1b}$$

This is called *scaling by accumulation* (Figure 7.1).

Alternatively, we can scale the counts to subsections of a single transect by varying the bin size ***ni***, usually within a fixed transect length:

$$\frac{N_A}{N_B} = \left(\frac{x_A}{x_B}\right)^{\beta}\left(\frac{n_A i}{n_B i}\right)^{\beta} \tag{7.1c}$$

This is the basis of *coarse-graining* (Figure 7.1), a form of iterative scaling. Figure 7.1 shows a third form of iterative scaling, called *lagging*. In this form of spatial scaling the quantity of interest is scaled to the separation between two points. All three forms of iterative scaling will be described again in later chapters.

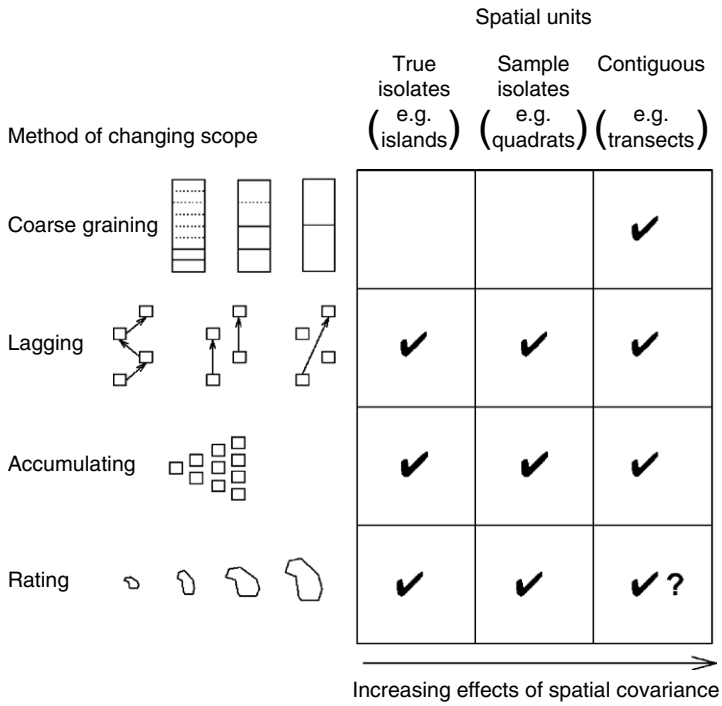


FIGURE 7.1 Comparison of Four Scaling Maneuvers. Iterative scaling is by accumulation, coarse graining, or lagging. Non-iterative scaling is by rating. Check marks signify allowable spatial units, given the method.

ANOTHER LOOK AT SECTION 7.4

For a quantity of interest to you, provide a range of typical values, then provide typical magnitudes for unit vectors ***i***, ***j***, and ***k*** needed to quantify position in Cartesian coordinates.

7.5 Notation

Good notation is a tremendous aid in working with scaled quantities, with the spatial and temporal attributes of quantities, and with derived and ensemble quantities described in the next two chapters. Ecology lacks a clear and consistent notational system for scaled quantities with spatial and temporal attributes. This makes it tempting to create a logical and consistent notation that, once proposed, will, of course, be immediately adopted. The lessons of history are otherwise. New notation, if noticed at all, typically passes into oblivion (Cajori, 1929). Mathematical signs and symbols are invented by individuals, but only two individuals, G. Leibniz and L. Euler, have invented more than two signs or symbols that subsequently passed into common usage. This is largely due to the effort required to learn new notation. Once a notation has been learned (as with the qwerty keyboard), it resists replacement by another (such as the more efficient and comfortable Dvorak keyboard).

A second reason for slow diffusion of new notation is that only experience will tell whether a sign is indispensable. It could take years or decades to find out whether a symbol is worth keeping. Why go to the effort of learning a new notation unless it is already known to aid comprehension or simplify computation?

A third factor that slows diffusion is that some forms of notation are hard to set into type, even though they are clear and easily written by hand. Leibniz realized this and advocated symbols that could be set onto a single line of type, dy/dx, instead of symbols that required two tiers of type, such as $\dot{y}$. Three tiers required even more time to set mechanically because they must be placed on their own line:

$$\frac{dy}{dx}$$

This constraint has now disappeared as the graphics capabilities of computers have made it possible to set type in ways that come close to the ways that symbols are written by hand. Nevertheless, mathematical symbols placed on several tiers require time and effort. For example, the time taken to prepare the first edition of this book for camera-ready copy exceeded the time to write the text due to the extensive integration of signs and symbols into the text. Anything that reduces the time and costs of typesetting will contribute to the adoption of a sign or symbol.

Good notation would seem to require little thought. In fact, there are *principles of good notation* based on historical experience (Cajori, 1929) and familiarity with quantitative treatments of biology (Riggs, 1963). These authors state several criteria of good notation, which are combined in Table 7.2. The first criterion, consistency with current usage, is important because of the perpetual temptation to improve existing notation. The reason for listing this criterion first is that notational improvements do not take

Table 7.2 Criteria in Selecting Mathematical Notation

1. Consistent with current usage
2. Reduces the burden on memory
3. Demonstrated utility in quantitative work
4. Brevity
5. Lends itself to computer applications

Note: Modified from Cajori (1929).

root and flourish until they are adopted by a nonconferring group of specialists (Cajori, 1929). Notation that is consistent with current usage increases the communication of ideas. Novel signs and symbols usually hinder communication. At this point I now have to confess that in this book I have been using a novel symbol—adding an arrow to an equality sign (=>) to mean "calculated from". The modified equality sign shows that the right side was calculated from the left and that this is not equivalence by definition.

The second criterion listed in Table 7.2 is reducing the burden on memory. This is accomplished by several tactics. One is to use as few symbols as possible. Another tactic is to avoid ambiguous signs and symbols, such as St. Andrew's cross (×) for multiplication. This symbol is all too easily confused with the letter x, which conventionally stands for an unknown in algebra, for distance along a line in Cartesian coordinates, or for time since birth in population biology. Still another tactic is to use letters for quantities and signs for operations such as dividing, taking derivatives, or taking logarithms. Diacritical marks, such as dots over a symbol for operations on the quantity, reduce the burden on memory, provided there is some precedent for the mark. Yet another tactic is to connect symbols to the measured quantity rather than to secondary characteristics. For example, $N_{disturb}$ and $N_{control}$ are far easier to use and recall than D and C to represent numbers of species in disturbed and undisturbed plots of grassland.

The third criterion in the table, utility in application, can really only be determined from the experience of groups of people working on similar problems. This can take decades. Thirty years between the proposal and eventual adoption of a sign has not been unusual in the past (Cajori, 1929). User groups associated with computer software packages may change this trend. These groups now test the utility of a sign or notational convention in a few years rather than several decades.

The fourth criterion is brevity, provided this does not increase the burden on memory. Exact definition of a quantity in relation to other quantities requires notation that is sometimes unwieldy. But once a quantity has been defined, briefer notation has several advantages. Brevity speeds the recognition of quantities when they appear in a lengthy expression. Brevity increases the readability of equations by making the relation of one quantity to another more apparent. Several examples of this passage to brevity occur in earlier sections of this chapter.

One of Cajori's criteria is suitability for mechanical typesetting. This was omitted from Table 7.2 because typesetting is now digital, allowing notation that matches the flexibility of hand notation. In place of this criterion I have substituted suitability for making computations with a computer. An example is the use of several signs for the different meanings of "equal". Several different symbols are required to tell a computer how and when to make calculations. These signs distinguish definitions from other forms of equality, such as an instruction to begin calculation. The idea of separating the several

meanings of "equal" comes from a computer package (MathCad) that is widely used to teach mathematics to undergraduates. Much of the notation in this book comes from this package, making it easy to translate into computational code.

Many of the points made by Riggs (1963) about the use of mathematical notation in physiology apply with equal force to ecology. Riggs' first point is that biological systems are complex, hence the inescapable need to make assumptions and devise models, including mathematical expressions. The second point is that the brevity of an expression belies the effort required to devise, understand, or check the expression. The third point is the cardinal need for clear and unambiguous definition of terms and symbols. This point motivated the five-part definition of a quantity in Chapter 3. Riggs' next point is that in principle any symbol can be chosen for a quantity, but in practice much is accomplished by thoughtful choice based on several desiderata: ease of recognition, brevity, and conformity to standard usage. Consistency with current usage appears as the first criterion in Table 7.2 due to the lessons of history (Cajori, 1929). Riggs recommends substituting one symbol for a recurrent group, a device used frequently in this book. Another point made by Riggs is the need for consistency in the use of symbols. This is easily accomplished by ensuring that a symbol keeps the same units and procedural statement in any setting.

Riggs advocates (strongly!) the association of abstract symbols with concrete meanings. Any book in quantitative methods necessarily contains abstract symbols that should not, to quote Riggs, "float about in your mind like featureless wisps of mist above a marsh." They should be thought of as measurable quantities associated with vivid images and specific units. Whenever the generic symbol Q appears further on in this book, it should be replaced in the reader's mind with a familiar quantity to make it concrete and comprehensible.

ANOTHER LOOK AT SECTION 7.5

If you have ever employed mathematical notation in a written report, how did you select the notation? Did you use any of the criteria presented in Table 7.2?

7.5.1 Application: Spatial and Temporal Attributes

This section applies the principles of good notation to identify a set of symbols to represent scaled quantities having spatial and temporal attributes. The aim is not a definitive notation; the aim is to illustrate the process of choosing a notation that balances clarity and internal consistency with traditional use of symbols in ecology. Such notation facilitates quantitative treatment of quantities with geographical and chronological attributes: position, resolution, and extent in space and time.

Table 7.3 lists a set of conventions for expressing the spatial and temporal attributes of quantities. The conventions are based on the principles in Table 7.2, most notably consistency with current practice. To demonstrate the conventions, we apply them to an example of gypsy moth distribution and dynamics. Q is the generic symbol for a scaled quantity, but for the sake of specificity, here Q will temporarily represent the density of gypsy moth *Lymantria dispar* cocoons measured in units of cocoons $\cdot$ m^{-2} of tree bark.

Table 7.3 Conventions Used in Expressing Spatial and Temporal Attributes of Quantities

Q represents a quantity with units. Any symbol can be used in place of Q.
x, y, and z stand for geographic attributes of position and resolution.
t stands for temporal attributes of position and resolution.
Boldface type distinguishes vectors (which have direction) from scalars (which have no direction). To write a vector symbol by hand, we add a squiggle beneath to distinguish scalar quantities x from vector quantities $\underset{\sim}{x}$.
The attributes of measured values of Q are represented by subscripts.
Calculated or expected values of Q are shown as functions of spatial and temporal attributes.

$\boldsymbol{xi}$ represents the position of each measurement of cocoon density along the transect, in a segment of fixed length $\boldsymbol{i}$. By holding $\boldsymbol{i}$ to a fixed value and x to a sequence of integer steps, we allow $\boldsymbol{xi}$ to function as an address for each measurement. $Q_{\boldsymbol{xi}}$ stands for the set of measurements of cocoon density along the transect. This is shortened, in the interests of clarity, to Q_x if we have no plans to undertake rescaling by changing the step size $\boldsymbol{i}$. The symbol $Q_{\boldsymbol{xi}}$, or, equivalently, Q_x, stands as well for a map of the measurements along a transect. A plot of cocoon density Q_x against x shows the one-dimensional distribution of measurements.

If we happen to have a model or formula that lets us calculate Q_{xi} from a knowledge of position $\boldsymbol{xi}$, then a distinct symbol for the calculated values will be needed. An appropriate symbol is $Q(\boldsymbol{xi})$. This represents the "expected value of Q at position $\boldsymbol{xi}$" or "the expected value of cocoon density at location $\boldsymbol{xi}$." This calculated value is based on *expectation from a function*. The symbol $Q(\boldsymbol{xi})$ can be shortened to $Q(x)$ if we have no interest in changing the resolution scale $\boldsymbol{i}$. A plot of $Q(x)$ versus x shows the pattern of distribution expected from the model. Thus, $Q(x)$ will be a simplification of the data, based on a formal model that captures major features of Q_x rather than every detail. $Q(x)$ will only occasionally be exactly equal to Q_x. To evaluate the model against the data, we set the two quantities as equal by adding an error or residual term:

$$Q_x = Q(x) + \textit{residual} \tag{7.2a}$$

This symbolic expression is read as "the measured value of the quantity Q_x is exactly equal to the sum of the expected value of the quantity $Q(x)$ plus some residual variation". In ecology, the parameters of a model such as $Q(x)$ are usually estimated from data. An example is an estimate of an exponent for a scaling relation (Equation 2.5a) or measurement relation (Equation 2.6a) When parameter estimates are used to compute the model values, the result is marked as an estimate by placing a "hat" over the symbol:

$$Q_x = \hat{Q}(x) + \textit{residual} \tag{7.2b}$$

Statistical models use *statistical expectation*, which means that the value of $Q(x)$ arises from measurement many times, not just a few times. The symbol $E[Q]$ represents the expected value of Q. The notation appropriate to statistical evaluation is:

$$Q_x = E[Q(x)] + \textit{residual} \tag{7.2c}$$

where E[$Q(x)$] represents cocoon density calculated from the model. This is the research model, also called the *alternative model* relative to the *null model* H_o of no relation with cocoon density Q_x to distance x. The null model is:

$$Q_x = E[Q_x] + residual \tag{7.2d}$$

$E[Q_x]$ represents the expected or mean value regardless of location x, as obtained by repeated measurement. A well-developed set of statistical methods (e.g., Sokal and Rohlf, 1995) exists to decide whether to accept the research model in preference to the null model. Chapter 15 covers application of these methods to estimating and comparing the parameters of power laws.

If we make measurements on the transect at successive points in time, we have four attributes: spatial resolution $\boldsymbol{i}$, spatial range $\boldsymbol{xi}$, temporal resolution $\boldsymbol{h}$, and temporal range $\boldsymbol{th}$. The symbol $Q_{xi\ th}$ stands for cocoon density, measured at regular intervals along the transect at regular intervals in time. A more compact notation is $Q_{x\ t}$ if spatial and temporal resolution are held constant. The symbols $Q_{xi\ th}$ and $Q_{x\ t}$ stand for the set of measurements of cocoon density collected into a column vector for each point in time. The information represented by these symbols can be expressed as plots of Q_{xi}, for each regularly spaced point in time. If we happen to have a function that permits us to calculate the way in which the map Q_x changes with time, then as before we can compare the model written in functional notation $Q(x\ t)$ to the data represented by $Q_{x\ t}$.

If we have enough resources to measure cocoon density over an area of forest rather than just along a line, the symbol for this set of measurements is $Q_{xi\ yj}$. This becomes $Q_{x\ y}$ if resolution does not vary. These equivalent symbol represent data that can be plotted against position x and y within the area of interest. $Q(x\ y)$ represents the value expected from the model that relates cocoon density to position x and y. The model could be in the form of an equation. It might also be a contour map estimated from previous data. Such a map is an empirical model of the cocoon density. The measured value $Q_{x\ y}$ will differ from the model value $Q(x\ y)$ by some residual amount.

$$Q_{x\,y} = Q(x\,y) + residual \tag{7.3a}$$

As with the one-dimensional case, the null model of no relation of cocoon density to location is:

$$Q_{x\,y} = E[Q_{x\,y}] + residual \tag{7.3b}$$

The expression for a quantity as a function of location is $Q(\boldsymbol{xi}\ \boldsymbol{yj})$. This is called a *scalar field*, which is defined as a scalar quantity that is a function of a vector. In this case the vector is a position, and the scalar field is two dimensional. The calculation of distribution of some variable of interest (quantified as a scalar field) from knowledge of physical, chemical, and biological processes is one of the central challenges in environmental science. Examples of such calculations include forecasting the weather, projecting rises in sea level due to global warming, and evaluating the effects of climate change on forest and crop production.

If we have measurements of a quantity throughout an area at several points in time, we have six attributes, which I ask the reader to name and visualize at this point:

Symbol	Name, in Words
x	
$\boldsymbol{i}$	
y	
$\boldsymbol{j}$	
t	
$\boldsymbol{h}$	

Finally, the height of cocoons off the ground within the area of interest might be of interest. Cocoon mortality, for example, might depend on height. Now we have six spatial attributes, which is going to be cumbersome. So, the symbol $Q_{xi\ yj\ zk}$ will be shortened to $Q_{\boldsymbol{x}}$ to stand for the set of measurements of the quantity, cocoon density, with attributes of position $\boldsymbol{x}$ in a three-dimensional volume. The amount of information that we now have makes it a challenge to show $Q_{\boldsymbol{x}}$ as a map. One way to make such a map is to show position $\boldsymbol{x}$ as a three-dimensional perspective drawing on a two-dimensional computer screen and use color to designate the quantity, shading from blue at low values to red at high values. In addition to the data $\boldsymbol{Q_x}$ we might also have expected values, from either a dynamic model or a static model such as a contour plot through the data. $Q(\boldsymbol{x})$ stands for calculated values of Q at position $\boldsymbol{x}$. Calculated or expected values typically show a simplification of the data. They capture the major features of spatial variation in the quantity, rather than every detail. Values calculated from $Q(\boldsymbol{x})$ typically show a smoothed or cartoon representation of the quantity. $Q(\boldsymbol{x})$ is called a scalar field in three dimensions.

Perhaps we have available repeated measurements of the quantity in three dimensions. This brings us up to eight attributes. The symbol $Q_{\boldsymbol{x}\ t}$ stands for a measured quantity with complete specification of position in three-dimensional space through time. To draw this we need a series of pictures, which we could display in sequence on a page of paper or, better yet, on a computer screen. A visually effective format (Tufte, 1990) would be a three-dimensional projection onto the flat surface of a computer screen, where we can watch the change in colors, perhaps with blue representing low density, shading to red for high density of the quantity Q = density of gypsy moth cocoons. If we have a way of calculating cocoon density at any position at any time, the symbol for these values is $Q(\boldsymbol{x}\ t)$. This is a sequence of cartoon representations of the quantity that appear before us on the computer screen as a series of snapshots or as an animated sequence.

The aim of this excursion into the unfamiliar realm of mathematical notation was to return with a clear and consistent set of symbols for quantities with spatial and temporal attributes. A good set of symbols is as important as the familiar modes of verbal expression, if we want to measure or calculate these quantities. The symbol $Q_{\boldsymbol{x}\ t}$ stands for any geographically and temporally referenced quantity with temporal and spatial attributes of resolution (vectors $\boldsymbol{h}$, $\boldsymbol{i}$, $\boldsymbol{j}$, and $\boldsymbol{k}$) and position in time (th) and space ($\boldsymbol{x}i$, yj, and $z\boldsymbol{k}$). The compact symbol $Q_{\boldsymbol{x}\ t}$ demonstrates the capacity of mathematical

notation to represent complex concepts (requiring pages of text to explain). Symbols for spatial and temporal attributes are like actors that appear on stage, acquire meaning through explanation and context, then at each reappearance bring to the stage context and meaning. Symbols can represent not just a collection of numbers but a complex and valuable notion: that any quantity that we care to measure has attributes that set its spatial and temporal scales.

The symbol $Q_{x\,t}$ is so compact that it is inscrutable on first encounter. It needs a chance to state its role. Once the symbol's context and meaning are grasped, the symbol serves us well because it is so compact. With this symbol and its associated notation it becomes possible to make calculations based on ideas about the spatial and temporal scale of quantities.

7.6 Spatial and Temporal Scaling

When a quantity is coordinated to spatial or temporal attributes, we can scale it to time and space by altering the scope in any of several ways (Figure 7.1) that are consistent with the definition of scaling in Box 6.1. We can alter the scope by rating multiple units with respect to size or duration, as when we rate lakes by size and then scale fish catch to lake size. Noniterative scaling relations (Equation 2.5a) and noniterative measurement relations (Equation 2.6a) are usually based on this maneuver. We can alter the scope by *accumulating* units (Figure 7.1), as when we examine reproductive contribution with increasing age of a cohort. This maneuver is useful with data from *similarly* sized plots, but it's rarely used. We can alter the resolution, and hence the scope, by altering the separation (Figure 7.1) between units. This maneuver, called *lagging*, is accomplished with an iterative procedure, whereby a single set of measurements is used repeatedly. The units can be sequential in space or time, but they do not have to be. We can alter the scope by changing the measurement frequency (Figure 7.1), as when we convert from a series of daily counts to weekly or monthly counts. Technically speaking, we alter the resolution by altering the spatial and temporal extent of the unit (altering vectors $\boldsymbol{h}$, $\boldsymbol{i}$, $\boldsymbol{j}$, and $\boldsymbol{k}$). This is called *coarse graining* (refer to Chapter 10). This scaling maneuver is accomplished iteratively and so it requires units that are evenly placed in time or space, preferably by being contiguous (see Figure 7.1). If we have an exhaustively measured quantity, as in a time series with no gaps or in a satellite image with no missing data, then in principle lagging and coarse graining generate equivalent information. In practice this equivalence is not evident and conversion is often difficult to achieve. Further, there are strong historical precedents for the use of one maneuver within any given discipline. For example, oceanographers are often familiar with coarse graining, whereas terrestrial ecologists rely almost entirely on lagging to quantify scale-dependent patterns. Scaling by coarse graining, lagging, and accumulation are treated in more detail in Chapter 10.

Finally, we can alter the scope by *rating* (Figure 7.1), defined as comparing measurements made on distinct units that differ in size or some other characteristic (e.g., mass). An example of scaling by rating is change in size with change in time, as in the example used earlier:

$$M_t = [15.0_{\text{day 2}}\ 18.1_{\text{day 6}}\ 18.1_{\text{day 7}}\ 20.2_{\text{day 14}}\ 20.1_{\text{day 14}}] \cdot \text{grams}$$

The measurement relation (Box 6.1) is:

$$\frac{M_{i+\Delta i}}{M_i} = \left(\frac{t_{i+\Delta i}}{t_i}\right)^{\beta} \tag{7.4a}$$

Relative to the values at $i = 0$ ($t_i = 2$ days), there are four values of β, of which one is:

$$\frac{18.1 \text{ grams}}{15.0 \text{ grams}} = \left(\frac{6 \text{ days}}{2 \text{ days}}\right)^{0.105} \tag{7.4b}$$

The other three values of β are similar. Rearranging 7.1a, we have:

$$M_{i+\Delta i} = \left(M_i^{-1} t_i^{\beta}\right)\left(t_{i+\Delta i}\right)^{\beta} \tag{7.4c}$$

for which the estimate via regression (see Chapter 15) of β is 0.137.

Measurement relations such as Equation 7.4a require attributes on a ratio type of scale. Many temporal attributes are recorded on an interval type of scale (time of day, day of the year). However, the interval-scale quantity of time recorded from a clock can be rescaled to a ratio scale relative to a zero point, such as the beginning of a period of growth, beginning of a survey, and so on.

Another example of scaling by rating is that of Willet counts N_i, which are readily scaled to segment length L_i according to a measurement relation.

$$\frac{N_A}{N_B} = \left(\frac{L_A}{L_B}\right)^{\beta} \tag{7.5a}$$

Intuitively the scaling exponent is $\beta = 1$, which means that a doubling in area will double the number of Willets, trebling the area will treble the number of Willets, and so on. Though this is true of averages taken over many cases, it is not the case for a single set of measurements. For the Willet data, the scaling of numbers from the largest and smallest segments yields an exponent greater than 1:

$$\frac{7.4 \text{ Willets}}{2.1 \text{ Willets}} = \left(\frac{6.4 \text{ km}}{1.9 \text{ km}}\right)^{2.03} \tag{7.5b}$$

A better estimate can be obtained by regression. The scaling exponent via regression is $\beta = 1.45$. With only five counts, little can be said about whether this unexpected relation is due to chance. When more counts are analyzed, the relation often (but not always) turns out to be stronger than expected by chance. Why should numbers of Willets more often than not increase disproportionately with increase in unit area? The answer lies in the highly clumped distribution of Willets. The chance of encountering such a group increases with area searched. Consequently, any particular sequence of counts of Willets will, more often than not, show an increase in number with increase in area. Of course, this will not be the case over the long run because from time to time a large number of Willets will be found in a small segment of beach.

Defined Concepts and Terms for Review and Future Reference

____ functional vs. statistical expectation
____ geographic attributes of a quantity
____ LaGrangian vs. Eulerian data
____ null versus alternative model
____ principles of good notation
____ spatial resolution (grain)
____ scalar vs. vector quantity
____ scalar field
____ scaling by accumulation
____ scaling by coarse-graining
____ scaling by lagging
____ scaling by rating
____ temporal attributes of a quantity
____ unit vector

8 Quantities Derived from Sequential Measurements

Each point on a great ice body has its own numerical value for mass balance. Is the ice right here thicker or thinner than last year? Is the glacier, at this spot, thriving or dying? The collective profile of all those individual soundings—more ice or less? thriving or dying?—is called the gradient of net mass balance. This gradient tells, in broad perspective, what has been lost and what has been gained.

—David Quammen, *Strawberries Under Ice,* 1988

8.1 Synopsis

This chapter describes a series of ecologically important quantities calculated from sequential measurements. These derived quantities include the time rate of change $\dot{Q}$ the time rate of change as a percentage $Q^{-1}\dot{Q}$, the flux $[Q]\dot{x}$, the spatial gradient ∇Q, the spatial gradient as a percentage $Q^{-1}\nabla Q$, and the divergence $\nabla \cdot Q$.

All these quantities occur in the ecological literature, although not always under these names. An example is the spatial gradient in prey. Studies of foraging behavior often revolve on gradients in prey numbers or gradients in energy value.

The derived quantities in this chapter can all be calculated at several time and space scales. In this they differ from directly measured quantities, which are obtained at a particular resolution or frequency of measurement. For want of a better term, I have called a collection of derived quantities calculated at more than one scale a matrix of contrasts. The rows in this matrix show contrasts at a single scale. Columns show contrasts at a range of scales. Comparison across a row corresponds to scanning a quantity. Comparison upward in a column corresponds to zooming in on detail. Important clues about the dynamics of a quantity arise from adopting the roving viewpoint (scanning) combined with sequential changes in the scale of attention (zooming).

8.2 Time Rates of Change

Ecology, in the broad sense of organisms interacting with one another and with their environment, is about change. How fast does the mass of an organism change as it grows? How fast does total population size N change? Or how fast does local density $[N] = N/A$ change? These fundamental questions are addressed by measurement of the *time rate of change* in the quantity Q. (I hope readers will substitute the name of their favorite quantity for the symbol Q.) The symbol for the time rate of change in Q is $\dot{Q}$, which

doi: 10.1016/B978-0-12-627865-1.00008-8

is read as "the time rate of change in Q." The dot notation was invented by Newton, who placed it over the symbol for a quantity to represent the rate of change in that quantity with respect to another quantity. Newton's dot notation resisted replacement by Leibniz's notation d/dt until the early 19th century. The dot notation has now disappeared from mathematics because it proved cumbersome and unusable for third and fourth derivatives. However, the dot notation has persisted in the natural sciences as the symbol for the time rate of change. In physiology, for example, the volume of oxygen V_{oxy} absorbed by the lungs per unit time is $\dot{V}_{oxy}$. The dot stands for the time rate of change in a quantity, not for the operator d/dx, the derivative with respect to some variable x.

The time rate of change is calculated either from successive measurements or from a theoretical rule. Here is the time rate of change calculated from two measurements of a particular quantity, N_t = ant numbers:

$$\frac{N_2 - N_1}{t_2 - t_1} = \frac{50 \text{ ants} - 100 \text{ ants}}{5 \text{ days} - 3 \text{ days}} \Rightarrow -25 \quad \frac{\text{ants}}{\text{day}} \tag{8.1a}$$

This notation is accurate but unwieldy. More concise notation will be useful:

$$\dot{N}_t \equiv \frac{\Delta N}{\Delta t} = \frac{N_i - N_{i-1}}{t_i - t_{i-1}} \tag{8.1b}$$

The symbol $\dot{N}_t$ represents the measured value of change ΔN scaled to change in time Δt. The attribute t has been expressed as a quantity, which is coordinated with the quantity N by the unitless index i, as in Chapter 7. A sequence of three measured values of the quantity $\dot{N}_t$ looks like this:

$$\dot{N}_t = [-25 \ -20 \ -18] \cdot \text{ants/day}$$

The same notation applies to any quantity. Another example is the area occupied by a colony of ants at a series of points in time, for which an appropriate symbol is A_t with units of m^2. Placing a dot over this symbol generates a new symbol $\dot{A}_t$ for the measured rate of expansion or contraction in the area occupied by the colony.

The generic symbol for the measured time rate of change in any quantity Q is:

$$\dot{Q}_t \equiv \frac{\Delta Q}{\Delta t} \tag{8.2}$$

It will be useful to distinguish this from the time rates of change in a quantity calculated from a mathematical function that expresses an idea about dynamics. This activity will be described in more detail after the concept of a functional relation has been developed in Chapter 13. The symbol for the time rate of change calculated from a function $Q(t)$ according to some rule is:

$$\dot{Q}(t) \equiv \frac{dQ}{dt} \tag{8.3}$$

The symbol dt is the instantaneous change in time. The instantaneous time rate of change dQ/dt is calculated at any value of t, in contrast to $\Delta Q/\Delta t$, which is calculated only

Table 8.1 Notation Distinguishing Derived Quantities Based on Measurement from Those Based on Functional Expression

Time rate of change:	
$\dot{Q}$	From a rule or from measurements
$\dot{Q}(t)$	Calculated from a function
$\dot{Q}_t$	Calculated from measurements only
Fluxes:	
$A^{-1}\dot{Q}$	From a rule or from measurements
$A^{-1}\dot{Q}(t)$	Calculated from a function
$A^{-1}\dot{Q}_x$	Calculated from measurements only
Gradients:	
∇Q	From a rule or from measurements
$\nabla Q(\boldsymbol{x})$	Calculated from a rule
$\nabla Q_{\boldsymbol{x}}$	Calculated from measurements only

This notation applies to one dimension (x), two dimensions (x and y), or three dimensions ($\boldsymbol{x} = [x\ y\ z]$).

at the times when measurements were made. Table 8.1 shows notation that distinguishes between rates calculated from measurements and rates calculated from functions.

The symbol t in Table 8.1 has been used here in a way that is standard in the literature, not distinguishing step number t from position in time $t\boldsymbol{h}$. The product of step size $\boldsymbol{h}$ and number of steps t expresses the position in time as a displacement from a zero point. The product $t\boldsymbol{h}$ has units of time, whereas t is a number without units. As typically used, the quantity $\dot{Q}_t$ assumes a fixed time step such as $\boldsymbol{h} = 1$ day. When step size $\boldsymbol{h}$ is constant, t and $t\boldsymbol{h}$ are used interchangeably.

As a matter of completeness, the symbol for the time rate of change in a measured quantity $Q_{t\boldsymbol{h}}$ with attributes of resolution $\boldsymbol{h}$ and range $t\boldsymbol{h}$ is:

$$\dot{Q}_{t\boldsymbol{h}} \equiv \frac{\Delta Q}{\Delta t\boldsymbol{h}} \tag{8.4}$$

$\dot{Q}_{t\boldsymbol{h}}$ represents the time rate of change in a quantity derived from two successive measurements. When the step size $\boldsymbol{h}$ is fixed, the symbol $\dot{Q}_t$ will serve just as well.

The notation in Table 8.1 applies to any quantity. For example, earthworms bring volumes of soil ($V = \text{cm}^3$) to the surface at a daily rate of $\dot{V}$ in units of $\text{cm}^3\ \text{day}^{-1}$. The symbol for the measured rate is $\dot{V}_t$. Another example is a velocity $\dot{x}$, which represents the time rate of change in position x. Measured velocity $\dot{x}_t$ may well differ from the velocity $\dot{x}(t)$ calculated from a functional expression describing motion.

ANOTHER LOOK AT SECTION 8.2

Examine the definitions of the following quantities:

$$\dot{N}_t \qquad \dot{Q} \qquad \dot{Q}(\boldsymbol{t}) \qquad \dot{Q}_{t\boldsymbol{h}}$$

Which have definable time scales?

8.2.1 Application: Crude Rate of Change in Population Size

Population biologists and demographers sometimes employ *crude rates of change* in population numbers. For example, if a cohort of 1000 adult corn earworm moths, *Heliothis armigera*, eventually produces 1500 adults before dying, the crude rate of change over the generation time of one year is 500 moths yr^{-1}. The number of eggs produced by a cohort greatly exceeds the cohort size; mortality of eggs and caterpillars then reduces this number. Thus it is of interest to partition the net rate into components of recruitment and mortality. Box 8.1 shows the calculation of crude rate of change in numbers, the crude rate of recruitment, and the crude mortality rate for the example of corn earworm moths. The change in time is noted explicitly in these examples by an arrow in the subscript. $\Delta t_{1\to 2}$ signifies the change in time from observation 1 to observation 2. Normally a much terser notation appears: Δt_1 for the change in time beginning at the observation labeled 1.

Box 8.1 Calculation of Crude Rate of Recruitment, Mortality, and Change in Numbers

A cohort of 1000 moths produces 200,000 eggs, of which 1500 survive to form the next generation.

$$\begin{array}{ccccc} 1000 \text{ moths} & \longrightarrow & 200{,}000 \text{ eggs} & \longrightarrow & 1500 \text{ moths} \\ \text{time} = 0 \text{ days} & & \text{time} = 20 \text{ days} & & \text{time} = 365 \text{ days} \\ & \Delta t_{1\to 2} & & \Delta t_{2\to 3} & \end{array}$$

$$\dot{B}_t \equiv \frac{\Delta N_{1\to 2}}{\Delta t_{1\to 2}} = \frac{(200{,}000 - 1000) \text{ moths}}{20 \text{ days}} \Rightarrow 9950 \quad \frac{\text{moths}}{\text{day}}$$

$$\dot{D}_t \equiv \frac{\Delta N_{2\to 3}}{\Delta t_{2\to 3}} = \frac{(1500 - 200{,}000) \text{ moths}}{345 \text{ days}} \Rightarrow -575 \quad \frac{\text{moths}}{\text{day}}$$

$$\dot{N}_t \equiv \frac{\Delta N_{1\to 3}}{\Delta t_{1\to 3}} = \frac{(1500 - 1000) \text{ moths}}{365 \text{ days}} \Rightarrow 1.37 \quad \frac{\text{moths}}{\text{day}}$$

The relation of change in numbers $\dot{N}_t$ to mortality $\dot{D}_t$ and recruitment $\dot{B}_t$ is:

$$\dot{N}_t = \dot{B}_t \frac{\Delta t_{1\to 2}}{\Delta t_{1\to 3}} + \dot{D}_t \frac{\Delta t_{2\to 3}}{\Delta t_{1\to 3}} \tag{8.5}$$

In words, this equation says that the crude rate of change in numbers is due to the crude rate of recruitment adjusted for time, plus the crude rate of mortality adjusted similarly for time. The dot notation lends itself to translating between verbal and mathematical expression of this relation.

ANOTHER LOOK AT SECTION 8.2.1

Rewrite Expression 8.5 using the index i rather than its values of 1, 2, and 3 in this example.

8.3 Time Rate of Change as a Percentage

The time rate of change $\dot{Q}$ often depends in some fashion on the magnitude of the quantity Q. An example is absolute growth rate, which decreases as animals become larger. Another example is addition of new organisms to a population. Eighty caribou are expected to produce more calves than 20 caribou. When the time rate of change in a quantity depends on the magnitude of that quantity, the time rate of change as a percentage becomes of interest. Returning to the example of change in ant numbers, the *percent rate of change* calculated from two measurements is:

$$\frac{1}{N_1} \cdot \frac{N_2 - N_1}{t_2 - t_1}$$

$$\frac{1}{100 \text{ ants}} \cdot \frac{50 \text{ ants} - 100 \text{ ants}}{5 \text{ days} - 3 \text{ days}} \Rightarrow -25 \quad \frac{\%}{\text{day}} \tag{8.6a}$$

This is the relative rather than absolute or crude rate of change. A briefer notation for the relative or percentage rate of change (Eq. 8.6a), calculated from measurements, is:

$$\frac{1}{N}\frac{\Delta N}{\Delta t}$$

The change in number has been scaled to discrete time. A still terser notation, which puts emphasis on the symbol for the quantity, is:

$$\dot{N}_{th}/N \equiv \frac{1}{N}\frac{\Delta N}{\Delta t} \tag{8.6b}$$

The symbol $\dot{N}_{th}/N$ is compact and easy to interpret, making it useful when the quantity appears repeatedly. The subscript th has been added to signify that ΔN is scaled to discrete time Δt with time step h. This contrasts with the instantaneous rate $\dot{N}(t)$, which has no time step. The notation for the instantaneous rate of change, as a percentage, is:

$$\dot{N}(t)/N \equiv \frac{1}{N}\frac{dN}{dt} \tag{8.6c}$$

Percentage rates of change are ubiquitous in population biology, which usually works with percentage recruitment and mortality rates rather than with the crude rates of mortality $\dot{D}_t$, recruitment $\dot{B}_t$, and total change in numbers $\dot{N}_t$. The relative or per capita rates of mortality, recruitment, and change in numbers have units of % per unit time. Box 8.2 shows the calculation of per capita rate of change at two temporal scales.

Box 8.2 Calculation of Per Capita Rate of Change at Two Different Temporal Scales

A cohort of 1000 moths produces 200,000 eggs, of which 1500 survive to form the next generation.

Time step is $\boldsymbol{h} = 1$ day:

$$\dot{N}_{th}/N \equiv \frac{1}{N_1}\frac{\Delta N}{\Delta t} = \frac{(1500-1000)\text{ moths}\dagger}{1000\,\text{moths}\cdot 365\text{ days}} \Rightarrow 0.14 \quad \frac{\%}{\text{day}}$$

Time step is $\boldsymbol{h} = 1$ generation:

$$\dot{N}_{th}/N \equiv \frac{1}{N_1}\frac{\Delta N}{\Delta t} = \frac{(1500-1000)\text{ moths}\dagger}{1000\,\text{moths}\cdot 1\text{ generation}} \Rightarrow 50 \quad \frac{\%}{\text{generation}}$$

The temporal resolution $\boldsymbol{h}$ is used to rescale the rate:

$$\frac{\dot{N}_{th1}/N}{\dot{N}_{th2}/N} = \left(\frac{\boldsymbol{h}1}{\boldsymbol{h}2}\right)^1$$

$$\frac{50\%}{\text{generation}}\cdot\frac{1\text{ generation}}{365\text{ days}} = 0.137\%/\text{day}$$

Box 8.3 Calculation of Per Capita Rate of Recruitment, Mortality, and Change in Numbers

A cohort of 1000 moths produces 200,000 eggs, of which 1500 survive to form the next generation.
Time step is $\boldsymbol{h} = 1$ day:

$$\begin{array}{ccccc} 1000\text{ moths} & \longrightarrow & 200{,}000\text{ eggs} & \longrightarrow & 1500\text{ moths} \\ \text{time} = 0\text{ days} & & \text{time} = 20\text{ days} & & \text{time} = 365\text{ days} \\ & \Delta t_{1\to 2} & & \Delta t_{2\to 3} & \end{array}$$

$$\dot{N}_{th}/N = \frac{1}{N_1}\frac{\Delta N_{1\to 3}}{\Delta t_{1\to 3}} = \frac{(1500-1000)\text{moths}}{1000\,\text{moths}\cdot 365\text{ days}} \Rightarrow 0.137\frac{\%}{\text{day}}$$

$$\dot{B}_{th}/N = \frac{1}{N_1}\frac{\Delta N_{1\to 2}}{\Delta t_{1\to 2}} = \frac{(200{,}000-1000)\text{moths}}{1000\,\text{moths}\cdot 20\text{ days}} \Rightarrow 995\frac{\%}{\text{day}}$$

$$\dot{D}_{th}/N = \frac{1}{N_2}\frac{\Delta N_{2\to 3}}{\Delta t_{2\to 3}} = \frac{(1500-200{,}000)\text{moths}}{200{,}000\,\text{moths}\cdot 365\text{ days}} \Rightarrow -0.29\frac{\%}{\text{day}}$$

It is of interest to partition the per capita rate into components of recruitment and mortality, as it was for the crude rate (Box 8.1). In discrete time, these components can act in sequence rather than simultaneously. This results in interesting dynamics because the net change depends on the time scale. The computations are shown in Box 8.3. Per capita rates of change in discrete time (Box 8.3) are not as readily rescaled as crude rates of change in discrete time (Box 8.1). This is because the basis for computing the per capita rate (i.e., division by numbers present at the beginning of the discrete time period) changes from period to period. In Box 8.3, the basis changed from 1000 moths at the beginning of Period 1 to 200,000 moth eggs at the beginning of Period 2.

A widely adopted solution to this inconvenience is to estimate *instantaneous rates of change* for which the rate is independent of h, the time step or temporal resolution. This is reflected in the notation, which uses instantaneous time t rather than discrete time th. The computation is similar, except that the difference between two counts is replaced by the natural log of the ratio of the two counts. Box 8.4 shows estimation of instantaneous rates for the moth example.

Box 8.4 Estimation of Instantaneous Per Capita Rate of Change in Numbers, Recruitment, and Mortality from Successive Points in Time

A cohort of 1000 moths produces 200,000 eggs, of which 1500 survive to form the next generation.

$$\begin{array}{ccccc} 1000\,\text{moths} & \longrightarrow & 200{,}000\,\text{eggs} & \longrightarrow & 1500\,\text{moths} \\ \text{time} = 0\ \text{days} & & \text{time} = 20\ \text{days} & & \text{time} = 365\ \text{days} \\ & \Delta t_{1\to 2} & & \Delta t_{2\to 3} & \end{array}$$

$$\dot{B}(t)/N \equiv \frac{\ln(N_2/N_1)}{\Delta t_{1\to 2}} = \frac{\ln(200{,}000\,\text{moths}/1000\,\text{moths})}{20\ \text{days}} \Rightarrow 26.5 \quad \frac{\%}{\text{day}}$$

$$\dot{D}(t)/N \equiv \frac{\ln(N_3/N_2)}{\Delta t_{2\to 3}} = \frac{\ln(1500\,\text{moths}/200{,}000\,\text{moths})}{345\ \text{days}} \Rightarrow -1.42 \quad \frac{\%}{\text{day}}$$

$$\dot{N}(t)/N \equiv \frac{\ln(N_3/N_1)}{\Delta t_{1\to 3}} = \frac{\ln(1500\,\text{moths}/1000\,\text{moths})}{365\ \text{days}} \Rightarrow 0.111 \quad \frac{\%}{\text{day}}$$

Time t is used to rescale the per capita rate of change:

$$\frac{\dot{N}(1)/N}{\dot{N}(2)/N} = \left(\frac{t1}{t2}\right)^1$$

$$\left(\frac{0.111\%}{\text{day}}\right)\left(\frac{365\text{days}}{1\text{year}}\right) = \left(\frac{40.5\%}{\text{year}}\right)$$

Instantaneous rates allow the total reproductive contribution to be partitioned into its components of recruitment and mortality in much the same way that crude rates could be partitioned (Box 8.1). Box 8.5 shows the relation between the three per capita rates using the estimates from Box 8.4. The dot notation displays the derived quantities in a fashion that lends itself to translation of the equation into words. It aids in visualizing the relation between the per capita rates.

Box 8.5 Instantaneous Rate of Change in Population Numbers, Partitioned into Recruitment and Loss

Proportional Change in Numbers	=	Proportional Increase Due to Recruitment	+	Proportional Loss Due to Mortality
$\frac{\dot{N}(t)}{N} \cdot \Delta t$	=	$\frac{\dot{B}(t)}{N} \cdot \Delta t$	+	$\frac{\dot{D}(t)}{N} \cdot \Delta t$
$\frac{0.11\%}{\text{day}} \cdot 365\text{days}$	=	$\frac{26.5\%}{\text{day}} \cdot 20\text{days}$	+	$\frac{-1.42\%}{\text{day}} \cdot 345\text{days}$
40.5%	=	529.8%	+	−489.3%

Because of this convenient property—that the product of per capita recruitment and mortality yield the net change—instantaneous rates are regularly used in theoretical and applied population biology. Because recruitment and loss can be partitioned, these rates are taken as independent (e.g., Ricker, 1958). However, such mathematical beauty can be fatal. Instantaneous rates estimated from discrete points in time apply only to the time interval over which they are calculated. Once expressed as instantaneous rates, these estimates are all too easily applied to inappropriate time scales. And as we will see later, interesting dynamics are lost when rates are taken as instantaneous when in fact they act episodically.

ANOTHER LOOK AT SECTION 8.3.1

Have you ever encountered or used discrete time rates of change? Do you think your experience is typical of other biologists or ecologists?

8.4 Fluxes

In a loose or casual sense, the term *flux* refers to the passage of material from one place or compartment to another. For example, energy flow from a prey to predator population is called a *flux* in this casual sense. In its strict sense, a flux is the rate at which a quantity

passes at right angles through a surface, either imaginary or real. Almost any ecological process can be cast as a flux:

- The flux of light downward to the level of the forest canopy
- The flux of rain to the ground
- The flux of nutrients upward into sunlit waters
- The flux of seed propagules onto a cleared field
- The flux of prey entrapped in the mesh of a spider's web
- The flux of fish entrapped in a gill net
- The flux of energy to seabirds bringing prey up through the sea surface
- The flux of animal or human blood across the skin surface to a mosquito

As a generalization, the relation of any population to its resources can be described as a *flux* in the strict sense of exchange per unit time across a defined area. The area might be a flat surface, such as the flux of prey upward through the sea surface to birds. The area might be convoluted, such as the rock surface area scraped by limpets. Or the area might be limited to small segments, such as the skin area through which blood passes to a mosquito or the mouth area through which prey passes to a predator.

One symbol for a flux is the rate of change in a quantity per unit area $\dot{Q}/A$. As with the symbol for time rate of change, the symbol for a measured flux is usefully distinguished from the symbol for a flux calculated from a functional expression (Table 8.1). This notation omits one of the key characteristics of a flux, which is that it is a directed quantity oriented at right angles to a surface A. A flux must have a vector pointing at a plane. The flux of rain, for example, is understood to be downward (in the z direction); it has unit vector $\boldsymbol{k}$ at right angles to the x-y plane. One way to keep track of the unit vector and the plane through which it is directed is to use vector notation for position, as shown in Chapter 7. Vector notation for a flux contains the symbol for a velocity $\dot{\boldsymbol{x}}$, which is a directed quantity. The vector notation is $[Q]\dot{\boldsymbol{x}}$, the product of velocity $\dot{\boldsymbol{x}}$ and a concentration $[Q] = QV^{-1}$. This notation is less intuitive than the previous symbol, but it has the advantage of displaying the direction of the flux in all three directions, x, y, and z.

An example is the flux of plant seeds onto a recently burned area. The number of seeds that land in the burned area per unit of time is a flux oriented downward at right angles to the area. Of course the seeds are also moving laterally into and out of the burned area. And so there are lateral fluxes of propagules $[N]\dot{x}$ and $[N]\dot{y}$ as well as a vertical flux $[N]\dot{z}$. But from the point of view of colonization, the flux that is of interest is $[N]\dot{z}$, the number that eventually settle and come to rest in the area. The flux has units of propagule concentration times velocity. This flux can be visualized as seeds per unit volume, fluxing downward with velocity $\dot{z}$. This is equivalent to the rate at which propagules land per unit area.

$$[N]\dot{z} = \frac{\text{propagules}}{\text{volume}} = \frac{\text{distance}}{\text{time}} = \frac{\text{propagules}}{\text{area} \cdot \text{time}} \tag{8.7}$$

Another example is the upward flux of food energy through the sea surface to marine birds. The upward flux of energy (E = kilojoules) to a group of 10 wandering albatross, *Diomedea exulans,* that capture and consume 400 kJ of squid in a day over an area of 10 km · 10 km is:

$$[E]\dot{z} = \frac{1}{10\ \text{km} \cdot 10\ \text{km}} \cdot \frac{400\ \text{kJ}}{1\ \text{day}} \Rightarrow \frac{\text{kJ}}{\text{km}^2\text{-day}} \tag{8.8}$$

The vector component of a flux can be $\dot{x}$ in the x direction, $\dot{y}$ in the y direction, or $\dot{z}$ in the z direction. The overall velocity $\dot{\boldsymbol{x}}$ consists of three components:

$$\dot{\boldsymbol{x}} = [\dot{x}\ \dot{y}\ \dot{z}] \tag{8.9}$$

The use of boldface type in the symbol $[Q]\dot{\boldsymbol{x}}$ distinguishes flux relative to three coordinates from flux relative to a single coordinate shown in plain typeface $[Q]\dot{x}$.

The most efficient notation for a flux is tensor notation, which appears occasionally in meteorology and oceanography. Tensors have the advantage of keeping track of quantities in four coordinates (x, y, z, and t) in highly compact form. The disadvantage is that tensors are even more unfamiliar and abstract than vector notation for position. Tensor notation and the associated concepts of contraction and transformation may eventually prove useful in geographically explicit analyses of ecological problems. But for now we will stick with vector (directed) and scalar quantities, which are at the edge of notational sophistication for ecology.

8.4.1 Application: The Lateral Flux of Genes

A population geneticist might be interested in the total number of genes (G = number of existing copies of a particular gene) and the rate at which this number changes ($\dot{G}$ = change in number per unit time). But in fact, population geneticists are usually more interested in the number of gene copies relative to the total number of individuals that can carry a copy. If this ratio G/N is equal to 1, there is no genetic variation at the level of the gene within the chromosome. The gene is said to be *fixed*, and there is no opportunity for *evolution*, which is defined as change in gene frequency.

The ratio of the number of gene copies to the number of available locations is the gene frequency:

$$q \equiv G \cdot N^{-1} \tag{8.10}$$

In the strictly technical sense of the word, evolution is $\dot{q}$, the time rate of change in this ratio. The measured value of change in gene frequency over a period of time Δt is:

$$\dot{q}_t \equiv \frac{\Delta(G \cdot N^{-1})}{\Delta t} \tag{8.11}$$

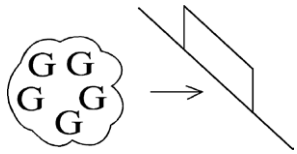
The quantity $\dot{q}$ can also be calculated from a functional expression. A convenient symbol for gene frequency calculated from a function, usually based on some theory, is $\dot{q}(t)$. This stands for the instantaneous rate of change in gene frequency:

$$\dot{q}(t) \equiv \frac{d(G \cdot N^{-1})}{dt} \tag{8.12}$$

Several processes alter gene frequency in natural populations. One of these processes is *migration*, the flux of genes from one population to another. Other processes that alter gene frequency occur within a population: mutation, selection, drift. So, here is the quantity $\dot{q}$ broken down into a flux (migration) and the remaining *in situ* rate:

$$\dot{q} = [q]\dot{\boldsymbol{x}} + \dot{q}_{in\ situ} \tag{8.13}$$

Let's examine the flux more closely. It is the flux of gene frequency q in the x and y direction. It is not a flux of genes in the x and y direction, for which the symbol is $[G]\dot{x}$. The flux of genes through an imaginary surface separating two populations is easy enough to picture:



But this flux of genes will not necessarily alter gene frequency. Imagine, for example, two populations with exactly the same gene frequency, $q = 50\%$ of population size N. If 100 organisms migrate northward at a velocity represented by $\dot{y}$, and exactly half are carrying a copy of a particular gene, there will be a flux of gene copies $[G]\dot{y}$, with no change in gene frequency due to migration. For there to be a change in gene frequency, there must be a difference in gene frequency between the two populations. That is, there must be a spatial gradient in gene frequency.

Calling the migration of genes a *flux* seems to be somewhat idiosyncratic, but in fact many of the models developed to understand this process use the concept of a flux (e.g., Roughgarden, 1979). This excursion into simple population genetics shows how the concept of a flux, once grasped, can be applied to any quantity, leading to sometimes novel and interesting ways of looking at ecology and ecological genetics.

8.4.2 Fractal Fluxes

Fluxes are conventionally viewed in Euclidean grids and boxes as the movement of a quantity through a flat plane. Flat planes are more characteristic of manufactured objects, whereas for natural objects contorted and convoluted surfaces are the rule. Each of the examples at the beginning of the section on fluxes can be viewed as a flux through a flat plane, but each can also be viewed as a flux through a convoluted plane. In the ocean, nutrients mix upward through the convoluted surface of the thermocline. In the forest, rain strikes the convoluted canopy, drips through that canopy to the ground, and then soaks across the ground surface, which is rough at all spatial scales.

The idea of a flux through a convoluted or fractal surface is not standard, but it is a promising way of describing the dynamics of ecologically interesting quantities such as nutrients, energy, and biomass. It will be interesting to see whether the idea of a *fractal flux* will live up to its promise as a way of computing flux at one scale from flux at another.

ANOTHER LOOK AT SECTION 8.4

For a quantity of interest to you, sketch a diagram showing its flux. Consider how you might measure this flux.

8.5 Spatial Gradients

Ecologically important quantities vary considerably from location to location. The *spatial gradient* measures this contrast relative to the change in position. The spatial gradient

is defined as the difference in the value of a quantity at two locations relative to the separation. For example, Hill (1973) reported density of acacia plants, *Acacia ehrenbergiana*, in a sequence of quadrats surveyed by Greig-Smith and Chadwick (1965). Here are the counts in five contiguous quadrats, each 10 m · 10 m in size:

$$N = [5 \quad 51 \quad 22 \quad 12 \quad 11] \cdot \text{seedlings}$$

Each count is assigned a position according to the center of the quadrat in which it occurs. As a result, $N_{x=5\,\text{m}} = 5$ seedlings, $N_{x=15\,\text{m}} = 51$ seedlings. The gradient in numbers between the first two sites is calculated in the following manner:

$$\frac{N_2 - N_1}{x_2 - x_1} = \frac{51\text{seedlings} - 5\ \text{seedlings}}{5\ \text{m} - 15\ \text{m}} \Rightarrow 4.6\,\frac{\text{seedlings}}{\text{m}} \tag{8.14a}$$

This notation is fine for calculation but unwieldy for use in reasoning about gradients. The measured gradient in acacia numbers (N = seedlings) is abbreviated to:

$$\frac{\Delta N}{\Delta x} = \quad 4.6\,\frac{\text{seedlings}}{\text{m}} \tag{8.14b}$$

A still more compact notation for the measured value of the spatial gradient uses the *del* sign ∇ in front of the symbol for the quantity. Here is the notation for the measured gradient, using the generic symbol Q:

$$\nabla Q_x \equiv \frac{\Delta Q}{\Delta \boldsymbol{x}} \tag{8.15}$$

$\Delta \boldsymbol{x}$ stands for the measured change in position; ΔQ stands for the measured difference in the quantity. ∇Q is read as "the gradient in" whatever Q signifies, whether ant numbers, rainfall, or any other quantity.

The sequence of five measured values of acacia density results in four measured gradients ∇N_x, derived from N_x. Collected together, the gradients are:

$$\nabla N_x = [4.6 \quad -2.9 \quad -1 \quad -0.1]\dagger\ \text{seedlings/m} \tag{8.16}$$

The compact symbol here represents a collection of four gradients. The symbol can just as easily represent a much larger collection.

Spatial gradients are also calculated from functional expressions rather than from pairs of measurements. The symbol for the gradient in the x direction, calculated from a rule, is:

$$\boldsymbol{i}\frac{dQ}{dx}$$

dx represents an infinitesimally small change in the x direction, and $\boldsymbol{i}$ is the unit vector. Not every gradient falls conveniently along a single axis, such as x, the axis that by convention is oriented, or pointed eastward. But every gradient can be resolved into its

x (eastward), y (northward), and z (upward) component. The symbol for a gradient calculated from a rule is the sum of these three components:

$$\nabla Q(\boldsymbol{x}) \equiv \boldsymbol{i}\frac{dQ}{dx} + \boldsymbol{j}\frac{dQ}{dy} + \boldsymbol{k}\frac{dQ}{dz} \tag{8.17}$$

where dx, dy, and dz are infinitesimally small changes in the x, y, and z directions. $\boldsymbol{i}$, $\boldsymbol{j}$, and $\boldsymbol{k}$ are the unit vectors in the x, y, and z directions, respectively, as described in Chapter 7. In the example of the acacia seedlings, the unit vector had units of meters, and only the spatial dimension x was used.

To distinguish between gradients calculated from a rule and gradients calculated from measurements, the conventions in Chapter 7 will be used: functional notation for a derived quantity calculated according to a rule, a subscripted symbol for a derived quantity calculated from measurements only (Table 8.1).

Here is an example of gradients in two dimensions, again using the acacia data. The count of seedlings in five contiguous quadrats along two adjacent transects, as reported by Hill (1973), is:

$$N = \begin{bmatrix} 5 & 15 & 22 & 12 & 11 \\ 7 & 8 & 12 & 2 & 2 \end{bmatrix} \cdot \text{seedlings}$$

As before, the gradients are calculated relative to distance, but this time the unit vector will be in 10 m increments, in both the $\boldsymbol{i}$ and $\boldsymbol{j}$ directions. The first gradient, in Row 1, is 46 seedlings/decameter rather than 4.6 seedlings/meter, as in the previous calculation. This change in the size of the unit vector will make it easier to follow the next set of calculations.

To show the source of each gradient in both the x and y directions, I have inserted the value of the gradient, in boldface type, between the observations from which it was calculated:

$$\begin{bmatrix} 5 & \mathbf{+46} & 51 & \mathbf{-29} & 22 & \mathbf{-10} & 12 & \mathbf{-1} & 11 \\ \mathbf{+2} & & \mathbf{-43} & & \mathbf{-10} & & \mathbf{-10} & & \mathbf{-9} \\ 7 & \mathbf{+1} & +8 & \mathbf{+4} & 12 & \mathbf{-10} & 2 & \mathbf{0} & 2 \end{bmatrix}$$

The display contains a mixture of scalar (normal typeface) and vector (boldface) quantities, so neither a symbol nor units can be assigned to the display. The gradients are vector quantities, either positive (pointing right along transects and downward across transects) or negative (pointing left along transects and upward across transects).

Next, I have erased the observations. This leaves the topography, which consists of a set of gradients collected together into a matrix represented by a single symbol, ∇N:

$$\nabla N = \begin{bmatrix} & \mathbf{+46} & & \mathbf{-29} & & \mathbf{-10} & & \mathbf{-1} & \\ \mathbf{+2} & & \mathbf{-43} & & \mathbf{-10} & & \mathbf{-10} & & \mathbf{-9} \\ & \mathbf{+1} & & \mathbf{+4} & & \mathbf{-10} & & \mathbf{0} & \end{bmatrix} \cdot \frac{\text{seedlings}}{\text{decameter}}$$

Now the notation developed in Chapter 7 comes into its own, leading to a clear and direct way of expressing the topography of this quantity: eight gradients in the

x direction and five gradients in the y direction, calculated from 10 observations at 10 positions. In this example the symbol ∇N stands for the entire topography of 13 gradients calculated from 10 contiguous observations. The same symbol easily represents a still larger collection of gradients. Hill (1973) reported counts from 32 parallel transects, each consisting of five contiguous quadrats. The topography of this larger set consists of 32·4 gradients in the x direction and 5·31 gradients in the y direction. The symbol ∇N easily represents all 283 gradients, calculated from acacia counts at 160 contiguous sites.

In many situations it is convenient to ignore the vertical dimension. We will not find most species of plant rooted directly above another, although we might find epiphytes or cliff-loving species where this is so. When only the horizontal gradients need be considered, the appropriate symbol is ∇_h, placed in front of the symbol for the quantity of interest. ∇_h is read as "the horizontal gradient." $\nabla_h s$ is read as "the horizontal gradient in species number s."

At this point the reader is invited to write out the symbol for the gradient in each of the following quantities, using either ∇_h for the lateral gradient or ∇ for the gradient in three dimensions:

Soil temperature	T	_____
Population density	$[N]$	_____
Gene frequency	q	_____
Population biomass	M	_____
Primary production	$\dot{M}$	_____

The reader is next invited to visualize each of these derived quantities, then to assign each quantity a name. The gradients in some of these quantities have more than one name. For example, the lateral gradient in gene frequency q goes by the name *cline* in the evolutionary literature.

Gradients sometimes appear as percentages rather than absolute values. The symbol for the gradient as a percentage, using the generic symbol Q, is $Q^{-1}\nabla Q$. One way of reading this out is "the gradient in Q as a percentage of Q." Returning to the example of the acacia seedling counts along a single transect, N stands for seedling number, ∇N_x stands for the measured gradient in seedling number, and $N^{-1}\nabla N_x$ stands for the relative gradient in numbers. The quantity N has units of entities, and so $N^{-1}\nabla N$ could also be called the *per capita gradient*.

The relative gradient of a quantity is calculated as follows:

$$N = [\ 5 \qquad 51 \qquad 22 \qquad 12 \qquad 11\ ] \cdot \text{seedlings} \tag{8.18a}$$

$$\nabla N_x = [\ 4.6 \qquad -2.9 \qquad -1 \qquad -0.1\ ] \cdot \text{seedlings m}^{-1} \tag{8.18b}$$

$$N^{-1}\nabla N_x = \left[\ \frac{460}{5} \qquad \frac{-290}{51} \qquad \frac{-100}{22} \qquad \frac{-10}{12}\ \right] \cdot \%\text{m}^{-1} \tag{8.18c}$$

$$N^{-1}\nabla N_x = [\ 92 \qquad -5.7 \qquad -4.5\dagger \qquad 0.83\ ] \cdot \%\text{m}^{-1} \tag{8.18d}$$

This notation applies to any symbol. For example, try writing the symbol for the relative gradient of q = gene frequency. Then try calculating the relative gradient in gene frequency at three positions, separated from each other by 100 kilometers, where q = [50% 85% 100%].

As another example of the gradient in a quantity, try visualizing the way in which the energy cost of territorial defense changes with increasing territory size for a circular territory around a nesting site. Then try writing the symbol for the percent gradient in energy cost using E = kJ day^{-1} for energy cost.

ANOTHER LOOK AT SECTION 8.5

In what contexts have you encountered gradients in ecology? Be sure to include contexts in which a gradient was used but not labeled as such.

8.6 Divergences

What about taking the gradient of a directed or vector quantity? What if we take the gradient in the lateral flux of genes at a series of positions? To visualize the gradient of a directed quantity, draw a checkerboard grid of 4 squares by 4 squares. North is up or away from you, east is to your right. Now place two arrows pointing north ↑↑ on two adjacent squares along the south side of the grid. When the arrows move to the square at which they are pointing, they retain their left-to-right spacing. Now place two arrows on two adjacent squares in the following arrangement: ↗ ↖. When these arrows move to the square at which they are pointing, they converge into a single square. Finally, here is a divergent arrangement ↖ ↗ with arrows pointing away from adjacent squares.

Given a divergent or convergent arrangement, the net gradient in a given direction can be calculated. Box 8.6 shows calculation of an eastward gradient for four different arrangements of two vector quantities. To learn something surprising about these divergent or convergent quantities, try calculating them from right to left (take left as positive) rather than from left to right (right positive), as in Box 8.6. Remember that the sign of the quantity changes, as does the order in which the gradient is calculated.

Box 8.6 Calculation of the Gradient of a Vector Quantity at Two Points

First, the eastward gradient. The eastward component is positive ↗, negative ↖, or zero ↑. For eastward components having a magnitude of 1 unit, the eastward gradient of eastward components is calculated for several arrangements.

Divergent Arrangement	Calculation				Gradient is:
↖ ↗	(↗)	− (↖)			
	(+1)	− (−1)	=>	+ 2	Positive
Convergent Arrangement	**Calculation**				
↗ ↖	(↗)	− (↖)			
	(−1)	− (+1)	=>	− 2	Negative
Parallel Arrangements	**Calculation**				
↑ ↑	(↑)	− (↑)			
	(0)	− (0)	=>	0	Zero
↖ ↖	(↖)	− (↖)			
	(−1)	− (−1)	=>	0	Zero

Now the westward gradient, taking left as positive: Calculating from left to right changes the sign of the quantity as well as the order in which the gradient is calculated.

Arrangement	Calculation	Gradient is:
↖ ↗	(↖) − (↗)	
↖ ↗	(__) − (__) => ___	_____

The final calculation in Box 8.6 shows that the sign of the gradient is independent of the direction in which the gradient is taken. Hence a positive gradient in a directed quantity always refers to a situation in which the arrows point away from each other, or diverge. Similarly, a negative gradient always indicates *convergence*. The *divergence* is defined as the gradient in a directed quantity. The gradient is either positive (divergent) or negative (convergent).

Now on to locust flux instead of arrows. In applications, fluxes are usually treated in two or even three spatial dimensions, not just one. For example, the eastwardly flux of locusts swarming over a grid of wheat fields is $[N]\dot{x}$ or $N\rightarrow$ for short. The northerly flux is $[N]\dot{y}$ or $N\uparrow$. Box 8.7 shows the flux gradient across the field for two patterns of flux as measured on all sides of the field. If the farmer observes that the flux arriving from the west side exceeds the flux leaving on the east, the farmer has a problem: a negative gradient in the flux of locusts that are converging on the farm. Similarly, if the flux coming from the north exceeds the flux leaving across the south side of the field, locusts are accumulating. The total lateral gradient in locust flux is the easterly gradient in the eastward flux plus the northerly gradient in the northward flux. The symbol for this total gradient is $\nabla_h \cdot [Q]\dot{x}$ This is the divergence, or gradient, in locust flux, which measures the magnitude of the problem of locust accumulation in a defined area.

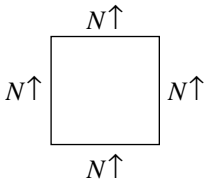
The total horizontal gradient in flux (total divergence) is the sum of the components where *negative divergence* is the same as convergence. Box 8.7 shows two patterns of total divergence. In the first, the total divergence due to simple northward flux of locusts is zero. In the second, the flux is convergent because the north-south divergence is negative, while the east-west divergence is zero.

The total horizontal gradient in flux is represented by the vector product $\nabla\cdot$ applied to a symbol for a vector or directed quantity (the gradient operator ∇ with no dot is applied to scalar quantities). The sign $\nabla\cdot$ in front of the symbol simply means to take the sum of the gradients in two or three directions, depending on the number of spatial dimensions specified for the quantity. This is a compact way of keeping track of ecologically interesting processes such as convergent and divergent fluxes of organisms, or genes, or energy fixed by photosynthesis and transferred to higher trophic levels.

The symbols for divergence and convergence of directed quantities are convenient for reasoning about quantities. But the symbols are abstract and hence need to be associated with images of divergent and convergent motion. Boxes 8.8–8.12 show patterns of divergent, convergent, and translatory motion of particles around a point in two spatial dimensions x (positive eastward) and z (positive upward). The reader is invited to verify that the mathematical symbols for each type of motion are correct by taking the sign of the x and z gradients on moving from one arrow to the next $\rightarrow\rightarrow$ in each diagram.

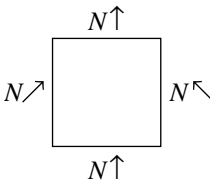
Box 8.7 The Lateral Flux Gradient

Easterly flux: $[N]\dot{x}$ or $N\rightarrow$
Northerly flux: $[N]\dot{y}$ or $N\uparrow$
Northeastward flux: $N\nearrow$
Northwestward flux: $N\nwarrow$



North to South $\dfrac{\Delta[N]\dot{x}}{\Delta x} = 0$

East to West $\dfrac{\Delta[N]\dot{y}}{\Delta y} = 0$



North to South $\dfrac{\Delta[N]\dot{x}}{\Delta x} = 0$

East to West $\dfrac{\Delta[N]\dot{y}}{\Delta y} < 0$

The total lateral gradient in locust flux is the easterly gradient in the eastward flux plus the northerly gradient in the northward flux:

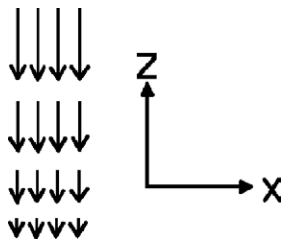
$$\nabla_h \cdot [\dot{Q}] = \frac{\Delta[Q]\dot{x}}{\Delta x} + \frac{\Delta[Q]\dot{y}}{\Delta y}$$

Box 8.8 Vertically Convergent Motion.

z is vertical (positive upward), x is horizontal (positive to the right).

$$\frac{\Delta[Q]\dot{x}}{\Delta x} = 0$$

$$\frac{\Delta[Q]\dot{z}}{\Delta z} < 0$$

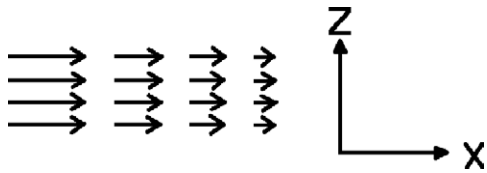


Box 8.9 Horizontally Convergent Motion.

z and x as in Box 8.8:

$$\frac{\Delta[Q]\dot{x}}{\Delta x} < 0$$

$$\frac{\Delta[Q]\dot{z}}{\Delta z} = 0$$

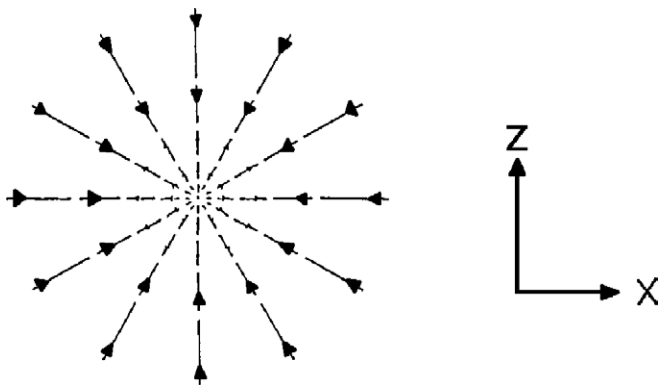


Box 8.10 Horizontally and Vertically Convergent Motion.

z and x as in Box 8.8:

$$\frac{\Delta[Q]\dot{x}}{\Delta x} < 0$$

$$\frac{\Delta[Q]\dot{z}}{\Delta z} < 0$$

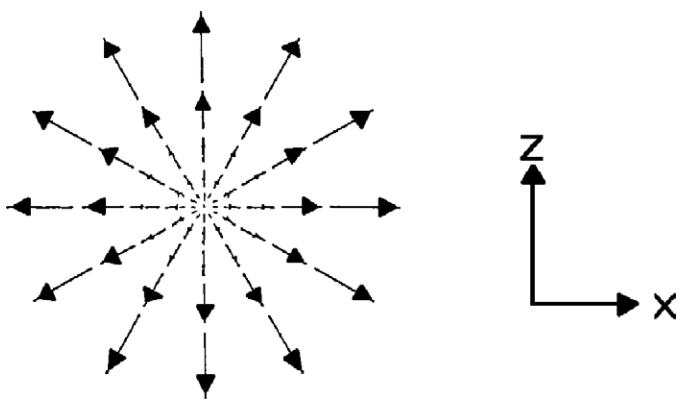


Box 8.11 Horizontally and Vertically Divergent Motion. Vertically Convergent Motion z is vertical (positive upward), x is horizontal (positive to the right).

z and x as in Box 8.8:

$$\frac{\Delta[Q]\dot{x}}{\Delta x} > 0$$

$$\frac{\Delta[Q]\dot{z}}{\Delta z} > 0$$

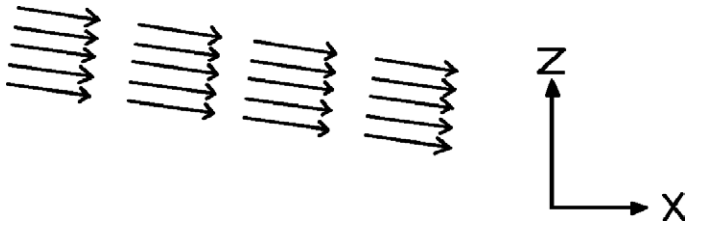


Box 8.12 Translatory Motion, No Divergence (Convergence).

z and x as in Box 8.8:

$$\frac{\Delta[Q]\dot{x}}{\Delta x} = 0$$

$$\frac{\Delta[Q]\dot{z}}{\Delta z} = 0$$



Convergent and divergent motions generate smaller-scale patchiness in natural populations during dispersal stages and then throughout the lifetime of mobile organisms. At larger scales, patchiness results from convergent and divergent motion over generations. The divergence (and convergence) of a population, defined as the gradient in the numerical flux $\nabla \cdot [N]\dot{\boldsymbol{x}},$, expresses the idea of coalescence and dispersal in a way that, though not intuitive, does allow calculations of change in patchiness.

An extension of this idea is that the rate of change in the volume occupied by a group of organisms is equal to the divergence of their lateral velocities. Here is a formal expression of the idea of change in volume due to divergence:

$$\dot{V} \equiv \nabla \cdot \dot{\boldsymbol{x}} \tag{8.19}$$

This is the *divergence theorem*, which relates the time rate of change in volume V of a parcel to quantity $\nabla \cdot \dot{\boldsymbol{x}}$, the divergent motion resolved into three velocity components $\dot{\boldsymbol{x}} = [\dot{x}\ \dot{y}\ \dot{z}]$. Meteorology texts (e.g., Dutton, 1975) contain concrete treatments of this abstract idea. In words, the theorem says that the time rate of change in the volume of a parcel is equal to the divergence of the velocities perpendicular to the surface boundary of the parcel. If the boundary is defined as a series of triangular surfaces connecting peripheral organisms, the boundary stretches outward if an organism crosses outward through one of these triangular surfaces. If an organism at the edge moves inward, the edge shrinks inward. Movement inward has a negative divergence, a backward way of saying that movement is convergent. As a result, the occupied volume contracts ($\dot{V} < 0$).

Divergence of a quantity depends on spatial scale as measured by the unit vectors $\boldsymbol{i}$, $\boldsymbol{j}$, and $\boldsymbol{k}$ that underpin the velocity component $\dot{\boldsymbol{x}} = [\dot{x}\ \dot{y}\ \dot{z}]$. At a small scale the divergence in the vicinity of a waterspout or tornado may be considerable. At a larger scale, with larger unit vectors, the divergence in the atmosphere and the ocean can be substantial due to west-to-east passage of weather systems and to the major oceanic gyres. In animal populations, the divergence may be large at small scales, as in the vicinity of feeding aggregations. At larger scales the divergence in animal populations will usually be small, except in highly migratory species.

Divergent motion captures the idea of coalescence and dispersal at multiple scales. Divergent motion leads to changes in patchiness, crowding, and frequency of contact. At a small scale, convergence increases local density, thereby increasing the opportunity

for contact, defined as the number of potential interactions within some fixed distance (see Chapter 10). Convergence thus alters the potential for interactions requiring direct contact: predation, competition, and gamete exchange. These processes require direct contact, unlike gravity and electromagnetic forces, which act at a distance according to inverse square laws. The rate at which contact-dependent ecological processes proceed varies with the opportunity for interaction, which in turn depends on local density and degree of aggregation. Divergent and convergent motions alter local density and hence modify the opportunity for interaction of organisms among and within populations.

ANOTHER LOOK AT SECTION 8.6

Would you use the same unit vector to describe vertical and horizontal divergence of water motions in a lake, a river, and the ocean? How would you choose the vector?

8.7 Curls

Ecological interactions occur within the fluid envelopes of the atmosphere or the ocean. One of the characteristics of a fluid is that it cannot resist mechanical stresses. When force is applied, water and air rotate rather than resisting the force. Rotation occurs at small scales, as in the eddies that form behind trees in a wind. Rotation occurs at enormous scales, as in the North Atlantic gyre, which carries water clockwise past North America toward Europe. The perpetual rotary motion of the fluid envelope inhabited by life has important effects. For example, the eddies that form over rough soil areas allow seeds to settle more readily than in smoother areas. Migratory birds use the rotary motions of weather systems to accomplish migration over phenomenal distances.

These rotary motions, and the interaction of organisms with rotary motions of the surrounding atmosphere and ocean, are calculated relative to the flux gradient, as with the divergence. But now the change in flux relative to the gradient, called the *curl*, will be taken at right angles to the flux rather than along the same axis. For example, the northerly gradient in the eastward component is taken instead of the northerly gradient of the northward component. The result is called, appropriately enough, the *curl* of the directed quantity. The curl is the sum of the gradients taken at right angles to the flux gradients. The curl, like the divergence, is positive (clockwise), negative (counterclockwise), or zero (no rotary motion). An example of a biological flux with positive curl is the elliptical migration of bird populations. Migrants from the Arctic tend to follow routes that are elliptical because of displacement to the west during northward migration and displacement to the east during southward migration.

The sign that designates the curl of a vector quantity is, again, the gradient operator ∇ but this time accompanied by the sign for the vector cross-product $\nabla\times$ rather than the sign $\nabla\cdot$ for the vector dot product. Throughout their life cycles, organisms interact with rotary motions of the surrounding air and water. So, the curl $\nabla \times Q$ is a quantity that is ecologically important.

Texts on vector and matrix arithmetic show how cross-products and curls are calculated. These calculations might seem familiar to readers who have mastered the computational basis of multivariate analysis, a sophisticated form of statistics that has been a

popular way of searching for patterns in ecological data. The reason that these computations may seem familiar is that the curl of a quantity defines an axis of rotation in addition to the degree of rotation. Multivariate statistical analysis uses similar forms of calculation to find a series of best-fitting or "canonical" axes, all at right angles to one another. The curl, in a geographic application, defines an axis at right angles to two previously defined axes. A vivid image is the central axis of a tornado, which tilts this way and that.

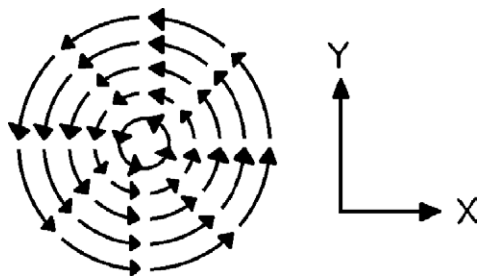
The next two boxes show two forms of curl, or rotary motion. The reader is again invited to verify that the mathematical formula matches the picture by calculating the gradients in successive arrows in the diagram. Box 8.13 shows simple rotary motion in the horizontal plane.

Box 8.13 Simple Rotary Motion, Looking Down on the Horizontal (x-y) Plane.

x and y are positive in the direction of the arrows:

$$\frac{\Delta[Q]\dot{y}}{\Delta x} > 0$$

$$\frac{\Delta[Q]\dot{x}}{\Delta y} < 0$$



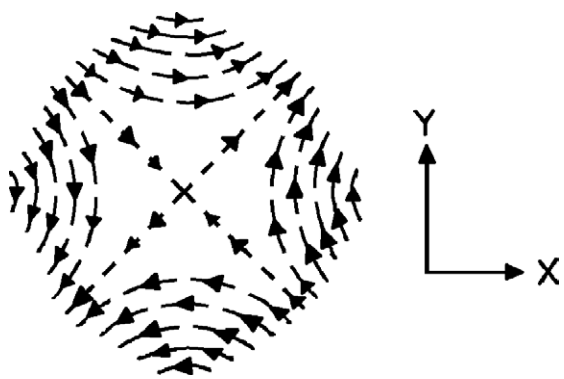
The next box shows *shearing*, again looking downward on the horizontal plane. This is the motion that occurs at the boundary of two air masses blowing past one another or two currents flowing past one another.

Box 8.14 Shearing Motion, Looking Down on the Horizontal (x y) Plane.

x and y are positive in the direction of the arrows:

$$\frac{\Delta[Q]\dot{y}}{\Delta x} > 0$$

$$\frac{\Delta[Q]\dot{x}}{\Delta y} > 0$$



The motions of ecologically interesting quantities do not occur in the pure form shown in these boxes. Rather, motions of nutrients, water, or food particles consist of a combination of these forms. Often, a form of motion at one spatial scale lies embedded in a larger-scale motion. For example, bottom fish migrate by moving up off the seafloor at one stage of the tide, which sweeps them along part of an ellipse for a period of time, because tidal currents away from the coast are rotary in form. The fish then move back to the bottom, remaining in place as the tide sweeps around the remaining compass points in its cycle. The net result of a series of smaller-scale rotary motions with the tide is a larger-scale translatory motion that accomplishes migration with minimum swimming effort (Harden-Jones, Walker, and Arnold, 1978).

ANOTHER LOOK AT SECTION 8.7

Compare the strength of the horizontal curl of water velocity in the narrow upper reaches of a river with that in the wider lower reaches of a river. Give an approximate unit vector for your description.

8.8 Contrasts

Thus far all these derived quantities have been calculated at a fixed resolution. Any of these quantities can be calculated at several different resolutions. For example, the gradient in acacia seedlings can be calculated at a resolution of two quadrats ($2\boldsymbol{i} = 20\,\text{m}$) rather than one quadrat ($\boldsymbol{i} = 10\,\text{m}$). Or it can be calculated at a resolution of three quadrats ($3\boldsymbol{i} = 30\,\text{m}$). All these calculations can be gathered together into a triangular matrix. Here is the complete set of gradients, at four different spatial scales, for the five acacia counts along a single transect:

$$\begin{bmatrix} +46 & -29 & -10 & -1 \\ +17 & -39 & -11 & \\ +7 & -40 & & \\ +6 & & & \end{bmatrix} \cdot \frac{\text{seedlings}}{i}$$

The first row in the matrix (at resolution of $\boldsymbol{i} = 10\,\text{m}$) has already been calculated and displayed previously. The second row in the matrix shows the gradient at a resolution of 20 m. This is calculated as the difference between every other count, divided by one unit and also expressed in terms of this unit. The third row shows the gradient calculated from every third quadrat. Only one gradient can be calculated at a resolution of four quadrats, as shown in the last row of the matrix.

The pattern of contrasts shows the complete topography of the five counts. This topography can be examined by panning or zooming. Panning corresponds to examination of the matrix from left to right within a row. In the preceding example, the gradient at the resolution scale of one quadrat decreases from left to right. Comparison across the first row shows that positive gradients are confined to the beginning of the transect.

Zooming in on detail corresponds to examining the matrix from bottom to top within a column. In the first column, for example, zoom comparisons are made relative to the first count (5 acacia seedlings). The gradient changes from 46 per unit ($\boldsymbol{i} = 10\,\text{m}$ increments) to 17 per unit ($2\boldsymbol{i} = 20\,\text{m}$). The gradient then decreases to 7 per unit ($3\boldsymbol{i} = 30\,\text{m}$) and finally drops to 6 per unit ($4\boldsymbol{i} = 40\,\text{m}$). The gradient, relative to the first count, declines with increase in the separation between locations.

For want of a better term, I have labeled this a *matrix of contrasts*. The quantity calculated in the matrix could be called the contrast in Q, by analogy to terms such as the gradient in Q. Taking this a step further, a symbol can be assigned to the matrix, Cntrst(∇N), to indicate that contrasts in the quantity N have been calculated at several resolution scales. The symbol ∇ stands for $\Delta/\Delta x$, the operation of comparing adjacent values of the quantity N as a difference.

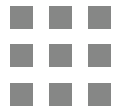
The contrast is a derived quantity that permits one to zoom in on detail or zoom back on the larger-scale structure. Changes in scale of attention are as informative and necessary in the natural sciences as they are in art.

> *There is a kind of indeterminacy … which lies in the fact that we can neither consciously sense nor think of very much at any one moment. Understanding can only come from a roving viewpoint and sequential changes in the scale of attention.*
>
> —C. S. Smith, *Structure in Art and in Science*, 1965

Calculation and display of all the gradients of acacia counts permit the roving viewpoint (scanning comparisons made from left to right) to be combined with sequential changes in scale of attention (zooming comparisons made from top to bottom).

Defined Concepts and Terms for Review and Future Reference

____ crude rates of change
____ curl $\nabla \times Q$
____ divergence (convergence) $\nabla_h \cdot Q$
____ divergence theorem $\dot{V} \equiv \nabla \cdot \dot{x}$
____ flux $\dot{Q}/A = [Q]\dot{x}$
____ instantaneous rates of change
____ matrix of contrasts
____ percent rate of change
____ spatial gradient ∇Q
____ time rate of change $\dot{Q}$



9 Ensemble Quantities: Weighted Sums

On one occasion Kelvin made a speech on the overarching importance of numbers. He maintained that no observation of nature was worth paying serious attention to unless it could be stated in precisely quantitative terms. The numbers were the final and only test, not only of truth but about meaning as well. He said, "When you can measure what you are speaking about, and express it in numbers, you know something about it. But when you cannot—your knowledge is of a meagre and unsatisfactory kind."

… Kelvin may have had things exactly the wrong way around. The task of converting observations into numbers is the hardest of all, the last task rather than the first thing to be done, and it can be done only when you have learned, beforehand, a great deal about the observations themselves. You can, to be sure, achieve a very deep understanding of nature by quantitative measurement, but you must know what you are talking about before you can begin applying the numbers for making predictions.

—Lewis Thomas, *Late-Night Thoughts on Listening to Mahler's Ninth Symphony*, 1983

9.1 Synopsis

This chapter describes quantities that are obtained by summing the values of variable quantities. These ensemble quantities include the time average $\bar{Q}_t$ and the spatial average $\bar{Q}_x$. These are biologically interpretable quantities. For example, the downward flux of seeds to the ground, summed over a mosaic of cleared and forested areas, gauges the colonization capacity of a plant population.

The rules for summing the values of a quantity differ from those for summing unitless numbers. The only certain procedure is weighted summation, together with biological or physical reasoning to determine whether there has been a change in scale. Summation across values of a scaled quantity occurs either by juxtaposing or by superposing. Juxtaposing the values of a quantity changes the scale by extending the range. An example is summing the density of seeds across a sequence of adjacent plots to obtain density over all plots. Superposing values leaves the scale unchanged. An example is repeatedly adding handfuls of seeds to a plot of ground.

The time average of a quantity typically consists of a juxtaposed sum that represents a longer time scale than individual measurements. Similarly, the spatial average is a single-valued quantity that represents a larger scale than the values from which it is calculated.

doi: 10.1016/B978-0-12-627865-1.00009-X

9.2 Notation

Ecology is the study of aggregate outcomes. How will hickory populations respond to global warming? What is the primary production by phytoplankton over millions of square kilometers during a spring bloom? The quantitative operation fundamental to an aggregate result is summation, which generates a single-valued quantity from a collection of values. The result is an interpretable quantity. For example, the sum of termite lifetimes over a yearlong period (T in units of termite-days per year) gauges the voracity of the colony in consuming wood. Summation of the number of pairwise combinations of 4 skuas with 100 kittiwake nests in a colony delimits the potential for nest predation in the colony. These sums deserve a name: an *ensemble quantity* is defined as the biologically interpretable sum of a sequence of values of a scaled quantity.

Care is needed in calculating ensemble quantities, because the rules for summing scaled numbers differ from those for unitless numbers. Summing the values of a scaled quantity often produces a different result than summing a series of numbers. Summation changes the spatial or temporal scale in some cases but not others.

As with any quantitative method, the operation of summing the values of scaled quantities will benefit from a consistent and readable notation. Chapter 7 expressed a preference for using spatial and temporal attributes as labels, but settled for the use of a unitless index consisting of integer numbers to coordinate measured values of a quantity with its spatial and temporal attributes. Table 9.1 develops this notation further as it applies to taking sums.

In Table 9.1 the generic symbol i is used to represent any index. Symbols other than i often appear, but certainly the commonest set is i, j, and k. These indices differ from the unit vectors $\boldsymbol{i}$, $\boldsymbol{j}$, and $\boldsymbol{k}$, which typically have units. The position vector $\boldsymbol{x} = x \cdot \boldsymbol{i}$ contains the unitless number x, which can serve to index a quantity measured at a regular series of locations. Position in time $\boldsymbol{t} = t \cdot \boldsymbol{h}$ contains the unitless symbol t that similarly can serve as an index.

Table 9.1 Notation for Taking the Sum of Unitless Numbers

Summation is over an index, which has a beginning point, a fixed increment, and an ending value n that defines the range. An index does *not* have units.

$i :=$ start, increment, …. end
$i := 1,10,1000$
$i := 1,1, \ldots n$ shortened to $i := 1 \ldots n$
q represents a vector of unitless numbers.

The summation sign represents taking the sum.

$$\sum q = \sum_{i=1}^{n} q_i = (q_1 + q_2 + \ldots + q_n)$$

The sign $\sum$ operates on a collection of numbers arranged in a row or column vector. Detail is added to the symbol as needed.

Index implied.	$\sum q$
Index implied, with upper limit of n; implies a starting point of $i = 1$.	$\sum_{n} q$ $\quad$ $\sum_{n} q$
Index explicit, but limits defined elsewhere.	$\sum q_i$

ANOTHER LOOK AT SECTION 9.2

Apply the operator $\sum_{i=1}^{n}$ to the quantity N = number of plant species in 10 contiguous quadrats, each 1 m^2, along a 10 m transect. What is the relation of the index $i = 1 \ldots n$ to the resolution and range of the quantity N?

9.3 Sums and Weighted Sums of Numbers

With notation established, the next step is to examine the rules for summing numbers having no units. Many readers will have encountered these rules already, perhaps in making statistical calculations. The rules in Table 9.2 will always work for numbers without units. Box 9.1 shows calculations based on these rules.

Table 9.2 Rules for Summing Numbers Without Units

q represents a collection of numbers gathered into a vector.
p represents another collection of numbers.
k represents a constant number.

$$\sum^{n} k = n \cdot k$$

$$\sum k \cdot q = k \cdot \sum q$$

$$\sum (q + p) = \sum q + \sum p$$

Box 9.1 Sums Calculated According to Rules for Numbers.

Equality by definition $k := 5$ is distinguished from equality by calculation $2 + 2 \Rightarrow 4$.

$$\begin{aligned} k &:= 5 \\ q &:= [1\ 2\ 3] \\ p &:= [0.1\ 0.2\ 0.3] \end{aligned}$$

$$\begin{aligned} \sum^{n} k &= 5 + 5 + 5 && \Rightarrow 15 \\ n \cdot k &= 3 \cdot 5 && \Rightarrow 15 \end{aligned}$$

$$\begin{aligned} \sum c \cdot q &= 5 \cdot 1 + 5 \cdot 2 + 5 \cdot 3 && \Rightarrow 30 \\ c \cdot \sum q &= 5 \cdot 6 && \Rightarrow 30 \end{aligned}$$

$$\sum (q + p) = (1 + 0.1) + (2 + 0.2) + (3 + 0.3) \Rightarrow 6.6$$

$$\sum q + \sum p = (1 + 2 + 3) + (0.1 + 0.2 + 0.3) \Rightarrow 6.6$$

The rules for adding numbers (Table 9.2) are a special case of *weighted summation*. Box 9.2 compares weighted with unweighted summation.

Box 9.2 Weighted and Unweighted Summation

Formula: $\bar{q}_w \equiv \dfrac{\sum q_i w_i}{\sum w_i}$

Example: $q = [4\ 5\ 6]$, $n = 3$

Unweighted summation: $\sum q = 4 + 5 + 6 \Rightarrow 15$

Equally weighted summation:

$$n^{-1}\sum q = (4 + 5 + 6)/3 \Rightarrow 5$$

$$\sum n^{-1} q = 0.33{\cdot}4 + 0.33{\cdot}5 + 0.33{\cdot}6 \Rightarrow 5$$

$w_i = n^{-1}$ $\qquad (\sum w_i)^{-1}\,(\sum q_i w_i) \Rightarrow 5$

Unequally weighted summation:

$p_i = (70\%\ 10\%\ 20\%)$

$$\sum p_i q = 0.7{\cdot}4 + 0.1{\cdot}5 + 0.2{\cdot}6 \Rightarrow 4.5$$

$w_i = (\ 7\ 1\ 2\ )$

$$(\sum w_i)^{-1}(\sum q_i w_i) = 10^{-1}(28 + 5 + 12) \Rightarrow 4.5$$

One way to visualize a weighted sum is as the pivot point of a stick balanced on a fulcrum. As stones are placed on the stick, the balance point shifts, depending on the mass of a stone and its distance from the center of the stick. All the points on the stick have the same weight, but when the weight at a site is increased by placing a stone, that point will have a greater influence on the pivot point than its neighbors. To give an example, the mean (point of balance) in the sequence $q = [\ 4\ 5\ 6\ ]$ is normally 5, but the weights (percentages) in Box 9.2 shift the point of balance to 4.5. Not all weighting factors are expressed as percentages, so the general formula (Box 9.2) takes this into account by using the sum of the weights $\sum w_i$ in the denominator. A convenient symbol for the weighted sum is to draw a line (which can be imagined as the pivoting stick) directly over the symbol for the quantity.

ANOTHER LOOK AT SECTION 9.3

To see how weighted summation works, guess the approximate value of the weighted sum of $q = [4\ 5\ 6]$ if weights are $w = [2\ 1\ 7]$. Then compute.

$$\bar{q} = \frac{\_\cdot\_ + \_\cdot\_ + \_\cdot\_}{\_ + \_ + \_} = \_$$

9.4 Sums of Scaled Quantities

We encounter inconsistencies if we apply a single summation rule to all quantities. For example, if a honeybee flies northward for distances of 7 m, 8 m, 8 m, 9 m, and 8 m in five successive periods of time, distance flown is:

$$y = [7\ 8\ 8\ 9\ 8] \cdot \mathrm{m}$$

The total distance flown is $\sum y = 40\,\mathrm{m}$, an easily visualized quantity. If a bee flies at around $8\,\mathrm{m\ s^{-1}}$, five successive determinations of this quantity might look like this:

$$\dot{y} = [7\ 8\ 8\ 9\ 8] \cdot \mathrm{m\ s^{-1}}$$

The sum $\sum \dot{y} = 40\,\mathrm{m\ s^{-1}}$ is clearly incorrect.

Definitions from the literature provide a name for the problem, but no remedy. For example, Shugart (1998, p. 94) describes the principle of superposition for linear systems as meeting the condition that the response of a system to several inputs in time is equal to the sum of the responses of the inputs acting alone. This condition for linear systems matches the rules for adding unitless numbers (Table 9.2) if we take q as the input and k as the response. This gives us a name, *superposable*, for quantities that follow the same rules as taking the sum of numbers (Table 9.1). But the definition serves only to label the distinction. It does not provide a practical criterion for distinguishing superposable quantities before they are summed.

The rules for weighted summation, combined with reasoning about units, will provide that criterion. The procedure will be to use weighted summation (Box 9.2) with units as weights. Reasoning about the sum of the weighting factors $\sum w$ will guide the computation of the sum of multiple values of a scaled quantity Q. Here is an example, using bee distances, for which the sum $\sum y = 40\,\mathrm{m}$ is a readily interpretable ensemble quantity. The first step is to choose a weighting factor that results in unitless numbers that can be summed. For the bee distances, this factor is $\mathrm{m^{-1}}$:

$$\bar{y} = \frac{7\ \mathrm{m} \cdot \frac{1}{\mathrm{m}} + 8\ \mathrm{m} \cdot \frac{1}{\mathrm{m}} + 8\ \mathrm{m} \cdot \frac{1}{\mathrm{m}} + 9\ \mathrm{m} \cdot \frac{1}{\mathrm{m}} + 8\ \mathrm{m} \cdot \frac{1}{\mathrm{m}}}{\sum \mathrm{m^{-1}}}$$

The weighting factor $\mathrm{m^{-1}}$ normalizes the measurements (7 m, 8 m, and so on) to dimensionless numbers: $(7\,\mathrm{m})(\mathrm{m^{-1}}) = 7$:

$$\bar{y} = \frac{7 + 8 + 8 + 9 + 8}{\sum \mathrm{m^{-1}}}$$

The next step is to interpret the sum of the weighting factors, $\sum \mathrm{m^{-1}}$. If the values are juxtaposed (which means the bee gets somewhere), the spatial frequency $\mathrm{m^{-1}}$ does not change. It is one measurement per meter. The formal expression of this is $\sum \mathrm{m^{-1}} = \mathrm{m^{-1}}$. Now when the calculation is made, the spatial sum is:

$$\bar{y} = \frac{40}{\mathrm{m^{-1}}} = 40\ \mathrm{m}$$

Table 9.3 Generic Recipe for Summing Scaled Quantities

1. Write out the quantity as a vector with units.
2. State whether summation will occur with respect to distance, time, or some other dimension.
3. State a weighting unit that clears units from the quantity. Write this as a vector of weights w_i.
   Weights can be equal $w = [1\text{U}\ 1\text{U}\ \ldots\ 1\text{U}]$
   Weights can be unequal $w = [q_1 \cdot 1\text{U} \quad q_2 \cdot 1\text{U} \ldots]$
4. Apply the weighting, according to the following formula:

$$\bar{Q}_w = \frac{\sum Q_i w_i}{\sum w_i}$$

5. Determine whether summation changes the scale of the weighting factor:
   If the scale is unchanged: $\sum w = 1\text{U}$ (equal units)
   $\sum w = 1\text{U} \cdot \sum q$ (unequal units)
   If the scale is changed: $\sum w = n \cdot 1\text{U}$ (equal units)
   $\sum w = n \cdot 1\text{U} \cdot \sum q$ (unequal units)
6. Complete the calculation.

If the values are superposed (5 bees starting at the same point), the spatial frequency changes. It becomes 5 measurements per meter and so $\sum \text{m}^{-1} = \text{n} \cdot \text{m}^{-1}$. With this calculation, the spatial sum is:

$$\bar{y} = \frac{40}{5 \cdot \text{m}^{-1}} = 8\ \text{m}$$

Summing a scaled quantity resulted in different outcomes, depending on how summation was interpreted relative to the scaling unit, which in this case was spatial frequency. The same applies to the summation of bee velocities at the beginning of this section. Given the way measurements were taken, the summation $\sum \dot{y} = 40\ \text{m s}^{-1}$ was incorrect; the summation $\sum \dot{y} = 8\ \text{m s}^{-1}$ was correct.

Now that an example is at hand, we move to a detailed recipe for summing quantities. The recipe goes like this: state the weighting factor, find units that clear this factor from the units of the quantity, rescale the quantity according to this factor, sum the rescaled quantities, interpret the result of summing the scaling unit, and divide by the sum of the scaling units. Table 9.3 presents the generic recipe.

The recipe for weighted summation is next applied to bee velocities, where a nonsensical result of $\sum \dot{y} = 40\ \text{m s}^{-1}$ resulted from simple summation with no attention to units. Box 9.3 shows the calculations, following the recipe in Table 9.3, for summation by *juxtaposition* and *superposition*.

The generic recipe (Table 9.3) for summing the values of a scaled quantity has three major ingredients: a statement of the weighting unit, a formula for applying the weighting units, and a statement of whether summation changes the scale of the weighting unit. Stating how summation is to occur (Step 2) depends in large part on the purpose of the calculation. In Step 3 the units are most commonly time, distance, and area (or their inverses, temporal and spatial frequency). The weighting units are chosen in a way that clears units from the numerator. If the quantity has units of km^2, the weighting factor will be km^{-2}. The fifth step, interpreting the sum of the weights, is the most

Box 9.3 Calculating the Sum of Scaled Quantities.

For Equal Units

1. Write the quantity as a vector with units:

$$\dot{y} = [7\ 8\ 8\ 9\ 8] \cdot \mathrm{m\ s^{-1}}$$

2. Measurements are at five contiguous positions, so summation will occur by juxtaposing distance, not time.
3. Two weighting factors are needed to clear units:

$$w_1 = \mathrm{s} \text{ and } w_2 = \mathrm{m^{-1}}$$

4. Apply formula for weighted summation:

$$\bar{\dot{y}} = \frac{\left(7\frac{\mathrm{m}}{\mathrm{s}}\right)\frac{\mathrm{s}}{\mathrm{m}} + \left(8\frac{\mathrm{m}}{\mathrm{s}}\right)\frac{\mathrm{s}}{\mathrm{m}} + \left(8\frac{\mathrm{m}}{\mathrm{s}}\right)\frac{\mathrm{s}}{\mathrm{m}} + \left(9\frac{\mathrm{m}}{\mathrm{s}}\right)\frac{\mathrm{s}}{\mathrm{m}} + \left(8\frac{\mathrm{m}}{\mathrm{s}}\right)\frac{\mathrm{s}}{\mathrm{m}}}{\sum \mathrm{m^{-1}} \sum \mathrm{s}}$$

5. Interpret summed weights:
   $\sum \mathrm{m^{-1}} = 5 \cdot \mathrm{m^{-1}}$ because summation is by juxtaposition
   $\sum \mathrm{s} = 1 \cdot \mathrm{s}$ because summation is by superposition
6. Complete the calculation:

$$\bar{\dot{y}} = \frac{7 + 8 + 8 + 9 + 8}{5\ \mathrm{m^{-1}} \cdot 1\mathrm{s}} \Rightarrow 8\ \mathrm{m \cdot s^{-1}}$$

For unequal time units:

$$w_1 = [10\ 20\ 20\ 20\ 10]\ \mathrm{s}$$
$$w_2 = [1\ 1\ 1\ 1\ 1\ ]\ \mathrm{m^{-1}}$$

5. $\sum w_1 = 1 \cdot \mathrm{s} \cdot \sum \mathrm{q} = 80\,\mathrm{s}$ summation by superposition
   $\sum w_2 = 1 \cdot \mathrm{m^{-1}}$ summation by superposition
6. $\bar{\dot{y}} = \frac{70 + 160 + 160 + 180 + 80}{\mathrm{m^{-1}} \cdot 80 \cdot \mathrm{s}} \Rightarrow 8.1\ \mathrm{m \cdot s^{-1}}$

work. This step requires reasoning about the quantity and about whether the summation used in Step 3 changes the scale of the quantity. The concept of juxtaposition versus superposition helps to make this less abstract. As with any calculation, it is important to ask whether the result is reasonable and consistent with expectations formed before making the calculation. There is no substitute for thinking about the quantity and for direct visualization of what happens when the values of a quantity are summed.

Ensemble quantities have a scale that goes into calculating them. This is as much a part of an ensemble quantity as the units of the ensemble. The scale, with both resolution and range, is implied in the summation sign, even though the index used to calculate an ensemble quantity has no units. The increment of the index is associated with a resolution scale, whether resolution by time, length, area, energy content, body mass, or any other basis for weighted summation. The numerical range of the index is $i = 1 \ldots n$. This range can be for time, length, mass, and so on.

Ensemble quantities have a scale, but this scale is not necessarily the same as that of an individual value. Single values of a quantity Q have a scale set by their resolution (minimum resolvable unit = 1U) and their range (q = number of steps). Summation in some cases changes the resolvable unit. Table 9.4 expresses this idea.

In reasoning with scaled quantities, the effects of summing must be considered explicitly. Summing will change the scale in some cases but not others:

If the scale is changed, then: $\sum 1\text{U}^{-1} = n \cdot 1\text{U}^{-1}$
If the scale is unchanged, then: $\sum 1\text{U}^{-1} = 1\text{U}^{-1}$

Here is an example with juxtaposed lengths:

$$\begin{array}{l} \longrightarrow \\ \rightarrow\rightarrow\rightarrow\rightarrow\rightarrow\rightarrow \end{array} \qquad n = \frac{\sum L}{L} = EF$$

The symbol $\sum L$ stands for the total length, composed of five short lengths $\vec{L}$. The length $\sum L$ represents a coarser resolution than each component L. The ratio of $\sum L$ to L is a unitless number, the number of steps. It is also the unitless expansion factor EF (described in more detail in Chapter 12, Section 3).

This example shows how an ensemble quantity scales a series of measurements up from many local values to a single larger-scale value. A convenient term for this is *scaling up by summation*. The average is computed from the sum, so an equivalent term is *scaling up by averaging*.

Table 9.4 Scale of an Ensemble Quantity Compared to Scale of Component Values

A single value of a quantity Q is the product of the number of steps q and a unit 1U:

$Q = q \cdot 1\text{U}$

The sum of a series of values is:

$$\sum Q_i = \sum (q_i \cdot 1\text{U}) = \frac{\sum Q \cdot 1\text{U}^{-1}}{\sum 1\text{U}^{-1}} = \frac{\sum q_i}{\sum 1\text{U}^{-1}}$$

ANOTHER LOOK AT SECTION 9.4

List two quantities of interest to you—one that is summed by superposing values and one that is not.

9.5 The Mean Value of a Quantity

In ecology the most common use of summation is to compute a mean. The mean value of a series of numbers is straightforward—it is the sum of the numbers, divided by the number of observations. The mean value of a scaled quantity requires more thought; it is a weighted sum or total. The formula for the mean value of a quantity is:

$$\text{mean}(Q) \equiv \frac{1}{n}\sum^{n} Q \tag{9.1}$$

As we have seen, care is needed in calculating $\sum Q$ the sum of a scaled quantity. In most applications, equal weighting summation by superposition (see Table 9.3) is used. The units of the mean are the same as the units of the quantity.

The notation used in Table 9.3 includes a bar over the symbol for the quantity, subscripted by another symbol representing the weighting factor. For example, the spatial average in the energy value of prey is $\bar{E}_x$, whereas the average energy value of prey over a period of, say, a year is $\bar{E}_t$. The generic symbols are $\bar{Q}_x$ and $\bar{Q}_t$ for the *spatial and temporal averages*, respectively.

The overbar was used to represent the operation of taking a weighted sum. It is used for both equally and unequally weighted sums. An average, as it is usually calculated, is an equally weighted sum for which w has no units and $\sum w = n$, the number of observations (Box 9.2). For example, the equally weighted bee velocity was calculated at $8\,\text{m s}^{-1}$. The unequally weighted bee velocity was calculated at $8.1\,\text{m s}^{-1}$. Both are averages, and so the symbol $\bar{\dot{y}}_x$ applies to both. The ensemble quantity $\bar{\dot{y}}_x$ is the time rate of change in position in the y direction, averaged over distance in the x direction. Try writing out the names of the following ensemble quantities, then visualizing each.

$$\bar{\dot{y}}_z$$

$$\bar{\dot{y}}_t$$

In statistics, standard notation involves placing a bar over the symbol to signify the average of several observations. This works satisfactorily for numbers, but it will lead to ambiguity in working with scaled quantities; is the average taken over space? Over time? Adding the subscript removes the ambiguity by making it clear how the quantity was averaged.

In geophysics, a different notation is sometimes used to distinguish spatial averages $<Q>$ from temporal averages $\bar{Q}$. This distinction is useful, though it seems unlikely to make much headway in ecology. The overbar for all averages, including the spatial average, is already widely established in ecology and in the statistical literature. The overbar lends itself well to distinguishing the observed average $\bar{Q}_x$ from an average $\bar{Q}(x)$ calculated with a functional expression $Q(x)$ for the spatial average. The overbar reduces the burden on memory because one symbol designates the operation of taking an average. The distinction between spatial and other averages can be made by retaining the appropriate subscript, such as $\bar{Q}_x$ for spatial average.

Ensemble quantities, like the quantities from which they are derived, should be visualized and associated with a vivid mental image that captures the essence of the

quantity. One good way of visualizing a spatial average is to picture a graph of the quantity as a series of lines projecting upward from a plane, much like trees projecting from the forest floor. Then picture the spatial average as an imaginary surface, parallel to the ground, through which some trees project while other trees fall short of the surface suspended above them. The spatial limits on the ensemble quantity should also be held in mind, for the measured average applies only within some limited spatial range. Time averages should also be visualized, with the start and stop points an explicit part of the mental image. For example, average rainfall over a year is pictured as a single fixed value in units of *V/A* (volume per unit area) with observations falling above and below this value from day 1 to day 365. The 20-year average is pictured from day 1 to day 7305. The range is a key part of ensemble quantities and so should be held in mind whenever thinking about an ensemble quantity.

ANOTHER LOOK AT SECTION 9.5

Imagine a computational program that automatically applies the rules presented in Table 9.3. What prompts would the program need to calculate the average of scaled quantities?

9.6 Sums of Quantities on Nonratio Scales

What happens if quantities on nominal or ordinal types of measurement scales are summed? The same principles (Table 9.3) apply. In the case of nominal quantities, the dimensionless numbers q (see Table 9.4) are either 1 or 0, and the sum of a set of values q is:

$$\sum q = 0 \quad \text{if all } q = 0$$
$$\sum q = 1 \quad \text{if any } q = 1$$
$$\text{Also: } \sum 1\text{U}^{-1} = 1\text{U}^{-1}$$

That is, the scale never changes.

Quantities on a nominal scale are often summed to obtain a frequency or proportion. For example, 10 foraging attempts might be scored as 2 successful ones, hence $f = 2$ out of 10 and $p = 20\%$. Once nominal scale quantities are summed in this fashion, they become ratio scale quantities.

Quantities on a rank scale are usually handled by calculating the sum of the ranks $\sum$R, then interpreting the sum of the "units" in either of two different ways. One approach is to treat ranks as a series of observations $\sum 1\text{U}^{-1} = n$. In this case the sum of a set of ranks is not itself a rank. Another approach is to ignore the rank scale entirely, treating the ranks as integers $\sum 1\text{U}^{-1} = 1$. Here is an example. The first and second ranking items are scored as:

$$\sum \text{R} = 1 + 2 \Rightarrow 3$$

The number 3 cannot be interpreted as a rank.

Now an ecological example: three types of coral are competing for space. Species A takes space from B, and both A and B take space from C, but species A and B together may take space from C or lose space to C rather than having the same rank as C. In this example the rank of C relative to (A + B) cannot be calculated from the rank of each item. However, the sum of ranks can be calculated as an integer. Rank-based statistical tests use this number, the sum of ranks, to evaluate the probability of a particular arrangement of ranks.

Defined Concepts and Terms for Review and Future Reference

____ ensemble quantity
____ juxtaposition vs. superposition
____ scaling by summation and averaging
____ spatial average
____ temporal average
____ weighted summation

10 Ensemble Quantities: Variability

It would certainly be a mistake to say that the manipulation of mathematical symbols requires more intellect than [does] original thought in biology.

—R. A. Fisher, *The Genetical Theory of Natural Selection*, 1930

10.1 Synopsis

One of the major contributions of biological thought is that variability is a quantity subject to loss and gain. This view of variability distinguishes biology from the other natural sciences. Genetic variability, for example, is the raw material on which natural selection operates to generate new species. Variability in other ecological quantities is of just as much interest, whether the quantity is density, production rate, or a biogeochemical flux. What factors generate variability in these quantities? What factors reduce variability in these quantities?

Variability is measured by an ensemble quantity called a deviance, which is the weighted sum of the deviations of individual values from an average. Any weighting scheme can be used. One of the most commonly encountered is to weight deviations according to their own magnitude, resulting in a variance var(Q). Weighting the deviations of one quantity by the deviations of another quantity results in the covariance cov(Q,R), which is the basis of correlation and regression. Deviations on a logarithmic scale result in the information statistic G$(Q|\beta)$, which has become increasingly important in statistical analysis over the past three decades. Deviations are interesting quantities in themselves, and taken as an ensemble (variance, covariance, etc.), they become important quantities in ecology.

Deviances are multiscale comparisons. They gauge the difference between local observation and an expected value of greater extent. The temporal variance, for example, gauges the difference between each short-term value and the long-term value of a quantity. In a similar fashion, the spatial variance measures the average difference between local values and a mean value of greater extent.

Variances are often interpretable quantities. For example, the variance in the velocity of a fish measures the kinetic energy, per unit mass, of swimming. The variance, viewed as an interpretable quantity, is often of more interest than the average or summed value of an ecological quantity simply because of the highly variable response of natural populations to heterogeneous environments.

doi: 10.1016/B978-0-12-627865-1.00010-6

The temporal variance measures the vigor of change in a quantity. Much of the early work on temporal variance focused on the search for periodic variation. Variance in ecological time series exists at all scales, and hence rate of increase in variance with increase in temporal scale becomes of more interest than whether the temporal variance reaches a maximum at some scale.

The spatial variance in organism density measures patchiness at the scale of measurement resolution. Early work focused on the search for characteristic scales of patchiness, but it is now evident that spatial variance increases with scale rather than being concentrated at a characteristic scale. The spatial variance in the number of organisms is closely related to mean crowding $M^*(N)$, which measures the potential number of pairwise contacts between organisms at a given spatial scale. The concept of mean crowding is readily extended to the yet more general concept of the potential contact, which measures the opportunity for interaction that depends on direct contact between individuals—predation, parasitism, mutualism, gene exchange, and some forms of competition.

Spatial and temporal variances are not static quantities. They grow and shrink through production and loss. For example, spatial variance in gene frequency is generated by mechanisms that isolate populations; at the same time it is eroded by processes that promote the lateral spread of a gene. One of the major research challenges in ecology is understanding the creation and erosion of spatial variability as a function of spatial scale. This includes the question of the degree to which variance generated at one scale is transformed into variance at another scale.

10.2 Deviations

One of the characteristics of ecological quantities is their high degree of variability. Repeated measurements of the size or speed of lifeless objects vary from the average due to measurement error. Repeated measurements of the size or speed of an organism will vary far more than measurement error. In working with ecological quantities, it is often of as much interest to examine the deviation from the average as it is to examine the average or ensemble value.

A simple way of examining this variability is to compute contrasts or differences between pairs of values, shown in Chapter 8. Another way of examining variability is to examine the contrast between the average value and the individual values that contributed to the average. The difference between a measured value and its larger-scale mean value will be called a *deviation*. Deviations can be expressed on any of the four types of measurement scale, as shown in Box 10.1.

Deviations on a ratio type of scale (Box 10.1) appear repeatedly and so are assigned a unique symbol:

$$\text{dev}(Q) \equiv Q - \text{mean}(Q) \tag{10.1}$$

For the generic quantity Q, the symbol is $\text{dev}(Q)$, which represents the collection of deviations. This is read as "the deviations in the quantity Q from the average," or "the deviations of Q" for short. Of course, some specific quantity would usually be read, not Q. In Box 10.1 the quantity $\text{dev}(\dot{y})$ is read "the deviations in bee velocity $\dot{y}$."

Box 10.1 Deviations Calculated on Four Types of Measurement Scale

The quantity is bee velocity: $\dot{y} = [5\ 6\ 7\ 8\ 9] \cdot \mathrm{m\ s^{-1}}$

The mean value: $\mathrm{mean}(\dot{y}) = 7\ \mathrm{m\ s^{-1}}$

Deviations on a nominal scale: $\dot{y} - \mathrm{mean}(\dot{y}) = [-\ \ -\ +\ +\ +]$

Differences are scored as + or −.
A zero difference has been scored as +.
The pattern of deviations show only that the bee picked up speed.

Deviations on a rank scale: $\dot{y} - \mathrm{mean}(\dot{y}) = [5\ 4\ 3\ 2\ 1]$

The pattern of deviation on this scale shows that speed increases steadily.

Deviations on a ratio scale: $\dot{y} - \mathrm{mean}(\dot{y}) = [-2\ -1\ 0\ 1\ 2] \cdot \mathrm{m\ s^{-1}}$

Deviations on this scale measure the acceleration.

ANOTHER LOOK AT SECTION 10.2

For a quantity of interest to you, state a typical deviation from the mean. What is the ratio of the largest deviation relative to the mean?

10.3 The Total Deviation

One of the important concepts that emerged in statistics in the late 1970s and early 1980s is the distinction between exploratory and confirmatory analysis. *Exploratory analyses* seek pattern; *confirmatory analyses* establish whether a particular outcome is due to chance. Both use deviations but in different ways. The strategy of exploratory data analysis is a cycle of repeated calculation and examination of deviations (Tukey, 1977). The cycle continues until the deviations show no pattern. Good confirmatory analyses also rely on deviations. In a confirmatory analysis the only way to ensure that the assumptions are met is to check for patterns in the deviations of the data from the model (see Chapter 15). Checking the observations rather than the deviations from the model accomplishes nothing; it is voodoo statistics.

The total deviation $\Sigma\mathrm{dev}(Q)$ is an ensemble quantity called the *deviance*. This deviance is zero unless a weighting factor is applied. The weighting factor can be on any type of scale. A weighting factor of historical interest is the sign of the deviation, which is on a nominal type of scale. The result of this weighting (Box 10.2) is the *average deviation*. It is the sum of the absolute values of the deviations:

$$\mathrm{avdev}(Q) \equiv \Sigma\mathrm{dev}(Q) \cdot w_{nominal} = \Sigma|\mathrm{dev}(Q)| \qquad (10.2a)$$

An equivalent expression uses sums rather than the deviations directly:

$$\text{avdev}(Q) = \Sigma Q \cdot w - n^{-1} \cdot \Sigma Q \tag{10.2b}$$

Box 10.2 shows computations according to Equation 10.2b.

Box 10.2 Weighted Deviances (avdev(Q), SS(Q), and G(Q)) Computed from Deviations

N_x is the number of Willets in seven contiguous stretches of beach, each 1 km in length, from Chapter 7.

$$\text{mean}(N_x) = \bar{N}_x = n^{-1} \cdot \Sigma n_x = 8.86 \text{ Willets km}^{-1}$$

N_x	= [	9	4	8	9	31	0	1] · Willets km^{-1}
dev(N_x)	= [	+0	−5	−1	+0	+22	−9	−8] · Willets km^{-1}
$w_{nominal}$	= [	+	−	−	+	+	−	−]
w_{ratio}	= [	+0	−5	−1	+0	+22	−9	−8] Willets·km^{-1}

1. The average deviance, avdev(Q).
   Deviations on a ratio scale.
   Weighting is on a nominal scale. $w_{nominal}$ = + or − deviation.

   $$\text{avdev}(N) = \sum (N \cdot w_{nominal}) - n^{-1} \sum N \Rightarrow 36 - (-8.86) = 45 \text{ Willets·km}^{-1}$$

   $$\text{avdev}(N) = \sum (dev \cdot w_{nominal}) = 44.86 \text{ Willets·km}^{-1}$$

2. The sum of squares, SS(Q).
   Deviations on a ratio scale.
   Weighting on a ratio type of scale. $w_{ratio} = \text{dev}(N)$

   $$\text{SS}(N) = \sum (N \cdot N) - n^{-1} \cdot (\sum N)^2 \Rightarrow 1204 - 549.14 = 655 \text{ Willets}^2 \cdot \text{km}^{-2}$$

3. The G-statistic, $G = 2I$.
   Deviations on a logarithmic scale.
   Weighting on a ratio type of scale. $w_N = N$

   $$\text{lndev}(N) = \ln(N) - \ln(\bar{N}) = \ln(N/\bar{N})$$

   $$\text{lndev}(N) = [\ 0.016\ -0.80\ -0.10\ +0.016\ +1.25\ -0\ -2.2\ ]$$

   $$I = \sum [w_N \cdot \text{lndev}(N)] = \sum [N \cdot \text{lndev}(N)] = 32.94$$

   $$\text{G}(Q|\beta) = \text{G}(Q|8.86) = 2\sum [N \cdot \ln(N/\bar{N})] \Rightarrow 65.9 \text{ Willets}^1 \cdot \text{km}^{-1}$$

The most frequent weighting factor is the deviance itself. That is, $w = \text{dev}(Q)$. This results in a *sum of squared deviations* SS(Q):

$$\Sigma\ \text{dev}(Q) \cdot w = \Sigma\ \text{dev}(Q) \cdot \text{dev}(Q) = \Sigma\ [\text{dev}(Q)]^2 \tag{10.3a}$$

$$\text{SS}(Q) \equiv \Sigma\ [\text{dev}(Q)]^2 \tag{10.3b}$$

This particular deviance is so widely used that it has its own name and symbol: the *sum of squares*, or *SS*. Sums of squares of several quantities are distinguished via functional notation for one quantity SS(Q) versus another SS(Z). An equivalent expression for SS(Q) uses sums rather than the deviations directly:

$$\Sigma\,[\mathrm{dev}(Q)]^2 \;=\; \Sigma Q^2 \;-\; \mathrm{n}^{-1}\;(\Sigma Q)^2 \tag{10.4}$$

Box 10.2 shows computations using Equation 10.4.

The sum of squares *SS* is often an interpretable quantity. The bee velocity deviations (Chapter 9) have units of $\mathrm{m}^1\ \mathrm{s}^{-1}$. So the units of the sum of squared deviations will have units of $\mathrm{m}^2\ \mathrm{s}^{-2}$. These units gauge the kinetic energy kE per unit of bee mass. That is:

$$\text{Kinetic energy} \;=\; kE \;=\; \text{mass}\cdot\text{velocity}^2$$
$$\text{Mass-specific kinetic energy} \;=\; kE/\text{mass} \;=\; \text{velocity}^2$$

In the (Chapter 9) Willet example, the sum of squares *SS* has units of $\#^2\ \mathrm{km}^{-2}$. This is proportional to the number of pairs per unit area, and hence SS(N_x) in this example is a measure of the potential for interactions between Willets within each 1 km segment of the beach. In general, the deviance of a count of organisms will be of interest in any study concerned with locally density-dependent interactions among populations and among species.

Deviations are also of interest on a logarithmic scale. For these deviations the natural logarithm ln = $\log_e$ is always used:

$$\ln[\mathrm{dev}(Q)] \;\equiv\; [\ln(Q) - \ln\bar{Q}] \;=\; \ln(Q/\bar{Q}) \tag{10.5}$$

Each deviation on a log scale ln ($Q/\bar{Q}$) measures the departure of an observation from the mean value $\bar{Q}$. If this ratio is weighted by Q, the result is a measure of the information in each value of Q relative to the information in the mean. When the departure is small, most of the information is in the mean. When the departure is large, most of the information is in the data. The sum of the weighted deviations measures the amount of information in the data, given the mean (Kullback and Leibler, 1951).

$$\mathrm{I}(Q|\beta) \;\equiv\; \sum Q\,\ln(Q/\bar{Q}) \tag{10.6}$$

The symbol I($Q|\beta$) stands for the information in the values of quantity Q, given the parameters β. Box 10.2 shows calculations for the case of the most commonly encountered parameter, the mean value of Q.

If there is no information beyond the mean (all values are equal to $\bar{Q}$), then I($Q|\beta$) = 0. When the information I($Q|\beta$) in the data is small, the likelihood of the estimate $\bar{Q}$ is high. Consequently, minimizing I($Q|\beta$) maximizes the likelihood of the estimate $\bar{Q}$, given the data (Akaike, 1973). Because the principle of maximum likelihood is central to modern statistics, the information measure I($Q|\beta$) is a key component of statistical analysis. Twice the information $G = 2\cdot\mathrm{I}(Q|\beta)$ is Wilks' likelihood ratio statistic (1935), also known as the non-Pearsonian chi-square, or goodness of fit G-statistic (Sokal and Rohlf, 1995). The *G-statistic* G($Q|\beta$) is a weighted deviance that measures the information in the values of Q, given the parameter values β. This measure has become increasingly important in statistics since the work of Nelder and Wedderburn (1972), who demonstrated its use in applying maximum likelihood when deviations are not homogeneous, and ratios of sums of squares cannot be used to compute probabilities as

in classical methods based on the sum of squares SS(Q). McCullagh and Nelder (1989) present a comprehensive list of deviances and rescalings used in statistical analysis.

ANOTHER LOOK AT SECTION 10.3

For a given set of data, what happens to information I remaining in the data as the estimate of $\bar{Q}$ moves from a poor value (low likelihood) to a good value (maximum likelihood)? What happens to SS(Q) and avdev(Q) as I($Q|\beta$) decreases?

10.4 The Variance

The *weighted deviances* in Box 10.1 are the foundation for the measurement of variability. Most measures of variability in ecology are built on the SS(Q), the sum of the squared deviations. Table 10.1 defines five common measures, including the *variance* var(Q). The simplest is the mean value of the squared deviations, msd(Q). The mean squared deviation of the bee velocities comes to msd$(\dot{y})$ $=>$ $(2^2 + 1^2 + 0^2 + (-1)^2 + (-2)^2)/5 = 2 \cdot \text{m}^2\ \text{s}^{-2}$. The mean squared deviation of the Willet counts comes to msd(N_x) $=>$ $654.86/7 = 93$ Willets^2. To obtain a measure in the same units as the variable, the root mean squared deviation is often used (Table 10.1). In biology it is standard practice to use an estimate of the true variance rather than the observed variance. The conventional symbols are σ^2_Q for the true variance and s^2_Q for the estimate from the sample. This notation, which reduces a symbol to a subscript, becomes ugly for quantities that are already subscripted. This notation is to some degree misleading, since we may be interested in the variance, not in the parameter σ of the normal distribution. A more satisfactory convention is to use a "hat" over the Greek symbol to distinguish

Table 10.1 Commonly Encountered Measures of Variability

The mean squared deviation	$\text{msd}(Q) \equiv \frac{1}{n}\sum[\text{dev}(Q)]^2$
	$\text{msd}(Q) \equiv n^{-1} \cdot \text{SS}(Q)$
The root mean squared deviation	$\text{rmsd}(Q) \equiv \sqrt{\text{msd}(Q)}$
The variance	$\text{var}(Q) \equiv \frac{n}{n-1}\ \text{msd}(Q)$
	$\text{var}(Q) \equiv \frac{1}{n-1}\text{SS}(Q)$
The coefficient of dispersion	$\text{CD}(Q) \equiv \frac{\text{var}(Q)}{\text{mean}(Q)}$
The standard deviation	$\text{std}(Q) \equiv \sqrt{\frac{SS(Q)}{n-1}}$
The coefficient of variation	$\text{cv}(Q) \equiv \frac{\text{std}(Q)}{\text{mean}(Q)}$

the parameter σ from its estimate $\hat{\sigma}$, then note that for the normal distribution the estimate of σ^2 is var(Q), as in Table 10.1.

$$\hat{\sigma}^2 = \text{var}(Q) \tag{10.7}$$

In the example of the Willet data (Box 10.2) the estimate is var(N_x) => 654.86/6 = 109 Willets^2. This estimate uses $n - 1$ rather than n to adjust for the fact that the population will almost certainly contain several deviations that exceed those encountered in the sample. Hence the observed mean squared deviation msd(Q) is scaled up to a slightly larger value according to a factor $n/(n - 1)$. In the example of bee velocities, the mean squared deviation is msd(y) = (2 m^2 s^{-2}), whereas the estimate of the true variance is var($\dot{y}$) => 2 m^2 s^{-2}(5/4) =2.5 m^2 s^{-2}.

The variance has units that are the square of the quantity. To obtain a measure in the same units, the square root is taken. Taking the square root of the estimate of the variance yields an estimate of the standard deviation std(Q), shown in Table 10.1. To permit comparison of variability among series with different units, the standard deviation is normalized to the mean value. This results in a dimensionless ratio called the *coefficient of variation*. Table 10.1 shows the formula for the estimate of the true value of the coefficient of variation. The coefficient of variation is often viewed as a fraction and hence multiplied by 100 to express it as a percentage.

These measures of variability, like the mean value, represent *multiscale comparisons*. They all entail a form of "zoom" rescaling, whereby each deviation (representing a limited area and time) is compared to a single value representing a larger area or time. The local scale is set by the resolution of the observations at hand; the larger scale is set by the range of the observations. An example is the temporal coefficient of variance cv(Q_t), a single number that gauges the explosiveness of a rate of change across the temporal range of all the measurements.

Estimates var(Q) of the population variance σ^2 imply scale-up from the spatial range of the sample to the range of the population. Scale-up depends on whether the sample comes from a finite or infinite population. In many ecological applications, samples are defined relative to an infinite population of all possible measurements that could be generated by the procedural statement (refer back to Chapter 3). Another definition is that the sample is taken randomly from a defined set of enumerable units called a *frame*. If the frame and units are defined spatially, the inference from sample to population entails spatial scale-up based on the ratio of area covered by the sample to area of the frame. Scale-up, as a ratio, is undefined for an infinite population.

The variances of ecological quantities are sometimes interpretable quantities. An example is the variance in the velocity v of an organism:

$$\frac{\text{kinetic energy}}{\text{mass}} = \frac{mv^2}{m} = v^2 \tag{10.8}$$

This measures kinetic energy per unit mass of an organism. Another example is the variance in counts of organisms, which has units of entity · entity. As we saw in Chapter 5 (Equation 5.1), this can be interpreted as the potential number of pairs or the potential number of pairwise combinations in a group. In Box 10.2, Willet^2 gauges the number of potential competitors. The unit entity^2 looks strange, but it usefully measures the potential for interaction between the objects being counted.

ANOTHER LOOK AT SECTION 10.4

For a quantity of interest to you, state units of the mean, the variance, and the information statistic G (Section 10.3). Is the variance an interpretable quantity? Is the information statistic G an interpretable quantity?

10.5 Scale-Dependent Variance

Measurements referenced to time and location vary according to temporal and spatial scale. Recall from Section 7.5.1 that the symbol $Q_{x\,t}$ stands for any geographically and temporally referenced quantity with temporal and spatial attributes of resolution (vectors $\boldsymbol{h}$, $\boldsymbol{i}$, $\boldsymbol{j}$, and $\boldsymbol{k}$) and position in time ($t\boldsymbol{h}$) and space ($x\boldsymbol{i}$, $y\boldsymbol{j}$, and $z\boldsymbol{k}$). We can vary the scale either by altering the separation between units or by altering the spatial and temporal extent of the unit (altering vectors $\boldsymbol{h}$, $\boldsymbol{i}$, $\boldsymbol{j}$, and $\boldsymbol{k}$). We can alter the extent by either changing the measurement frequency (equivalently, altering the bin size over which measurements are averaged) or accumulating units in sequence (e.g., average over 10 sequential measurements, then add next measurement and recompute the average, etc.). These three operations are commonly called *lagging*, *coarse graining*, and *accumulation*. Iterative calculation of the variance according to each of these three operations displays the change in variability with change in scale.

Table 10.2 shows the commonly used measures of scale-dependent variability. In this table the symbol $Q_{x\,t}$ has been abbreviated to Q. The table also shows the relation of these ensemble quantities one to another. Most of these measures were developed for the analysis of patchiness and spatial variance in density of organisms, but they can be applied with any quantity referenced to location, to time, or both.

The commonly used measures of scale-dependent variability are the autocorrelation (as a function of lag), the semivariance (as a function of lag), the variance *MSA* (as a function of group or block size i), and spectral density (as a function of measurement frequency in either space or time). The *autocorrelation* $\mathrm{acf}_k(Q)$ is the correlation of a variable with itself at lags of 1, 2, and so on. Plots of the autocorrelation versus lag are called *correlograms*. Sokal (1979) discusses the use and interpretation of correlograms in ecology. Another measure of autocorrelation is the *semivariance* $\gamma_k(Q)$. The semivariance is some part of the overall variance; it is the variance at any one lag. The semivariance is the same as the variance if all observations are assigned to the same lag. Thus the semivariance of the bee data is 2.5 $\mathrm{m}^2\ \mathrm{s}^{-2}$ (all five measurements at the same lag). If all seven measurements of Willet density are assigned to the same lag, the semivariance as calculated from the formula in Table 10.2 is 109 Willet^2.

The semivariance, once calculated over more than one lag, is then plotted against lag to produce a *semivariogram*. When data are autocorrelated in space or time, the semivariance is less at small lags than large ones. The rate of increase in the semivariance with increase in lag quantifies the spatial autocorrelation. This autocorrelative information can be displayed as a function of measurement frequency rather than lag by computing either the *mean square variance among groups* ($\mathrm{MSA}_i(Q)$ in Table 10.2) or the *spectral density* (variance per unit frequency; see Table 10.2). The plot of mean square variance per unit block against block size has no name and so for convenience might called a *Greig-Smith plot*. The plot of spectral density versus frequency is called a *periodogram*.

The next two sections compare these measures, using specific examples in either space or time.

Table 10.2 Measures of Variability as a Function of Scale

Definitions	
$\boldsymbol{i}$	Unit of length, area, or volume
x	Location, 1 to x
t	Time, 1 to t
$\boldsymbol{x}_{max} = \boldsymbol{i}\,x$	Spatial range
$\boldsymbol{t}_{max} = \boldsymbol{i}\,t$	Temporal range
$n = \boldsymbol{x}_{max}/\boldsymbol{i}$	Number of spatial units
$n = \boldsymbol{t}_{max}/\boldsymbol{i}$	Number of temporal units
$f = 1/2\boldsymbol{i}$	Measurement frequency
$k = 1 \ldots n$	Lag (separation) from 1 to n
$\sum^{\boldsymbol{i}} Q$	Total, within unit $\boldsymbol{i}$
$(\sum^{n} \sum^{\boldsymbol{i}} Q)$	Total over all units = $\sum\sum Q$
Measures	
mean$(\sum^{\boldsymbol{i}} Q)$	Spatial mean per unit $\boldsymbol{i}$
var$(\sum^{\boldsymbol{i}} Q)$	Spatial variance per unit $\boldsymbol{i}$
CD$(\sum^{\boldsymbol{i}} Q)$	Coefficient of dispersion
$\text{MSA}_{\boldsymbol{i}}(Q)$	Mean squared deviation among groups of size $\boldsymbol{i}$
SpD(Q)	Spatial spectral density
$\gamma_k(Q)$	Semivariance at lag k
$\text{acf}_k(Q)$	Autocorrelation at lag k
Relations	
mean$(\sum^{\boldsymbol{i}} Q)$	$= n^{-1} \sum\sum Q$
CD$(\sum^{\boldsymbol{i}} Q)$	$=$ var$(\sum^{\boldsymbol{i}} Q)$/mean$(\sum^{\boldsymbol{i}} Q)$
$\text{MSA}_{\boldsymbol{i}}(Q)$	$=$ var$(\sum^{\boldsymbol{i}} Q)/\boldsymbol{i} =$ CD$(\sum^{\boldsymbol{i}} Q) \cdot \sum\sum Q/\boldsymbol{i}$
SpD(Q)	$= f^{-1}\,\text{MSA}_{\boldsymbol{i}}(Q) = 2\boldsymbol{i} \cdot \text{MSA}_{\boldsymbol{i}}(Q) = 2 \cdot$ var$(\sum^{\boldsymbol{i}} Q)$
cov(Q_x, Q_{x+k})	$=$ var$(Q_x + Q_{x+k}) -$ var$(Q_x) -$ var(Q_{x+k})
$\text{acf}_k(Q)$	$=$ cov(Q_x, Q_{x+k})/(var(Q_x) + var(Q_{x+k}))

ANOTHER LOOK AT SECTION 10.5

Draw 12 contiguous boxes in a line, assign the value of a measured quantity to each box, assign a symbol to the quantity, and then state whether the 12 boxes represent a time series or data from a transect. Then redraw the 12 boxes, connecting all pairwise neighbors with an arc (lag 1 pairs). Redraw the 12 boxes again, connecting all lag 2 pairs. Redraw the diagram showing 6 boxes, each with a new value, the sum of the 2 neighboring values assigned to the box. Redraw the diagram showing 4 boxes, each with a new value, the sum of the 3 neighboring values assigned to the box.

10.6 The Temporal Variance

The temporal variance measures the vigor of change in a quantity. If the quantity of interest is number of moth pupae N in 60 successive midwinter counts in a managed forest from 1880 to 1940 (Varley, 1949), the variance measures the degree to which each value deviates from the long-term or average value. The greater the variance, the more

rapid the change from year to year. The four populations reported by Varley (1949) all had large temporal variances, reflecting changes among years on the order of 10^4 or more. Warner et al. (1995) found that the coefficient of variation (standard deviation scaled to the mean) was higher for population time series than for abiotic time series.

Temporal variance is generated by periodic and episodic processes. For example, fluctuations in numbers of snowshoe hares (*Lepus americanus*) produce periodic fluctuations in numbers of lynx predators (*Lynx canadensis*) at time scales of 10 years (Elton, 1924), as shown in Figure 10.1a. In Peru, El Niño events generate episodic variability in numbers

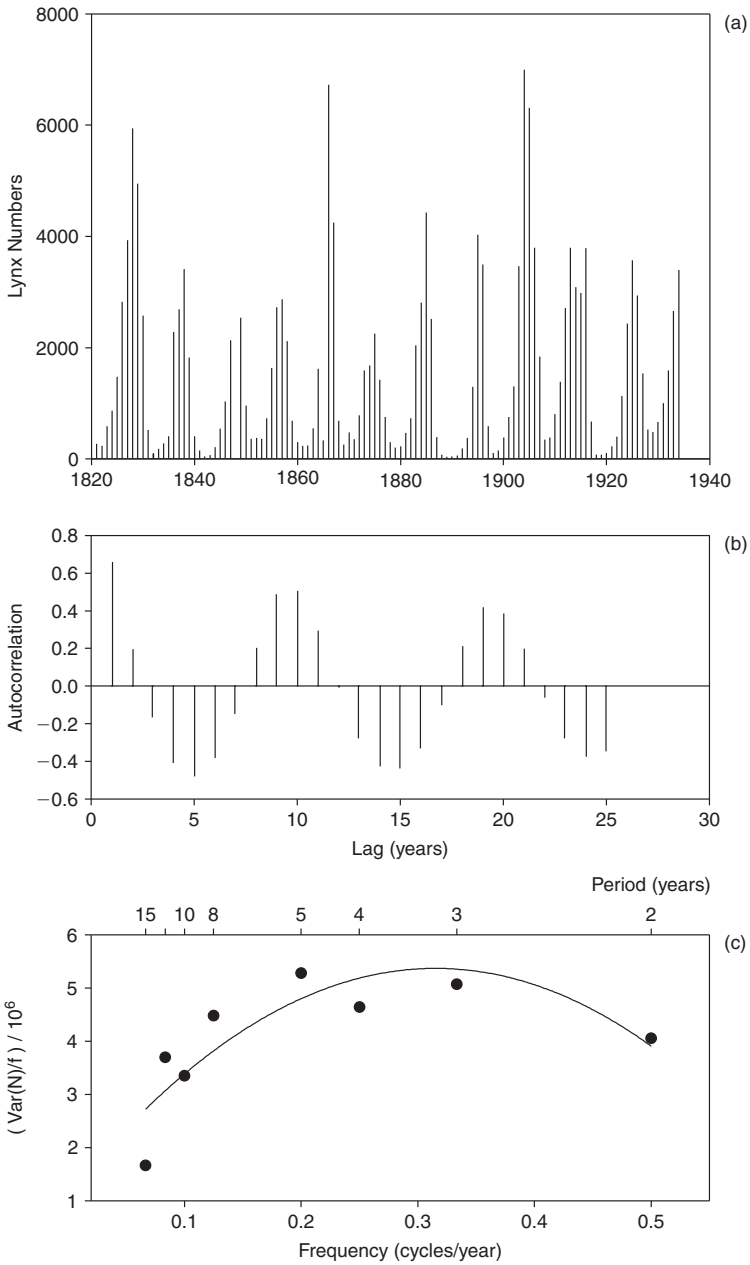


FIGURE 10.1 Temporal Variance in Lynx Numbers. Data from Andrews and Herzberg (1985); (a) Annual number trapped; (b) Autocorrelation in lynx numbers; (c) Change in variance with time scale, as measured by a periodogram.

of nesting guano birds and hence in annual rates of guano deposition (Figure 10.2a) at the time scale of two to five years. These two time series differ in that variation in the lynx series is confined to a characteristic scale of 10 years, whereas variation in the guano series is smeared across a variety of scales because of the irregularly episodic nature of the El Niño events that drive nesting success and hence guano deposition.

There exist a variety of *measures of temporal variance*, most of which are deviances (Table 10.1, 10.2). Temporal variance is commonly quantified as an autocorrelation. Figures 10.1b and 10.2b shows the autocorrelation (Table 10.2) in the lynx and guano time series. The autocorrelative structure in both series is typical of ecological data in that autocorrelation is positive at the small lags because adjacent measurement tend to be similar (high values near other high values, low near other low). The structure of both series is typical in decaying toward zero with increasing separation between measurements. The zero crossing is taken as the time scale at which a quantity is independent of previous dynamics. If there is periodicity in the series, the autocorrelation will drop to substantial negative values at lags on the order of half the dominant period, then rise to positive values at the lag corresponding to the period (10 years in Figure 10.1b). If variability is smeared across multiple scales, the autocorrelation will show little fluctuation after crossing zero, as in Figure 10.2b. There may be substantial variability at these larger separations in time, but this variability will not be evident, because it is distributed over a large number of lags.

Box 10.3 Calculation of Variance as a Function of Measurement Frequency

Values in lower part of table are means per period at periods of 2, 3, and 4. Variance computed from means within a column is shown at top of column.

		2	3	4	1/*f*	Period
		0.5	0.333	0.25	*f*	Frequency
Year	Lynx	3.86	4.09	2.33	var*(N)*	Variance
	(Thousands)	7.71	12.28	9.32	var*(N)*/*f*	Spectral density
1821	0.269	0.25	0.36	0.49		
1822	0.231					
1823	0.585	0.73				
1824	0.871		1.72			
1825	1.475	2.15		3.54		
1826	2.821					
1827	3.928	4.94	4.94			
1828	5.943					
1829	4.950	3.76		2.04		
1830	2.577		1.07			
1831	0.523	0.31				
1832	0.098					

Temporal variance is also quantified relative to the frequency of measurement in the series rather than the lag between values. The measurement frequency of a series ranges from 0.5 cycles per observation (small scale at 2 observations per cycle) to fewer cycles per observation (e.g., 0.1 cycles per observation at the larger scale of 10 observations per cycle). Box 10.3 shows the calculations for an abbreviated series, the first 12 years of the lynx data.

Figure 10.1c shows a periodogram for the spectral density (variance per unit frequency) in the lynx data. There is strong periodicity at the scale of a decade, hence a peak in variance at 10year/2 = 5 years. Enright (1965) showed that periodogram analysis is able to detect strong periodicity at a single scale but cannot reliably detect multiple periodicity at multiple scales. Periodograms have a number of known problems, including dependence on starting point in the series and inaccurate estimate of the frequency of peak variance. Spectral analysis (Platt and Denman, 1975) is a sophisticated form of periodogram analysis, one that addresses many of the known problems with periodogram analysis. It is thus better than periodogram analysis at detecting and estimating periodicity in time series. Spectral analysis requires time series that are of adequate length, regularly spaced in time, and free of zero values. Legendre and Legendre (1998) cover recently developed methods that address these problems.

Time series in ecology usually have no strong periodic component (Warner et al., 1995). Instead the variance increases in a regular fashion, from high frequencies (short periods) to lower frequencies (longer periods). Figure 10.2c shows a periodogram for the spectral density (variance per unit frequency) in the guano data. There is no strong periodicity. The variability increases from the right side of the figure (high frequency, small spatial scales) to the left side of the graph (low frequency, large spatial scales). The increase in variance often scales with frequency according to a power law. Here is the measurement relation for the power law relating spectral density to frequency:

$$\frac{\text{var}(Q_{long})/f_{low}}{\text{var}(Q_{short})/f_{high}} = \left(\frac{f_{low}}{f_{high}}\right)^{\beta} \tag{10.9a}$$

To illustrate the concept, variances and measurement frequencies from Figure 10.2c are substituted into this expression:

$$\frac{4.5 \times 10^9/0.5}{3.8 \times 10^9/0.1} = \left(\frac{0.5}{0.1}\right)^{\beta} \tag{10.9b}$$

$$\hat{\beta} => \left(\frac{\ln(4.5/3.8) - \ln(5)}{\ln(5)}\right) = -0.894 \tag{10.9c}$$

Equations 10.9b and 10.9c illustrate the procedure, but the estimate of β is highly uncertain because only two values were used in the estimate. As estimated by regression (Figure 10.1d), the variance per unit frequency scales with frequency according to an exponent of $\hat{\beta} = -0.901$. The negative exponent indicates that variance grows as the time scale increases and as frequency of measurement decreases. In the guano data (Figure 10.2a), variance is generated at the two- to five-year scale by drops in production following El Niño events. At larger time scales, human activity generated an

additional component of variance by excluding predators from coastal nesting sites, producing a long-term increase in guano production after 1940 (Figure 10.2a). New sources of variance are uncovered as the time scale increases (larger period, lower frequency of measurement). The variance grows from small scales (right side of Figure 10.1c) to large

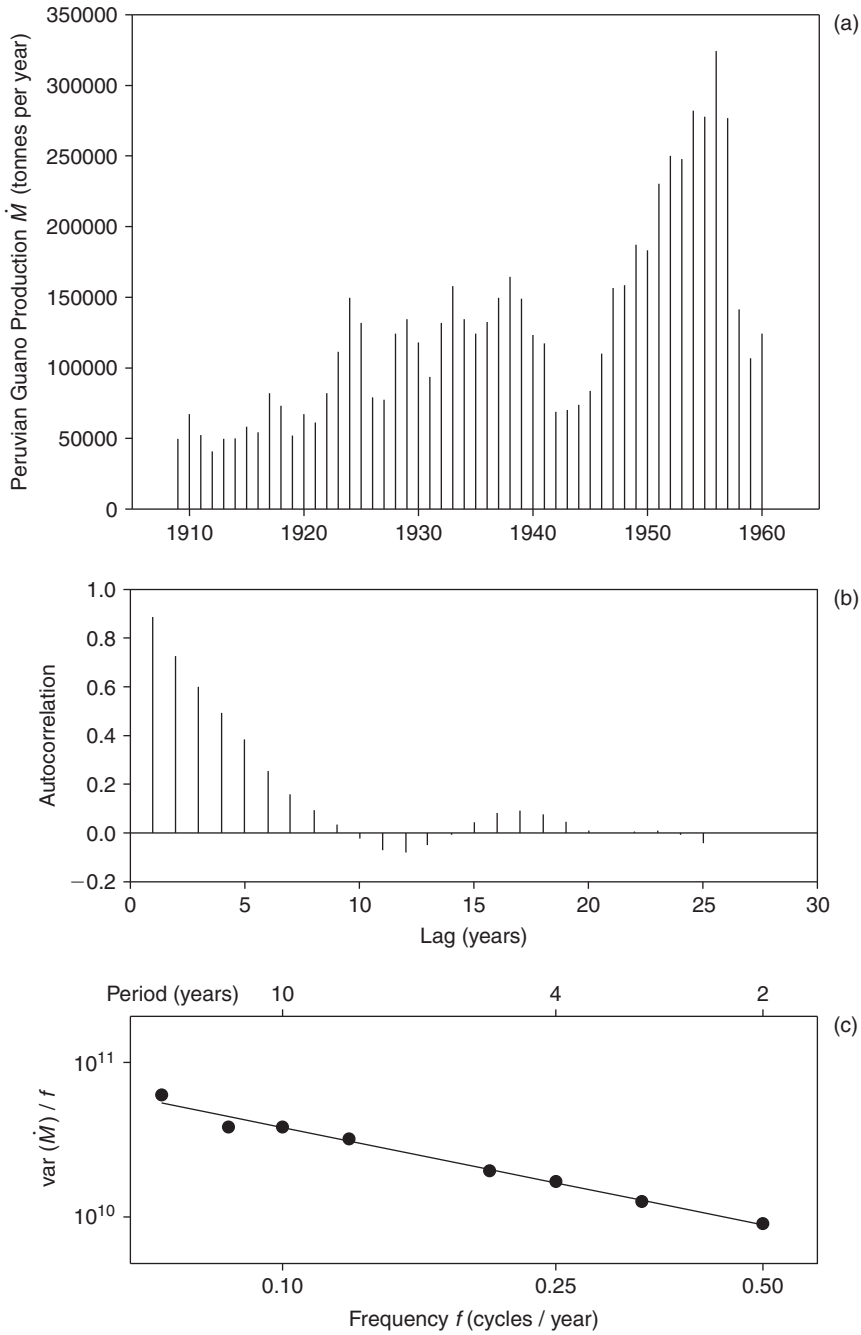


FIGURE 10.2 Temporal Variance in Guano Production from Peru. Data from Schneider and Duffy (1988). (a) Annual production. (b) Autocorrelation in guano production. (c) Growth in variance with time scale, as measured by a periodogram and by spectral density.

scales (left side of Figure 10.1c). The exponent of the spectral density plot measures the rate of increase.

Early work on temporal variance in ecology was devoted to the search for periodic variation (e.g., Enright, 1965). More recently, an analysis of the available long-term series (circa 30) showed that temporal variance usually does not peak at a particular scale and hence is not periodic (Warner et al., 1995). Instead, variance exists at all scales (Powell and Steele, 1995). The question becomes: At what rate is variance added with increase in temporal scale? rather than: At which scale is temporal variance maximum? Because new sources of variance are added at longer time scales, it is important to place any ecological study (necessarily of limited duration) into the context of long-term changes (Magnuson, 1990; Golley, 1993; McClaran et al., 1995; McDowell et al., 1995; Stohlgren, 1995; Venrick, 1995). One of the central challenges in ecology is developing an understanding of temporal variance in population and ecosystem variables as a function of temporal scale (Powell, 1989; Steele, 1991b).

ANOTHER LOOK AT SECTION 10.6

Name 10 variables of interest to you. Of these, which show periodic variance?

10.7 The Spatial Variance

The spatial variance is another important ensemble quantity in ecology. This quantity is at least as interesting and far more informative than the spatial average. The spatial variance is visualized as the spatial "roughness" or "graininess" of a quantity. If the quantity is vegetation height z, the mean or average value mean(z) represents a plane floating above the earth's surface; the spatial variance is the degree to which local values deviate above or below this idealized surface (the average height) above the forest floor. The greater the roughness of the vegetation canopy (easily visualized), the greater the variance in its height. Spatial variance is not a static quantity. Rather, it is generated and eroded by antagonistic processes. Variance in canopy height, for example, is generated by windthrows that tear holes in the canopy. Trees growing into the gap then reduce the variance.

There exist a variety of *measures of spatial variance*, most of which are deviances (Tables 10.1 and 10.2). Table 10.2 displays commonly used measures of spatial variance and the relation of one measure to another. All these measures quantify the strength of spatial gradients relative to the spatial mean. The stronger the gradients, the greater the variance.

Figure 10.3 displays the behavior of four different measures of spatial variance (the autocorrelation, the semivariance, the cumulative variance, and the variance per unit frequency) in two contrasting spatial series, counts of the seabird *Uria aalge* and counts of its prey, a schooling fish called the capelin (*Mallotus villosus*). The salient differences between the series are that there are several clusters of capelin and one broken cluster of murres, which is displaced approximately 2 km further along the transect than the largest cluster of fish (Figure 10.3a). This pattern is typical of seabirds in relation to

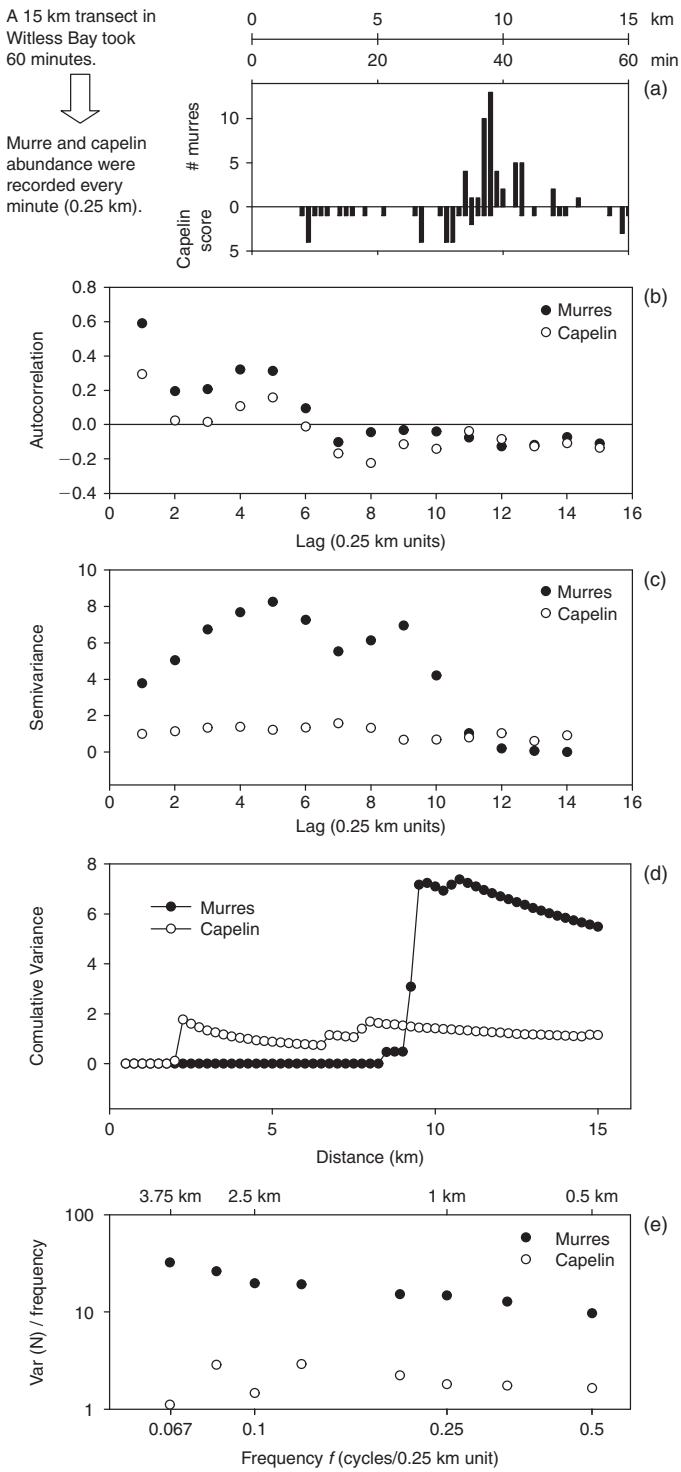


FIGURE 10.3 Spatial Variance in a Seabird, the Murre *Uria aalge*, in Relation to Spatial Variance in its Highly Mobile Prey, the Capelin *Mallotus villosus*. Data from Schneider and Piatt (1986); (a) Distribution along a single transect; (b) Change in variance of murres and capelin with change in scale, as measured by autocorrelation; (c) Change in variance with scale as measured by semivariance; (d) Change in variance with scale as measured by cumulative variance; (e) Change variance with scale as measured by spectral density.

highly mobile prey as birds converge on a school, feed intensely, and then the two clusters begin to drift apart until the birds reaggregate over a school (Hoffman et al., 1981).

Patchiness as a function of scale is often quantified via spatial autocorrelation of densities N/A. In one dimension a simple form of analysis is to plot the autocorrelation (Table 10.2) versus the lag k, as in Figure 10.3b. This displays the degree to which densities are correlated if they are neighbors (lag $k = 1$ step), neighbors at one remove (lag $k = 2$ steps), and so on. As with temporal autocorrelation, the association is typically high for near neighbors, then declines with distance.

Clustering creates the positive autocorrelation at short lags, evident in both birds and fish in Figure 10.3b. In both series, autocorrelation crosses zero just beyond a lag of 6, and thus the radius of the average patch size along the transect is $0.25\,\text{km} \times 6 = 1.5\,\text{km}$. The average diameter is 3 km, which corresponds roughly with the visual impression from Figures 10.3a. The average diameter along a transect is an estimate of chord length (not diameter) because the transect will usually pass through part of the patch rather than directly through the center of the patch. The autocorrelation does not rise again to positive values (as with the lynx data in Figure 10.1), and thus there is no indication of periodicity in either spatial series.

The semivariance (Figure 10.3c), as a descriptive statistic, yields an interpretation similar to the autocorrelation. The semivariance increases up to lags 5 (murres) and 7 (capelin), corresponding to an average half patch size along the transect of 1.25 km (murres) to 1.75 km (fish), or chord lengths of 2.5 km and 3.5 km. The semivariance drops to nearly zero at lags on the order of the chord length; thus there is no indication of the periodicity or nested clustering (patches within patches) that keep the semivariance at the level to which it initially rose. A geostastical model (Journel and Huijbregts, 1978; Davis, 1986) of the semivariance as a nugget (estimate of semivariance at zero lag), sill (estimate of level to which semivariance rises), and range (lags over which the rise from nugget to sill occurs) would clearly not be appropriate for these data.

The cumulative variance of both murres and capelin (Figure 10.3d) shows the characteristic behavior of highly clustered counts. The variance jumps suddenly, then slowly sinks until it jumps again at the next cluster. One such jump occurs in the murre series, two jumps occur in the capelin series. In highly aggregated species this behavior continues as the series grows in length, and the variance never converges on a single value, as it does for spatially independent data. The growth in variance with increase in scale is also evident in the plot of the variance scaled to the frequency of measurement (Figure 10.3e). In murres, this measure of variability rises from the smallest scale of 0.5 km (highest frequency of 0.5 cycles per unit) to the largest scale of 3.75 km (lowest frequency of 0.067 cycles per unit). The spectral density of murres, estimated crudely via the periodogram in Figure 10.3e, rises with frequency according to a power law, as would be expected in a spatial series with a single cluster, no periodicity, and no clusters within clusters. A good fit to a power law also results from clusters within clusters, so the only generalization possible is that periodicity produces periodograms that rise and fall (as in Figure 10.1), whereas other forms of variability produce periodograms fit by power laws. The spectral density as estimated by the periodogram in capelin (Figure 10.3e) rises for small scales (0.5 km) to large (3.75 km), with some indication of a drop at larger scales. However, the estimates at the largest scales in Figure 10.3e are uncertain, since they are the result of variance calculated across a small number of averages. Given the biology of schooling capelin, the suggestion of periodicity in Figures 10.3a and 10.3e is more a matter of happenstance than anything else.

ANOTHER LOOK AT SECTION 10.7

Of the four common measures of spatial variability (autocorrelation, semivariance, spectral density, and variance per unit block), how many have you encountered?

10.7.1 Patchiness

The analysis of patchiness is one of oldest topics in ecology. Definitions of patchiness are numerous, but most are based on spatial variance in density var(N/A) or a related measure (Table 10.2). Early work (Fisher et al., 1922; Clapham, 1936; Blackman, 1942) used the *coefficient of dispersion* (Table 10.2), for which the expected value is 1 for organisms that are randomly distributed. This ratio thus came with a convenient landmark for characterizing patchiness as uniform (CD(N) < 1), random(CD(N) = 1), or clumped (CD(N) > 1). Dependence of this measure on quadrat size was recognized early, leading Morisita (1954, 1959) to propose an index independent of patch extent. Morisita's I_β index is independent of quadrat size as long as the spatial unit (quadrat) is less than the patch extent and organisms are randomly distributed within patches (Patil and Stiteler, 1974).

An interesting quantity related to the coefficient of dispersion is Lloyd's (1967) measure of *mean crowding*. Lloyd wanted a measure of social interaction within areas determined by the ambit of an animal. Each of N animals in an ambit potentially encounters $N - 1$ other animals. Mean crowding M^* is the number of potential encounters ($N_x - 1$) in an ambit, weighted by the number in the ambit (N_x), averaged over a series of ambits. The formal expression of this idea is:

$$M^*(N_x) \equiv \frac{\Sigma N_x(N_x - 1)}{\Sigma N_x} \tag{10.10}$$

From examination, Lloyd's measure of the average number of animals encountered has units of $N^2/N = N$.

Mean crowding can be calculated directly from the mean density and observed mean squared deviation in density:

$$M^*(N_x) \equiv \frac{\text{msd}(N_x)}{\text{mean}(N_x)} + \text{mean}(N_x) + 1 \tag{10.11}$$

Mean crowding is usually calculated from the estimated variance var(N), even though the biological definition indicates that the mean squared deviation should be used. All four terms in Expression 10.11 have the same units (an example of the apple/orange principle from Box 4.2). Hence mean crowding must have the same units as the mean, whether entities, entities per unit area, or entities per unit volume.

Mean crowding is not a static quantity. It is expected to increase as animals converge into limited areas, raising the potential for interaction. It is expected to decrease as animals diverge away from crowded areas, thus becoming more evenly spaced. Territorial behavior is one mechanism that reduces mean crowding, at least at the spatial resolution scale of defensible areas.

The time rate of change in mean crowding $\dot{M}^*(N_x)$ is linked to the time rate of change in density $[\dot{N}]$ via a coefficient that relates crowding to density:

$$\dot{M}^*(N_x) = \frac{\Delta M^*(N_x)}{\Delta[N]} \cdot [\dot{N}] \tag{10.12}$$

Estimates of this coefficient and how it varies with spatial scale would be useful in calculating the effects of change in population size on contact rate and social interaction.

Because mean crowding is equivalent to Morisita's index, Lloyd (1967) concluded that mean crowding was independent of unit size, and hence unit size could be set equal to the ambit or spatial extent of activity of the study organism. Iwao (1968) regressed mean crowding on mean density, then interpreted the intercept as an index of aggregation size. This method, once widely used (e.g., Byerly et al., 1978; Gutierrez et al., 1980), relies on the same assumption as Morisita's index.

Taylor (1961) proposed a generalization of the coefficient of dispersion—a power law relation between the variance and the mean:

$$\text{var}(N/A) = k \cdot [\text{mean}(N/A)]^{\beta} \tag{10.13}$$

If organisms are distributed randomly, the ratio of the variance to the mean is fixed and $\beta = 1$. The fit of *Taylor's power law* to data has been checked with a variety of taxa (Taylor et al., 1978). Many processes can generate a power law relation between the mean and the variance (Taylor and Taylor, 1977; Anderson et al., 1982; Kilpatrick and Ives, 2003). Perry and Woiwood (1992) showed that parameter estimates are best accomplished via nonnormal error structures within the framework of the generalized linear model (McCullagh and Nelder, 1989).

The mean (a first-moment statistic) and the variance (a second-moment statistic) are at best summaries of the information contained in a full-frequency distribution of data. This has led to the use of frequency distributions to characterize patchiness, rather than relying on just the spatial variance. The negative binomial distribution was one of the first theoretical distributions in biology (Student, 1907; Greenwood and Yule, 1920) and remains widely used in ecology (Anscombe, 1948; Taylor, 1953; Houser and Dunn, 1967; Hassell, 1978). The use of frequency distributions to characterize spatial variability regularly attracts review articles (Rogers, 1974; Greig-Smith, 1983; Horne and Schneider, 1995).

Quantifying patchiness as a function of scale is another topic with a long history in ecology. It can be traced back to agricultural uniformity trials, where plot sizes were systematically varied to isolate the "best" scale at which to conduct experiments (Mercer and Hall, 1911). Subsequent work aimed at empirical relations between plot size and variability among plots (Smith, 1938; Bliss, 1941). Pattern analysis (Greig-Smith, 1952) evaluates spatial variance as a function of spatial scale using a method that is equivalent to periodogram analysis. Organisms are counted in a grid or along a transect of contiguous quadrats, and counts in neighboring quadrats are combined repeatedly into blocks of 4 (2 for transects). This maneuver decreases the spatial resolution by a factor of 2^2 at each step. The mean squared variance among blocks (Table 10.2) is then computed at each step and plotted against block size. If the distribution is random at all scales, the variance among blocks (equivalent to the coefficient of dispersion) remains constant at a value of 1. A patch will produce a peak in the variance at roughly the scale of the patch.

Kershaw (1957) extended this method by using linear transects of rectangular, contiguous quadrats. Pattern analysis is thus equivalent to periodogram analysis (Table 10.2) and thus a crude form of spectral analysis.

The measurement relation for *pattern analysis* is the coefficient of dispersion scaled to block size i. Notation follows Table 10.2.

$$\frac{\mathrm{CD}(Q_{large})}{\mathrm{CD}(Q_{small})} = \left(\frac{i_{large}}{i_{small}}\right)^{\beta} \tag{10.14}$$

Because of the block size is the inverse of measurement frequency (Table 10.2) and because the spectral density can be computed from the CD using the relations in Table 10.2, this measurement relation is equivalent to that for spectral density. For single transects, the relation of the CD to spatial scale often does not follow a simple power law, as in Equation 10.14. However, when multiple transects are averaged, power laws tend to appear.

Pattern analysis has been used to quantify patchiness of a wide variety of plants (Greig-Smith, 1983), benthic invertebrates (Grassle et al., 1975; Schneider et al., 1987), marine birds (Schneider and Duffy, 1985), and fish (Schneider and Piatt, 1986; Piatt, 1990). In marine biology and oceanography, spectral methods are used to detect pattern and estimate change in variance with change in measurement frequency (resolution). Variance decreases with increasing measurement frequency (hence resolution) in phytoplankton (Platt and Denman, 1975), krill (Weber et al., 1986; Levin et al., 1989), capelin (Schneider, 1989; Rose and Leggett, 1990; Schneider, 1994b; Horne and Schneider, 1997), groundfish (Horwood and Cushing, 1978; Rose and Leggett, 1990; Horne and Schneider, 1997), and birds (Schneider, 1990; Logerwell et al., 1998). For passively drifting organisms, the rate of decrease is similar to physical factors such as temperature. For mobile organisms, the rate of decrease is flatter ($0 > \beta > -1$) than expected from turbulent mixing ($\beta = -2$) or random particle motion ($\beta = -1$). The explanation is that oriented swimming by larger organisms generates strong but highly episodic variance at small scales (Weber *et al.* 1986) while reducing variance at larger scales (Horne and Schneider, 1997).

Many of the problems with pattern analysis (Skellam, 1952; Hill, 1969 Pielou, 1969; Usher, 1969;) can be addressed with spectral analysis, which uses sophisticated techniques to estimate a simple quantity, the spatial density (spatial variance per unit frequency, Table 10.2), as a function of measurement frequency. Spectral methods perform as well or better than other methods of analysis of spatially contiguous counts (Ripley, 1981). The leading problem with spectral techniques (shared by simple pattern analysis of spatial variance) is sensitivity to low or zero values (Fasham, 1978; Horne and Schneider, 1997). Randomization methods (Mead, 1974) address the problem of computing p-values on the hypothesis that variance does not change with block size. Problems of estimation have not been addressed within the comprehensive framework of nonnormal error structures (McCullagh and Nelder, 1989).

Patchiness as a function of scale is also quantified via the semivariance (Figure 10.3), the spatial autocorrelation of densities N/A (as in Figure 10.3b), or one of its variants: Moran's I statistic, Geary's c-statistic. Plots of these statistics versus distance between values are called *correlograms* (autocorrelation, Moran's I, Geary's c) or *semivariograms* (semivariance). Comprehensive treatments of semivariogram analysis can be found in Cressie (1991) and Legendre and Legendre (1998).

Once spatial structure is described, it is natural to ask: What is the source of patchiness? When correlograms or periodograms have the same shape and there is a known causal connection between two sets of measurements (e.g., murres eat capelin), it is tempting to conclude that the structure in one variable arose from the other. The inference is far from certain (Horne and Schneider, 1995). In the northwest Atlantic, the spatial structure of cod (*Gadus morhua*) matched that of its capelin prey, yet the association between the two species was weak or nonexistent (Horne and Schneider, 1997) on most transects in the study. Measures of scale-dependent association are described in Section 10.8.

ANOTHER LOOK AT SECTION 10.7.1

Environmental scientists usually report scale-dependent variability as either a function of lags or a function of measurement frequency. List any exceptions known to you, then list some reasons that an individual would tend to use either lags or frequency, but not both.

10.7.2 Production and Erosion of Spatial Variance

The spatial variance is not a static quantity. It increases in response to some factors and decreases in response to others. For example, the motions of the earth's tectonic plates generate spatial variance in elevation above or below sea level while erosion acts to reduce this variance. A convenient symbol for this dynamic quantity is $\dot{\text{v}}\text{ar}(Q_{x\,y})$, read as "the rate of change in the spatial variance in the quantity Q". An example is $\dot{\text{v}}\text{ar}(z_{x\,y})$, the change in variance in the elevation of the ground above an x-y plane at sea level. This quantity depends on the time scale used to express it. At short time scales there occur brief yet violent changes in the variance in elevation, due to landslides, soil slumps, and earthquakes. At longer time scales the variance in elevation changes more slowly due to lateral gradients in weathering of rock or the isostatic rebound of continental platforms after the retreat of glaciers. The quantity $\dot{\text{v}}\text{ar}(z_{x\,y})$ depends on the resolution scale at which it is expressed.

Another example of variability as a dynamic quantity is the spatial variance in carbon fixation by green plants. Spatial variance in nutrient supply or light flux generates spatial variance in carbon fixation $\dot{C}_x$. The variance in production increases, represented symbolically by $\dot{\text{v}}\text{ar}\,(\dot{C}_x) > 0$. Acting against these processes are those that reduce spatial variance. One such process is intensive grazing in areas of high plant biomass, which reduces spatial variance in carbon fixation because production depends on standing stock of plant biomass. The outcome of this process, in terse symbolic form, is $\dot{\text{v}}\text{ar}(\dot{C}_x) < 0$.

The spatial variance in population density is another quantity subject to production and loss. An example is the generation and decay of patchiness in gelatinous zooplankton. Wind blowing over water creates cells of rotating water parallel to the wind. The flow at the water surface converges at regular intervals, at spacings on the order of 10 m in a light breeze and up to 100 m in a gale (Hamner and Schneider, 1986). Gelatinous zooplankton (sea-jellies from the phyla *Cnidaria* and *Ctenophora*) collect together at the convergences as they swim upward, forming windrows parallel to the wind. As the wind rises, Langmuir circulation intensifies, causing spatial variance in sea-jelly density to increase: $\dot{\text{v}}\text{ar}(N/A) > 0$. A shift in wind direction rapidly erases the spatial structure (Schneider and Bajdik, 1992), with a consequent decrease in spatial variance: $\dot{\text{v}}\text{ar}(N/A) < 0$.

Another example of the creation and destruction of spatial variance is the concomitant action of slow-moving or "bulldozer-type" predators compared to mobile predators that focus their activity in areas of high prey density. On intertidal sand flats, horseshoe crabs (*Limulus polyphemus*) dig up the surface of the sediment as they feed on invertebrates. When the tide recedes, the crabs angle downward into the sand and grind to a halt. Meanwhile shorebirds converge across the water from high-tide roosts to feed in areas of high invertebrate density. The bulldozers generate patchiness, the birds reduce patchiness (Schneider, 1992).

In all these examples (ground elevation, carbon fixation, animal density), spatial scale clearly matters. One of the major research challenges in ecology is to understand the creation and erosion of spatial variability as a function of spatial scale. Little enough is known at present about the factors that generate and remove variance at any given scale, let alone the rates of production and loss as a function of scale. An example of the current state of knowledge is an analysis of change in the spatial variance of an infaunal invertebrate, the bamboo worm *Clymenella torquata*, for which it was possible to predict the loss in spatial variance due to the distributional response of shorebird predators at the scale of flats (tens of hectares), whereas prediction was not possible at the scale of 1 ha plots (Schneider, 1992). Why were predictions successful at one scale and not at another? Were the distributional responses of predators ineffective at the smaller scale? Were there other factors operating?

A related question is the degree to which variance generated at one scale is transformed into variance at another scale. In fluid systems, spatial variance in velocity (i.e., the specific kinetic energy of flow) is transferred from large to small scales when larger-scale rotational structures are twisted and deformed into ever smaller eddies and swirls. This induces spatial variance in passively drifting plankton, but what about actively swimming nekton? At what scale do movements by active swimmers generate variability in density? Fish are known to interact with larger-scale fluid motions, generating spatial variance in density at the scale of hundreds and even thousands of kilometers. So, spatial variance in a swimming species arises through its interaction with large-scale flow structures, such as gyres.

An unexpected source of large-scale variance in density is local interaction of an organism with its environment. In theory, larger-scale structure can result from surprisingly small-scale patterns of oriented movement (Satoh, 1990; Hassell et al., 1991). How often does this propagation of small-scale structure to larger scales occur?

ANOTHER LOOK AT SECTION 10.7.2

For a quantity of interest to you, name processes that generate and that erode spatial variance.

10.8 Variance and Hue

When a multiscale examination of an ecological quantity is undertaken with the grouping maneuver, it turns out that variability in the quantity is often "red." *Red variability* is stronger at low measurement frequencies than at high frequencies. The variability is "red" by analogy with red light, which has more energy at low frequencies (long wavelengths) than at high frequencies (short wavelengths).

An example of a quantity for which variability is red with respect to time is rainfall. This quantity varies from hour to hour, due largely to the onset and end of rain events. Superposed on this variability is day-to-day fluctuation so that at the lower frequency of days there is more variance than from hour to hour. At a still lower frequency, at the time scale of months, there is an additional component of variability due to seasonal rain. Continuing on to still lower frequencies, there are components of year-to-year variability, decadal variability, and so on, to even longer time scales.

Variability in rainfall is red with respect to distance or area as well as with respect to time. Rainfall varies at the scale of tens of meters, which becomes evident when readings are taken from a set of rain gauges. At a lower spatial frequency, that of a watershed, another component of variability enters due to local differences in climate, such as the contrast in rainfall on the east and west side of a mountain range. At still lower spatial frequencies, major climate zones impose an additional component of variability.

Not all quantities have red variability. Some analyses carried out with the grouping maneuver turn up quantities with *green variability*. In other words, more variability appears at intermediate than at higher or low frequencies—hence "green" by analogy with green light. Examples of green variability can be found in pattern analyses of plant distribution and soil characteristics (Greig-Smith, 1983). Examples have also been reported for the spatial distribution of marine birds and fish, which show peaks in variability at scales on the order of several kilometers (Schneider, 1989). Variability can be green with respect to time as well as space. An example is a quantity with stronger seasonal than annual fluctuations.

If variability in ecological quantities can be red or green, why not *white variability* (same variability at all scales)? Or why not *blue variability* (more variability at high frequencies and short scales)? White variability is a convenient null model against which to test for statistically significant patterns, but like a lot of null models, its fate is to be rejected. A quantity with blue variability would have to fluctuate strongly at short time and space scales, whereas at larger scales variation would be damped out. There appear to be no examples from the ecological literature, unlike examples of red and green variability.

The hue that is evident in an analysis will depend on the scope. A quantity could look red if the scope is narrow, but if the scope is extended by either increasing the range or increasing the resolution, it might turn out that the red variability was the shoulder of a peak at some larger scale.

Several studies have shown that quantities can have *pink spatial variability*. That is, variability increases only gradually with increasing scale and less slowly than does red variability. Weber et al. (1986) used spectral techniques to quantify spatial variance in krill *Euphausia superba* at measurement frequencies ranging from half a cycle per kilometer to half a cycle per hundreds of kilometers. Within this scope they found only a shallow increase in variability with increasing scale (lower frequency). The rate of increase in variability was, on average, less than that of the red variability in properties (e.g., temperature) of the surrounding fluid. This means that krill form much stronger local aggregations than if they were passively coalesced and dispersed by the surrounding fluid. This stronger spatial variance at small scales presumably arises from schooling behavior. This hypothesis has been confirmed in examinations of the hue of spatial variability in another group of mobile marine organisms, fish (Schneider, 1994a; Horne and Schneider, 1995).

A close examination of the krill and fish studies shows that when many transects or repeated runs of the same transect are averaged together, the result is pink variability,

even though individual transects have green variability (Schneider, 1994a). Averaging together several cases of green variability produced pink variability because the green peaks do not coincide. Other examples of green variability, such as the pattern analyses reported in Greig-Smith (1983), may also turn out to be pink when averaged together.

The results of multiscale analysis of ecological quantities to date are conveniently summarized in terms of their *hue*; many quantities have red variability, some have green variability at a single time and place but turn out to be pink when averaged over several places or times. The concept of hue can be given a definite quantitative expression, but first we need some notation with an example.

Here is a made-up example of a series of counts with green variability. The variability is concentrated in two peaks of approximately 4 units. The peaks are on either side of the transect so that there is little contrast between the left and right sides. Contrasts are strongest at an intermediate scale of around 4 units, with less contrast at larger or smaller scales:

$$N := [1\ 2\ 0\ 8\ 4\ 20\ 1\ 0\ 2\ 0\ 2\ 5\ 10\ 0\ 4\ 1]$$

The mean density, coefficient of dispersion, mean squared deviation among groups, and spectral density were calculated at a resolution of $\boldsymbol{i} = 1$, according to the recipes in Table 10.1, then recalculated using groupings of $\boldsymbol{i} = 2$ contiguous units, $\boldsymbol{i} = 4$, and $\boldsymbol{i} = 8$ contiguous units. Box 10.4 shows the results. The spectral density shows green variability—the maximum spectral density occurred at a block size of $\boldsymbol{i} = 2$. Of course, this is an extremely crude estimate of spectral density; a better estimate could be obtained by using some of the more sophisticated features of spectral analysis, such as adjustment of the shape or size of smoothing windows. The purpose here is to display the hue of variability, not to estimate the scale of maximum variability, so I have kept the computational details as simple as possible. In this example I have used a population variance $\mathrm{Var}(\sum N)$ because this is not meant as a sample from a larger population.

Box 10.4 Green Variability at Resolution of $\boldsymbol{i} = 1$, 2, 4, and 8 Units

$\boldsymbol{i}$	=	1	2	4	8
n	=	16	8	4	2
f	=	2^{-1}	4^{-1}	8^{-1}	16^{-1}
mean($\sum N$)	=	3.75	7.5	15	30
var($\sum N$)	=	25.69	47.25	38	36
CD($\sum N$)	=	6.85	6.3	2.53	1.2
$\mathrm{MSA}_i(N)$	=	25.69	23.63	9.5	4.5
SpD(N)	=	51.37	94.5	76	72

The made-up example of green variability is "painted" red by moving the cluster of high counts over close to one end of the transect.

$$N := [8\ 4\ 20\ 5\ 10\ 0\ 4\ 1\ 0\ 1\ 2\ 0\ 2\ 1\ 2\ 0]$$

This generates large-scale variability, evident in the strong contrast in counts between the right and left. Spatial statistics are again computed via the grouping maneuver. These show (Box 10.5) that variability now increases as we go from small to large scales or from low to high frequencies of measurement. The variability approximately doubles for each doubling of block size or halving of measurement frequency. This is an example of red variability in the strict sense—the spectral density is negatively proportional to the square of the frequency. A graph of red variability will have a slope of minus 2 when spectral density is plotted on a log scale against frequency on a log scale. The formal expression of this relation is:

$$\mathrm{SpD}(Q) \sim f^{-2} \tag{10.15a}$$

An alternative way of expressing this relation is that the percent rate of change in spectral density, relative to the percent change in frequency, is equal to minus 2. This alternative relation is obtained by recasting the first relation into logarithms, then taking differences:

$$\ln[\mathrm{SpD}(Q)] \sim -2 \cdot \ln(f) \tag{10.15b}$$

Taking derivatives, as shown in Chapter 14, we have:

$$\frac{d\ln[\mathrm{SpD}(Q)]}{d\ln f} = -2 \tag{10.15c}$$

Box 10.5 Red Variability at Resolution of $\boldsymbol{i} = 1, 2, 4,$ and 8 Units

$\boldsymbol{i}$	=	1	2	4	8
n	=	16	8	4	2
f	=	2^{-1}	4^{-1}	8^{-1}	16^{-1}
mean($\sum N$)	=	3.75	7.5	15	30
var($\sum N$)	=	25.69	57.75	182	484
SpD(N)	=	51.37	115.5	364	968

The graphical interpretation of this information is that the logarithm of spectral density decreases by 2 units for every unit increase in the logarithm of frequency.

What does this mean in terms of grouping unit $\boldsymbol{i}$? The relation of frequency to grouping unit $\boldsymbol{i}$ is:

$$f = (2\boldsymbol{i})^{-1} \tag{10.16a}$$

After taking logarithms, we have:

$$\ln f = \ln(2) - 1\ln(\boldsymbol{i}) \tag{10.16b}$$

Taking the derivative, we have:

$$\frac{d\ln f}{d\ln \boldsymbol{i}} = -1 \tag{10.16c}$$

To obtain the relation of spectral density to grouping unit $\boldsymbol{i}$, we multiply Equation 10.15c by 10.16c:

$$\left(\frac{d\ln f}{d\ln \boldsymbol{i}}\right)\left(\frac{d\ln[\mathrm{SpD}(Q)]}{d\ln f}\right) = -1 \cdot -2 \tag{10.17}$$

Noting that:

$$\left(\frac{d\ln f}{d\ln f}\right) = 1 \tag{10.18}$$

the result is:

$$\left(\frac{d\ln[\mathrm{SpD}(Q)]}{d\ln \boldsymbol{i}}\right) = 2 \tag{10.19}$$

In words, the graph of spectral density rises two logarithmic units for each unit increase in the logarithm of group size $\boldsymbol{i}$. At this point, try drawing a graph of red variability.

With the notational machinery in place for red variability, other hues can now be expressed in formal terms that lend themselves to computation and testing. White variability is a situation in which spectral density remains unchanged with frequency:

$$\mathrm{SpD}(Q) \sim f^0 \tag{10.20}$$

The graph on a logarithmic scale is flat, with a zero slope. Try adding white variability to your picture of red variability.

Pink variability is somewhere between red and white, with a slope somewhere between 0 and -2. An interesting special case of pink variability results from a random walk, such as the Brownian motion of minute particles in a fluid. Random walks result in large- and small-scale displacements from a starting point. The spectral density of these random displacements, plotted against frequency on a log-log scale, will have a slope of -1, which can be taken as the value of pink variability. Blue variability can be defined as a plot with a positive slope. Finally, green variability will have a positive and a negative slope, with a maximum value in the middle. Try adding green variability to your graph.

The example developed here leads naturally to the idea of a general expression for multiscale analysis, treated in more detail in Chapter 14 (Table 14.3). A zoom or multiscale comparison, in formal terms, is calculated as the change in an ensemble quantity with a change in grouping interval $\boldsymbol{i}$. In scanning, the resolution is held constant, or in mathematical terms, the unit vectors $\boldsymbol{h}$ for time and $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$ for space are held constant.

An ensemble quantity can now be either scanned with respect to step number or examined with a zoom lens to focus in on detail or zoom back to bring out pattern. In the case of spatial gradients ∇Q, what happens if we alter the distance over which the gradient is calculated? Does the contrast grow stronger as the resolution changes? The formal notation (developed in Chapter 14) is a means for making calculations based on Smith's principle (Smith, 1983) that more is learned by panning and zooming than with either by itself.

The distinctions between red and green variability provide quantitative expression of a longstanding but rarely articulated assumption that there is characteristic or "correct" spatial or temporal scale at which to investigate ecological phenomena. The implication of green variability is that there is a characteristic spatial and temporal scale of

maximum variability. Once this has been found, investigation at this scale is more efficient than at other scales because confounding effects of larger- or smaller-scale variability will be at a minimum. Late 20th century literature on "scale" advocated the identification of an appropriate scale of analysis; this assumes that variability is green. Further, if green variability is detected, the scale of maximum variability offers a clue to the processes that generate this variability. In a graph of green variability versus scale, one can visualize variability as being injected at the scale underneath the peak and damped away at other scales on either side of the peak. Much of the work on pattern analysis (Greig-Smith, 1983) has been motivated by a search for green variability. Green variability has been demonstrated repeatedly, but always for single transects or time series. In those cases where an average over several transects or series has been calculated, it turns out that the green peaks do not coincide and that the overall result is pink (e.g., Weber et al., 1986).

The implication of red or pink variability is that there is no one characteristic scale of variability. This means that there is no single scale at which investigations will be most efficient. The implication of the emergence of pink variability from averaging over multiple transects or times series is that a scheme that works in a case where variability is green will not be a sure guide to investigation in other cases. For example, a stratified survey, which has the aim of capturing maximum variability, may work well at a particular scale (areal extent) of strata but not work well in another place or at another time using strata at the same scale. Pink variability, when single transects or time series are green, points at ecological processes that act intermittently at different scales rather than at a characteristic scale. The prevalence of red or pink variability over green variability in the ecological literature suggests that efforts to identify "characteristic scales of variability" will not reveal any "best" scale but rather will tend to uncover new variability with increase in scope (increase in either extent or resolution).

ANOTHER LOOK AT SECTION 10.8

As the Gulf Stream flows northward, it spins off eddies that can last for months. As an eddy rotates, it spins off smaller eddies, which in turn spin off still smaller eddies, until eventually the energy of rotation of the large eddy is dissipated. Draw a picture of the large eddy, with smaller eddies within eddies. If you were to measure the variance in velocity along a transect across a Gulf Stream eddy, what color variance would you expect? Why?

10.9 Codeviances, Covariances, and Other Measures

Box 10.2 shows several different ways of weighting deviances computed from a single quantity. What about weighting one set of deviations according to another set of deviations? For example, what about weighting deviations in glacier thickness according to deviations in elevation? Weighting the deviations in glacier thickness by elevation gauges the propensity of ice to flow downhill—thick ice and steep slopes cause more flow than thin ice and shallow slopes. The weighted sum tells, in broad perspective, what ice will be lost or gained, to return to the quote at the beginning of Chapter 8. Or, to take another example, what about weighting the temporal deviations in litter fall from

chaparral plants according to the temporal deviations in decomposer activity beneath the plants? The weighted sum tells, in broad perspective, the tendency for flammable litter to build up in volume.

The *codeviance* is defined (Table 10.3) as a set of deviations $(Q - \bar{Q})$ weighted by another set of deviations $(Z - \bar{Z})$. When summed, the codeviations result in an ensemble quantity, the *sum of the products* SP(Q,Z)), that is analogous to the sum of the squares. Scaling to the number of products results in a *mean sum of products*, which is similar in structure to the mean squared deviance msd(Q). Table 10.3 displays the computational recipe that reduces rounding error in the mean sum of products. The strength of the association between the quantities Q and Z is measured by the mean sum of products. In biology it is standard practice to consider measurements of Q and Z to be samples from a population for which the true association is measured by the mean sum of products. Hence an estimate, the *covariance* cov(Q,Z), is used in place of the mean sum of products. This estimate (Table 10.3) is the mean sum of products scaled to (n–1).

Measures of covariance, like measures of variance, are multiscale comparisons of a collection of local values to a single larger-scale average. The zoom factor is the ratio of the range of the collection to the resolution of a single value. The estimate of the covariance cov(Q,Z) scales the sample of observations at hand to a larger population. As with the variance, the scaling factor is $n/(n - 1)$.

In most applications, covariances are normalized to a unitless ratio called the *correlation coefficient* corr(Q,Z), shown in Table 10.3. Normalization by the standard deviation of both component quantities reduces the covariance to a unitless ratio ranging from -1 (perfectly negative association) through zero (no association) to $+1$ (perfectly positive association). The formula in Table 10.3 is an estimate of the true correlation, which is denoted by the Greek symbol ρ (rho), just as the true variance was denoted by the Greek symbol σ (sigma).

These measures of association are typically used in the frequency domain as the correlation is plotted as a function of measurement frequency or, equivalently, group size (Table 10.3). We can also examine the correlation of one variable with another, at

Table 10.3 Commonly Encountered Measures of Covariability

Codeviations	$\text{codev}(Q, Z) \equiv (Q - \bar{Q})(Z - \bar{Z})$
The sum of the products	$\text{SP}(Q,Z) \equiv \sum \text{codev}(Q,Z)$
	$\text{SP}(Q,Z) \equiv \sum QZ - \frac{1}{n}\sum Q \sum Z$
The mean sum of products	$\overline{\text{SP}(Q,Z)} \equiv \frac{1}{n}\sum \text{codev}(Q,Z)$
The covariance	$\text{cov}(Q,Z) = \frac{1}{n-1}\sum \text{codev}(Q,Z)$
The correlation coefficient	$\text{corr}(Q,Z) \equiv \frac{\text{cov}(Q,Z)}{\text{std}(Q)\cdot\text{std}(Z)}$
The cross-correlation at lag k	$\text{ccf}_k(Q,Z) \equiv \text{corr}(Q_x,\ Z_{x+k})$
The cosemivariance at lag k	$\gamma_k(Q,Z) = (2n)^{-1}\sum (Q_x - Q_{x+k})(Z_x - Z_{x+k})$

lags of 1, 2, and so on. This is called the *cross-correlation* $\text{ccf}_k(Q, Z)$. The other measure in the distance domain is the *cosemivariance* $\gamma_k(Q,Z)$.

Covariances lend themselves to biological or physical interpretation, as in the examples of glacier thickness and litter fall at the beginning of this section. The covariance of ice thickness with elevation was physically interpretable. The covariance of litter fall with decomposer activity was biologically interpretable. But many covariances will lack this interpretable quality. An example often cited in statistics texts is the temporal covariance between sunspot numbers and rodent numbers. The units of the covariance are animals · sunspots · year^{-1}, an uninterpretable unit. The easiest way to determine whether a covariance is an interpretable quantity or just a statistic of association is to write out the units and decide whether these are interpretable. The units might look strange, but they should not be dismissed on this count.

Intuitively, the scale at which one chooses to examine the association of one variable with another does not matter. The ecological literature prior to the 1980s assumes this, because associations are reported as either present or absent, with no statement of scale. In fact, however, associations that were reported as weak or absent may well have been missed because they were present at some scale other than that at which measurements were taken. Figure 10.4 shows an example. Marine birds feed on schooling fish, yet data from shipboard surveys usually showed weak or nonexistent association. Figure 10.4a shows the data at the scale at which they were recorded (1 minute intervals corresponding to 0.25 km transects). Figure 10.4b shows the same data at a coarser scale of 2.5 km intervals. The association at the small scale ($r = 0.012$, Figure 10.4a) is typical of values reported in the literature in the 1980s (e.g., Safina and Burger, 1985). The association is strong ($r = 0.885$) at the coarser scale of 2.5 km intervals (Figure 10.4b). The next logical step is to compute association at a range of spatial scales by aggregating adjacent pairs along the transect, then adjacent triplets, and so on. Figure 10.4c shows correlations at measurement frequency ranging from 0.5 cycles per measurement to 0.05 cycles per measurement, corresponding to bin sizes from 0.5 km to 5 km. This method (Greig-Smith, 1952), though easy to undertake, is known to perform unreliably, especially at coarse scales (e.g., large bins corresponding to a fifth, a quarter, or a third of the entire transect). As with the computation of variability with increasing block size, many of the problems encountered in analyzing associations are removed by employing the machinery of spectral analysis. This gives a better estimate of the covariance (Table 10.3) than does the crude approach based on increasing block size. For this data, the results from spectral analysis (Figure 10.4d) were the same as for Greig-Smith pattern analysis (Figure 10.4c), although this will not always be the case.

The association of murres with capelin prey can also be displayed in the distance domain, using one of the measures in Table 10.3. The cross correlation of murres with capelin prey (Figure 10.5b) shows maximum association at lags of $k = 6$ (1.50 km) to 7 (1.75 km) and again at $k = 11$ (5.25 km), when murres are taken as lagging behind (encountered later in the transect than) the major cluster of capelin. The lags estimate the separation between the murre patch and two of the capelin patches. Cross-correlation drops to small values at other lags, including lags where capelin lead murres in the direction the transect was traversed. The cosemivariance (Figure 10.5c) rises to a maximum at lags of $k = 4$ to 5 and again at lags of $k = 9$ to 10, again estimating the separation of the murre patch from two capelin patches. These scale-dependent patterns of association result from the feeding behavior of the birds, which aggregate episodically

at fish schools. The cycle of coalescence, feeding, and subsequent drift (Hoffman et al., 1981) results in stronger association at coarse than at fine scales as bird aggregations tend to drift away from prey clusters until birds reaggregate over a cluster to feed.

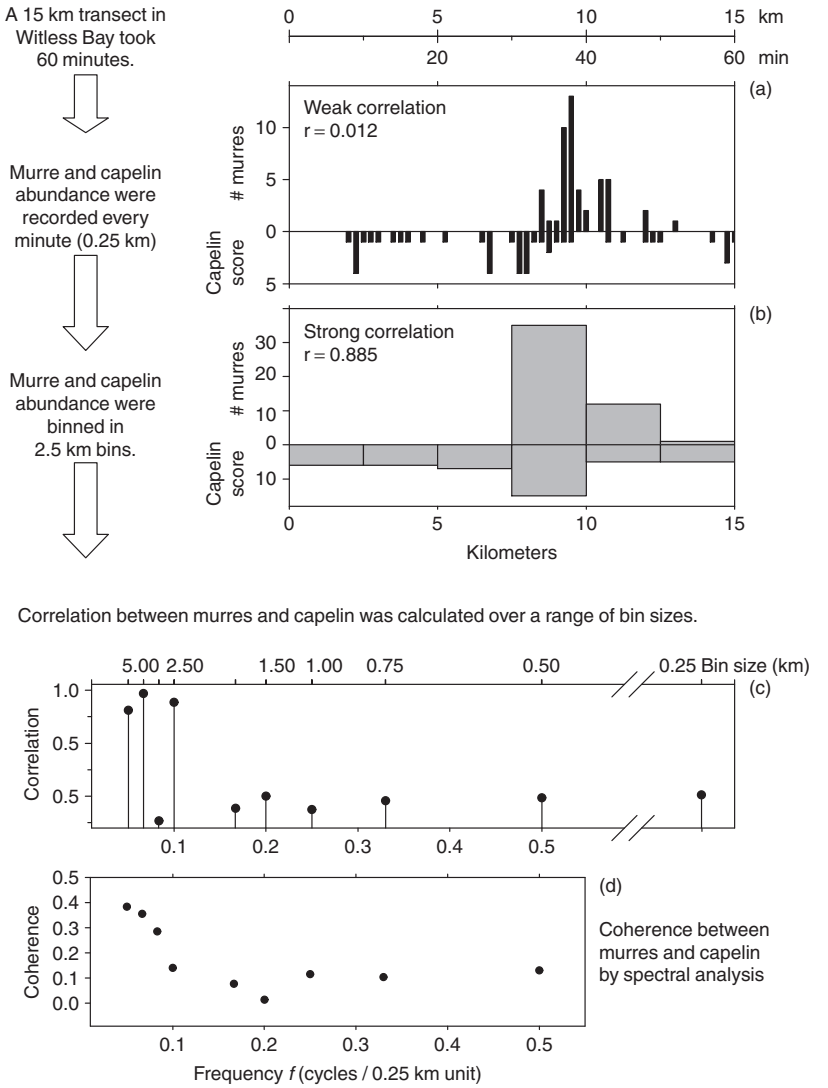


FIGURE 10.4 Spatial Association as a Function of Measurement Frequency. Same data as Figure 10.3. Redrawn from Schneider (2002); (a) Association of murres with capelin at high resolution (high frequency); (b) Association at coarse resolution (low frequency); (c) Association estimated by correlation; (d) Association estimated by spectral analysis.

ANOTHER LOOK AT SECTION 10.9

Pick two variables of interest to you. Make an educated guess as to the scale of association. Then, using one of the formats in Figures 10.4 or 10.5, sketch the suspected association as a function of lag or block size.

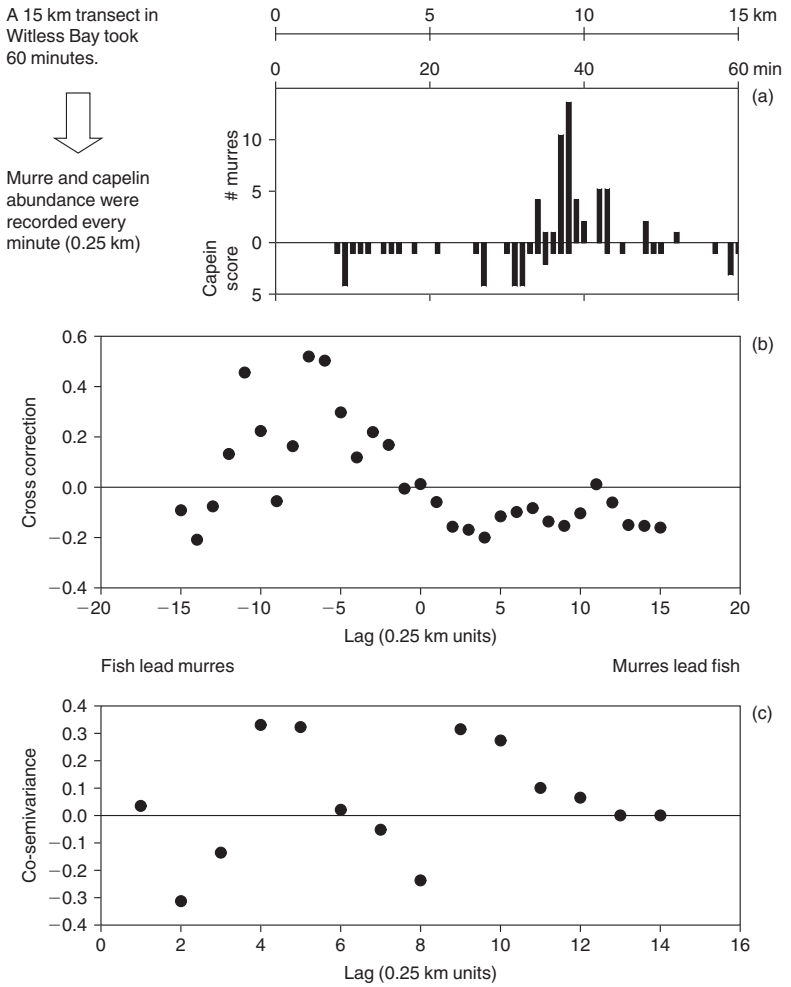


FIGURE 10.5 Spatial Association as a Function of Distance (Separation). Same data as Figure 10.3; (a) Distribution of murres relative to capelin at high resolution; (b) Association as measured by cross correlation. (c) Association estimated by co-semivariance.

10.9.1 Potential Contact

Biological interactions such as prey capture, parasitism, competition, gene exchange, and habitat selection proceed at rates that depend on the frequency of contact, a quantity that depends on the degree to which organisms N are aggregated, the degree of spatial heterogeneity in some environmental factor Z, and the covariance of N with Z. The dependence of ecological interaction on spatial structure is quantified as the potential contact PC_i at spatial resolution scale $\boldsymbol{i}$. *Potential contact* is defined as the product of local abundance N_x and local concentration of an environmental factor Z_x within the limits of some unit length or area:

$$PC_i \equiv \sum^{n} \frac{\sum^{i} N_x \sum^{i} Z_x}{\sum^{n} N_x} \tag{10.21a}$$

In this expression, the summation $\sum^n$ applies to a series of ambits. Each ambit is of size $\boldsymbol{i}$, which can have units of distance, area, or volume. The summation is over n ambits, so the total range is $n \cdot \boldsymbol{i}$. There is a further summation for each value of $\sum^i N_x$ and $\sum^i Z_x$. These represent sums within each stretch of size $\boldsymbol{i}$. In terser notation, Expression 10.21a becomes:

$$PC_i \equiv \sum \frac{\Sigma N \cdot \Sigma Z}{\Sigma \Sigma N} \tag{10.21b}$$

Box 10.6 demonstrates the concept of potential contact, with a series of worked examples. In the first example, potential contact of predators with prey is computed

Box 10.6 Computation of Potential Contact PC_i

The potential contact of predatory fish N with prey Z at the scale of tidal pools. $\boldsymbol{i}$ = pool.

			Pool 1	Pool 2	Pool 3	Pool 4	Pool 5
Prey	Z	=	10	0	1	8	50
Predators	$\sum N$	=	0	1	4	3	2
	$\sum N \cdot \sum Z$	=	0	0	4	24	100
	$\frac{\sum N \cdot \sum Z}{\sum \sum N}$	=	0	0	0.4	2.4	10

The potential contact summed over all pools:

$$PC_i = \sum \frac{\sum N \cdot \sum Z}{\sum \sum N} \Rightarrow 12.8 \text{ potential contancts per predator}$$

If the fish rearrange themselves at high tide so as to be more closely associated with prey, the potential contact within each pool changes.

			Pool 1	Pool 2	Pool 3	Pool 4	Pool 5
Prey	$\sum Z$	=	10	0	1	8	50
Predators	$\sum N$	=	3	0	0	2	5
	$\sum N \cdot \sum Z$	=	30	0	0	16	250
	$\frac{\sum N \cdot \sum Z}{\sum \sum N}$	=	3	0	0	1.6	25

The potential contact summed over all pools in this new situation rises.

PC_i = 29.6 potential contacts per predator

As the pools join at high tide, the scale changes from $\boldsymbol{i}$ = 1 pool to $\boldsymbol{i}$ = 5 pools joined. The potential contact rises:

$$\sum \frac{\sum N \cdot \sum Z}{\sum \sum N} = \sum \frac{10 \cdot 69}{10} = 69 \text{ potential contacts per predator}$$

for separate pools, then summed over the pools. In the second example, the fish aggregate relative to prey, which increases the potential contact at the scale of pools. Finally, potential contact can be calculated at the larger scale of five pools joined (not shown in Box 10.6). The numbers of predators and prey remain unchanged, whereas potential contact changes with spatial scale $\boldsymbol{i}$. Potential contact can be computed from a theory or a model; the symbol is $PC(\boldsymbol{i})$ to distinguish it from the measured value PC_i.

One of the more venerable concepts in population biology is that the rate of increase in a predator population depends on the interaction with prey. This is usually taken as the product of the total number of predators and prey, adjusted by a fixed coefficient. In the case of tide pool fish, the potential for interaction according to this notion would have to be computed as the product of all fish and all prey over the entire range of the predator, not just the one pool in which fish come into contact with prey. Spatial scale is assumed to be that of the entire population when in fact it is more local. Potential contact $PC_{\boldsymbol{i}}$ takes local spatial scale into account. A yet more realistic measure is realized contact, the number of prey that come into the visual range of a fish. This would be some percentage of the potential contact, depending on the behavior of the fish and its prey. Finally, capture rate would be some percentage of the realized contact.

The formal expression for potential contact is closely related to the spatial covariance (Schneider, Gagnon, and Gilkinson, 1987):

$$PC_{\boldsymbol{i}} = \frac{\text{cov}(N, Z)}{\text{mean}(N)} + \text{mean}(Z) \tag{10.22}$$

If the environmental factor is the number of conspecifics within the ambit ($Z = N - 1$), the general expression for potential contact becomes Lloyd's expression for mean crowding, or potential social interaction.

Potential contact depends on three quantities, each of which can increase or decrease the overall potential for contact. Potential contact increases if the covariance of density with the environmental factor increases, as in the tide pool fish example. Potential contact decreases as the overall mean density increases, assuming that covariance is not changed by an increase in larger-scale density. In most cases, however, the covariance will change with change in overall density. Potential contact increases as the mean value of the environmental factor increases, again assuming that this does not alter the covariance.

What units does potential contact PC_i have? It has the same units as the ratio of the covariance to the mean, that is, $\text{cov}(N,Z)/\text{mean}(N)$. The numerator of this ratio is visualized as contacts between N entities and the environmental factor Z. The units of Z might also be entities, such as the count of another species. The environmental factor can have other units—mass, concentration, and so on. The numerator $\text{cov}(N, Z)$ has units of $N \cdot Z$. The product of entities and the units of Z will be defined as a potential contact. The denominator of the ratio is a mean, which has units of entities. So the ratio of the covariance to the mean is a per capita contact, which has the same units as the environmental factor Z.

Contacts and per capita contacts apply to a wide range of ecological phenomena. Table 10.4 shows a diverse collection of examples drawn from situations in which the product $N \cdot Z$ has ecologically interpretable units.

Table 10.4 Partial List of the Units of Potential Contact $PC(i)$

N	Z	$PC(i)$
Parasites	Hosts	# · #
Population number	Competitors	# · #
Predator count	Prey count	# · #
Predator count	Prey biomass	# · Mass
Leaf number	Light flux	# · Einstein
Root number	Nutrient flux	# · Moles · Area^{-1} Time^{-1}
Whale number	Vessel hours	# · Time
Wild type allele	Mutant allele	# · #
Recessive mutant	Recessive mutant	# · #

Potential contact PC_i can be calculated from spatial data or from published estimates of spectral density and coherence. Table 10.5 shows the relation between potential contact and spectral statistics. Potential contact can also be computed from statistics calculated according to the methods of Greig-Smith (1983), in which the mean squared deviation MSA_i among groups of size i (Table 10.2) is plotted against group size. Bult et al. (1997) demonstrate the use of potential contact in quantifying density-dependent habitat selection. O'Driscoll et al. (2000) found that collapse of cod stocks in the northwest Atlantic in the 1990s was not associated with any reduction in potential contact with prey.

Table 10.5 Relation Between Potential Contact PC_i and Spectral Statistics

Definitions

Resolution	i				
Range	$n \cdot i$				
Frequency	$1/(2i)$				
Count spectrum	$\text{CS}(N)$	$\equiv$	$\text{SpD}(N)$	$\equiv$	$f^{-1}\,\text{var}(N)$
Spectral density	$\text{SpD}(Z)$	$\equiv$	$f^{-1}\,\text{var}(Z)$		
Coherence	$\text{coh}(N,Z)$	$\equiv$	$\dfrac{Cov(N,Z)}{CS(N)\cdot CS(Z)}$		

Potential contact is

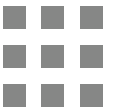
$$PC = \frac{2\,\text{coh}(N,Z)\cdot\sqrt{\text{CS}(N)}\cdot\sqrt{\text{CS}(Z)}}{\text{mean}(N)} - \text{mean}(Z)$$

ANOTHER LOOK AT SECTION 10.9.1

Sessile organisms such as trees and barnacles typically have a mobile life history stage. For the mobile stage of a sessile species, list factors for which potential and realized contact with those factors affect survival.

Defined Concepts and Terms for Review and Future Reference

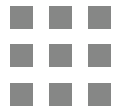
____ average deviation
____ autocorrelation
____ codeviations
____ coefficient of dispersion
____ coefficient of variation
____ confirmatory vs. exploratory analyses
____ correlation coefficient
____ correlograms
____ cosemivariance
____ covariance
____ cross-correlation
____ deviance
____ deviation
____ frame
____G-statistic
____ hue (red, green, white, pink, blue)
____ mean crowding
____ mean square variance among groups
____ mean sum of products
____ measures of spatial variance
____ measures of temporal variance
____ multi-scale comparisons
____ pattern analysis (Grieg-Smith plot)
____ periodogram
____ potential contact
____ semivariance
____ semivariogram
____ spectral density
____ sum of products
____ sum of squared deviations
____ Taylor's Power Law
____ variance
____ variogram
____ weighted deviances



PART

III

Scope



11 The Scope of Quantities

The rationale was that a science which neglects the reality and distinctive character of large-scale ecological processes provides a poor basis for tackling many of the more urgent problems in natural resources management.

—R. M. May, Preface, *Large-Scale Ecology and Conservation Biology*, 1994

11.1 Synopsis

We have already defined the scale of a quantity as the resolution within the range of measurement. The ratio of the range to the resolution is the scope, a number with no units and hence no dimensions. Scope can be thought of as the number of steps, given the step size. It applies to measurement instruments, the quantities measured in research programs, the research programs themselves, and the equations describing ecological patterns and processes.

A single measurement has a scope, which is the ratio of its magnitude to its precision. When applied to ratio, interval, ordinal, and nominal scale measurements, the concept of scope brings out how these four types of scale differ. These differences are readily displayed as a scope diagram.

Any measurement instrument has a scope, which is the ratio of its maximum reading to its resolution. Scope calculations are useful in comparing the capacity of measurement instruments.

The quantity chosen to measure natural phenomena will have a scope determined by a measurement relation, which scales measurement outcome to the scope of an operationally defined procedure. The principle of homogeneity of scope, as it applies to measurement relations, guides the development as well as application of power laws, including scaling relations estimated by regression of one quantity on another.

Strictly speaking, scale and scope are the result of measurement activities, not characteristics of the natural phenomena that are the object of these activities. However, scale (and hence scope) are routinely considered characteristics of natural phenomena, an approach that is of utility in evaluating planned programs of measurement. If a phenomenon has definable boundaries in space, its spatial scope can be defined as the ratio of diameters of the largest and smallest cases. We might, for example, speak of the spatial scope of low-pressure storm systems or the scope of bird migration. If a phenomenon is episodic, with discernable starting and ending points, its temporal scope can be defined as the ratio of duration of longest and shortest cases. Scope calculations are useful in comparing phenomena. Scope diagrams display these comparisons in a convenient and effective fashion. Scope diagrams need not be limited to time and distance, though these are the most common.

doi: 10.1016/B978-0-12-627865-1.00011-8

11.2 The Scope of a Measurement

Measurements occur on a scale (Box 6.1), and consequently a single measurement has a scope. The *scope of a single measurement* is defined as the ratio of the magnitude of a measurement to its precision. For example, a measurement of 10.2 cm implies a precision of 0.1 cm, from which the scope is calculated as 10.2 cm/0.1 cm = 102. If calibration of the measuring instrument shows that the precision was in fact no better than 0.25 cm, the scope is 10.1 cm/0.25 cm = 41.

What about measurements on nominal, ordinal, or interval types of scale? Does the concept of scope apply? Let's begin with interval scale measurements. These have a scope, just as much as ratio scale measurements do. For example, if the air temperature today is 5°C and we measured this to the nearest degree, the scope of the measurement is 5°C/1°C = 5. If we make this same reading to the nearest tenth of a degree, the scope of the measurement becomes much larger: 5°C/0.1°C = 50. Figure 11.1a shows the scope of two readings that have the same magnitude (5°C) but different scopes.

A *scope diagram*, such as Figure 11.1, displays the outer scale (extent, range) connected by a line to the inner scale (resolution, minimum value, grain). In Figure 11.1 a single reading is represented as a horizontal line that starts at the resolution and ends at the magnitude of the measurement. The length of the line represents the scope of the measurement. A base-10 logarithmic scale was used; consequently, the length of the line shows the number of tenfold increases relative to the step size marked by the left side of the line.

Figure 11.1b compares the scope of two readings made with the same resolution. The negative temperature in this example has a greater scope because it happens to be further from the arbitrary zero point, the freezing point of water. If we had chosen a temperature scale with a zero point at the freezing point of mercury (−38.8°C), the readings would become +28.8° and +43.8°, and the reading above the freezing would have the greater scope. The scope of an interval scale measurement depends on the arbitrary value where the counting of steps begins, as well as on resolution (step size).

The scope of an interval or ratio scale measurement is the number of steps away from the zero point, so the zero point itself cannot have a scope. In the case of temperature, the freezing point of water does not have a scope on the Centigrade scale. The freezing point of water does have a scope on the Kelvin scale. If the resolution is 1°K, the freezing point of water has a scope of 273.

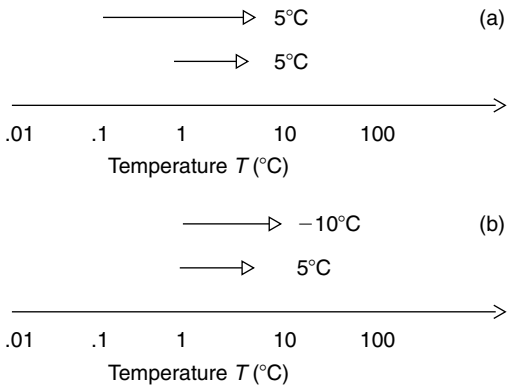


FIGURE 11.1 Scope of Measurement; (a) Scope of two temperature readings having the same magnitude and different resolution; (b) Scope of two readings having the same resolution.

Table 11.1 Scope of Measurement on Four Types of Scale

Types	Nominal	Ordinal	Interval	Ratio
Examples	Experimental treatment	Rank abundance	Calendar date	Age
Resolution	presence	rank	1u	1U
Range	presence	highest	S · 1u	S · 1U
Scope	1		S	S

What about measurements on other types of scales? The scope of an ordinal scale measurement is always equal to its numerical value because the resolution of all such measurements is the same: one step in rank. The rank is the scope, unlike measurement on a ratio scale. Because of the definition of step size, ordinal scale measurements cannot be rescaled, via the operation of division, to other measurements. Finally, a measurement on the nominal scale is either at the zero point or is one step away. The scope always comes out to be unity. On a nominal scale, all measurements have the same scope.

The concept of scope highlights the ways in which the four types of measurement scale differ (Table 11.1). The scope of a ratio scale measurement reflects both its resolution and its magnitude. The scope of an interval scale measurement reflects both its resolution and its distance from the zero point. The scope of an ordinal scale measurement is due only to the number of steps (ranks) from the starting point, because all measurements on this scale have the same resolution. And all nominal scale measurements have the same scope.

The four types of measurement scale differ in their information content, a situation that is reflected in the scope, or number of steps on that scale. The nominal scale, which is the least informative, has only one step. The ordinal scale, which is more informative than the nominal scale, has a restricted number of steps, no more than the number of objects being compared. Interval and ratio scales, which are still more informative, have far more steps. The number of steps goes up with increasing resolution and with increasing distance from the zero point. Table 11.1 recognizes these differences in information, which are designated by four types of units: 1U for ratio scales, 1u for interval scales, ranks for ordinal scales, and presences for nominal scales. Nominal and ordinal scales are less informative than ratio types of scales, but it does not follow that quantities on these scales are the same. Nor does it follow that quantities on these types of scale are merely numbers devoid of units.

ANOTHER LOOK AT SECTION 11.2

What happens to the cost of an instrument, in general, as the precision increases? Sketch a general relation between cost of an instrument and scope of each measurement from the instrument. Can you think of any exceptions to the relation you have drawn?

11.3 The Scope of Instruments

Telescopes, microscopes, oscilloscopes, stethoscopes, hygroscopes—all have a specific capability for measurement, set by upper and lower limits. These instruments, like any

other, have a capacity limited by their resolution and maximum attainable measurement. The *scope of an instrument* is defined as the ratio of the maximum measurement to the resolution. This is a dimensionless ratio that defines the capability of the instrument. The greater the scope, or number of possible steps, the greater the capability of the instrument. For example, a meter-stick has a capability or scope of 100 if marked in centimeters. It has a scope of 1000 if marked in millimeters. The scope of a measurement instrument is analogous to the scope of a musical instrument. A piano has a scope of 88. That is, it has a resolution of one-twelfth of an octave, within a range of a little over seven octaves.

The lower limit, or resolution, is often set by the "just noticeable difference" in reading instruments. The just noticeable difference on a simple caliper is about half a millimeter. Smaller differences, of the order of a tenth of a millimeter or less, are hard to read accurately. The addition of a vernier to a calliper extends the resolution to tenths of millimeters by making the divisions within a millimeter easily readable by eye.

The upper limits of an instrument sometimes depend on how large it is. A large balance, for example, can record a greater range in mass than a small balance. A much larger balance is needed to weigh a moose than to weigh a mouse. Upper limits also depend on the way the instrument is used. A meter stick has an upper limit of 1 meter when applied once, but when applied repeatedly in a straight line, it has an upper limit of several tens of meters. A surveyor's chain has by itself an upper limit of around 15 m. When applied repeatedly in a straight line by a surveying party with a level, the upper limit rises to tens of kilometers or more.

The capability of one instrument relative to another comes through clearly in a scope diagram. Figure 11.2 compares the scope of several instruments with differing capacities to measure distance. A logarithmic scale has been used, as in previous scope diagrams. The horizontal line for each instrument starts at its resolution, then ends at the upper limit of measurement. The separation between start and stop points shows the scope, or number of steps possible for that instrument, relative to its resolution. Similar diagrams have been constructed for instruments to collect remotely sensed data (Fuller 1989) and for techniques to collect paleoecological data (Schoonmaker, 1998).

Scope diagrams make clear at a glance the capability of techniques or instruments relative to quantities (Figure 11.2). If the diameter of eukaryotic cells (cD = μm) is of interest, a single instrument, the microscope, will suffice. If body length (bL = m) is of

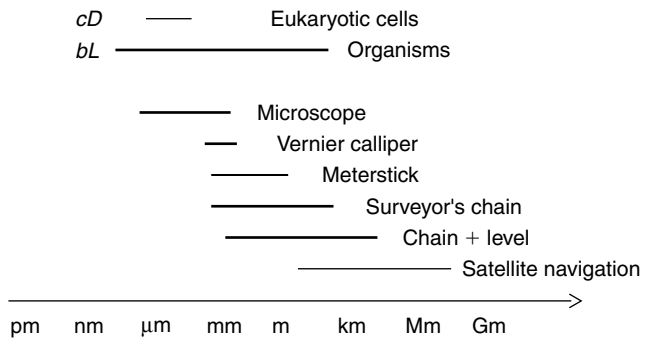


FIGURE 11.2 Spatial Scope of Instruments to Measure Lengths. The spatial scope in cell diameter *cD* and organism body length *bL* are shown for comparison.

interest, from the smallest free-living organism, *Mycoplasma*, to the largest, the blue whale (*Balaenoptera musculus*), several instruments will be required.

ANOTHER LOOK AT SECTION 11.3

For three instruments that you have used, state the resolution and maximum possible reading. Compute the scope and compare instruments.

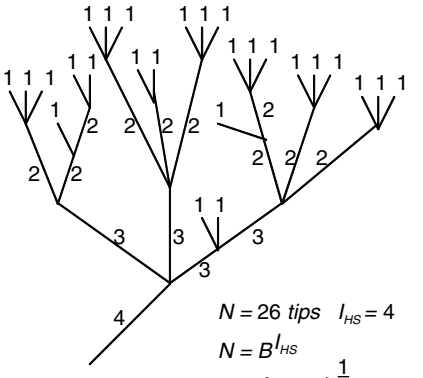
11.4 The Scope of Quantities

Scaled quantities are needed to read and express ecological ideas, just as a musical scale (steps per octave and number of octaves) is needed to read and perform music. The *scope of a quantity* will be set by the measurement protocol and relation that defines the quantity. The procedural statement for a quantity (refer back to Section 3.2) includes information about the unit of measurement, the way an instrument or technique was used, whether the protocol was iterative and whether iteration was used, and how it was applied (refer back to Figure 7.1).

Iterative protocols are essential for measuring complex phenomena (Box 2.1). Iterative protocols are conveniently grouped in three categories (refer back to Section 7.4, Figure 7.1). In coarse graining, the inner and outer scale are fixed while the frequency of measurement (resolution or grain) is progressively altered within this frame. An example is the treatment of spatial variance of murres (Figure 10.3e). Coarse graining gives us a range of small to large box sizes via repeated use of measurements at a single location. In lagging, the inner and outer scale are fixed while the separation (distance between two measurements) is progressively altered. An example is the spatial autocorrelation of murres (Figure 10.3b). Lagging gives us a range of small to large separations, again from a repeated measurements. In accumulation, the unit size is allowed to grow from a single small unit (inner scale) to a single large unit (outer scale). An example is a cumulative variance of birds along a transect, plotted against distance from starting point, as in Figure 10.3d. Accumulation gives us a range of small to large extents within the overall extent of the protocol.

Our concepts of measurement are strongly influenced by Euclidean schemes—steps along a straight line or square tiles on a grid. For many systems of interest to ecologists, there are patterns of directional connectivity missed by Euclidean protocols such as coarse graining. For dendritically connected systems (blood vessels, nerves, rivers and their basins), we can say that one point is upstream from another or that one basin is nested within another basin. For such systems, Horton-Strahler accounting (Horton 1945, Strahler 1952) will serve as a consistent and more informative measurement protocol than Euclidean segments on a line or boxes on a grid. Box 11.1 shows Horton-Strahler accounting, which is widely used for scaling laws in geomorphology (e.g., Rodriguez-Iturbe and Rinaldo, 1997; Veitzer and Gupta, 2000). The Horton-Strahler procedure is an iterative protocol that lends itself well to questions concerning scale in dendritically structured systems. The scope of the protocol is the HS index (Box 11.1), that is, the number of branching levels in the system.

Box 11.1 Horton-Strahler Accounting in Dendritically Organized Systems



In the diagram, the HS index is 4, and the branching ratio works out to be 2.26. For rivers, the branching ratio is between 3 and 5. Assuming a value of 4, the Mississippi River has $4^{12} = 16.8$ million starting points.

In this method the number assigned to each branch is obtained iteratively. Each tip (headwater, capillary, or leaf) is assigned a value of 1. Two first-order branches combine to produce a second-order branch, two second-order branches combine to produce a third-order branch, and so on. When a branch of smaller order combines with a branch of larger order, the order of the larger branch is not changed. The value reached at the root measures the size of the network. This value is called the *HS index*. The Mississippi River has an index of $I_{HS} = 12$. This index has the convenient property of relating the number of tips N to the branching ratio B.

The scope of an iteratively measured quantity depends on the measurement relation that establishes a one-to-one relation between a quantity measured at one scale and the same quantity measured at a different scale (Boxes 5.10 and 6.1). An example is the iterative measure of the area of a complex shape, such as an island (Equation 6.14):

$$\frac{A}{A_o} = \left(\frac{L}{L_o}\right)^{-D+1} \tag{11.1}$$

In this example, iterative measurement is by coarse graining. On the right side the overall scope is the ratio of the largest grid spacing (outer scale) to the smallest grid spacing (inner scale) nested within the largest grid. Within the overall scope we allow scope L/L_o to vary at ratios convenient for analysis. On the left side the overall scope is the ratio of observed area for the largest and smallest grid spacing. Within this overall scope we have ratios for each value of A/A_o. Equation 11.1 establishes a one-to-one relation based on an exponent. In this example, we have an iterative protocol for measuring island area.

We need to state our protocol in detail because it turns out that iterative protocols with the same scope produce different results. For example, the coastline length as measured by the number of box sides will differ in length from that measured by dividers, even if we use the same resolution (minimum divider step or side of box) and same scope. Even with the same instrument, different protocols produce different results. We can measure island perimeter by swinging a caliper clockwise, counting steps along the coast from the seaward side as the caliper walks clockwise along the coast. Or we can swing counterclockwise, counting from the landward side and walking counterclockwise along the coast. The protocols produce different results depending on whether the divider steps along the inside of an arc or along the outside of an arc. The effects are greater for small islands. To balance the effects, we can use a third method: Alternate the direction of swing at each step. With this protocol the caliper can walk either clockwise or counterclockwise around an island. As some readers might know, these three protocols (all using the same instrument set at the same sequence of step sizes) produce different results. Thus the result of measurement depends on some unexpected details of the measurement protocol. Iterative protocols are needed for measuring complex phenomena, but care is needed in applying protocols. Far more care is needed than with traditional noniterative measurement protocols that count steps in straight lines, tiles on a Cartesian grid, or cubes in a regular solid.

Noniterative measurement relations scale a quantity to fixed units such as survey quadrats or experimental tanks. An example is primary production in experimental tanks (mesocosms) that differ in experimentally fixed diameter and height (Chen et al., 1997). Another example is species numbers in relation to quadrats of different area (Equation 6.11, Box 2.3).

$$\frac{Nsp(A)}{Nsp(A_{ref})} = \left(\frac{A}{A_{ref}}\right)^{\beta} \tag{11.2}$$

On the right side of Equation 11.2, the overall scope is the ratio of the area of the largest to the smallest unit area (quadrat) in the study. In principle we can designate any size of quadrat as having the reference area A_{ref}. For convenience, we designate the smallest unit as having the reference area. The scope relative to the smallest quadrat depends on the size of the largest quadrat. On the left is the ratio of the count of species for units of area A and A_{ref}. Equation 11.2 establishes the one-to-one relation between these two ratios.

Noniterative scaling relations scale one quantity to another across a series of objects. The objects can be organisms, islands, lakes, ecosystems, or any other units that vary naturally. Scaling relations usually rest on at least two measurement protocols and hence on at least two measurement relations. For example, the overall scope of the noniteratively measured quantity body mass M is the ratio of the mass of the largest organism to the smallest organism in the study. The scope of respiration for the same organisms will be the ratio of the highest respiration to the lowest respiration, for which there is a different protocol. By combining studies, the scope of M can be extended to the maximum possible, which is the number of *Mycoplasma* mass units (the smallest organism) in a blue whale (the largest organism). This number is on the order of 10^{21} (Schmidt-Nielsen, 1984). Within this scope, we expect that several protocols will

be necessary, since no single protocol can be applied to measure body mass within the overall scope of this quantity. Nor can any one protocol be applied to measure respiration over this range of body sizes. Thus the scope of the scaling relation of respiration to body mass, over the entire range of body mass of multicellular organisms, will rest on several measurement relations for body mass and several measurement relations for respiration.

ANOTHER LOOK AT SECTION 11.4

List a phenomenon of interest to you. State how you would measure the phenomenon. With this measurement protocol, can you measure the phenomenon throughout its scope?

11.5 The Scope of Parameters and Equations

Parameters relate one variable quantity to another. The *scope of a parameter* is a ratio that scales the response variable (on the left side of an equation) to an explanatory variable (on the right side). The scope of the parameter is worked out from *homogeneity of scope:* The terms in an equation have the same scope as well as dimensions. For example, here is a simple equation relating puffin density [*N*] = count/km^2 to radial distance from a colony r = kilometers.

$$[N] = \beta_o + \beta_r \cdot r + \in \tag{11.3a}$$

The parameter β_r is a gradient—the rate of change in density with radial distance. A quick analysis (as in Chapter 6) shows that the gradient β_r has dimensions of # L^{-3}:

$$\begin{aligned} [N] &= \beta_o + \beta_r \cdot r + \in \\ \frac{\#}{L^2} &= \frac{\#}{L^2} + \frac{\#}{L^3} \cdot L + \frac{\#}{L^2} \end{aligned} \tag{11.3b}$$

The dimensions are interpreted as a *density gradient* $\beta_r = \nabla[N]$ rather than as the numbers per unit of a three-dimensional volume (L^3).

Next, let's use homogeneity of scope to examine each quantity and term in the equation. If measurements of density *[N]* and distance are made at a sequence of contiguous strips (each 5 km long and 0.2 km wide) along a transect of 50 km, the spatial range and resolution of both density *[N]* and radial distance r are the same:

$$\begin{aligned} [N] &= \beta_o + \beta_r \cdot r + \in \\ \frac{50\text{ km}}{5\text{ km}} &= \frac{50\text{ km}}{5\text{ km}} + 1 \cdot \frac{50\text{ km}}{5\text{ km}} + \frac{50\text{ km}}{5\text{ km}} \end{aligned} \tag{11.3c}$$

As a result, the parameter β_r has a scope of unity, which means there is no scale-up from the response on the left side to the explanatory variable on the right side of the equation.

However, the situation changes if 5 km counts are made at 10 km intervals along the transect, which is to say, only half the full 50 km transect is surveyed. In this situation the scope of the counts *[N]* is not the same as the scope of the distances, assuming that the estimate of the gradient in puffin density is taken to apply to the entire transect:

$$\begin{aligned} [N] &= \beta_o + \beta_r \cdot r + \in \\ \frac{50\text{ km}}{5\text{ km}} &= \frac{50\text{ km}}{5\text{ km}} + 2 \cdot \frac{50\text{ km}}{10\text{ km}} + \frac{50\text{ km}}{5\text{ km}} \end{aligned} \tag{11.3d}$$

The parameter β_r now has a scope of 2. This scope scales the data (at a limited resolution) up to the model (which applies to the entire transect, not just the measured sections).

It turns out that parameters perform several roles in data equations. One is to scale the units and dimensions of the explanatory (model) variable to the response (measurement) variable. Parameters act out this role in an equation according to the principle of dimensional homogeneity. Another role is to scale the explanatory variable to the same scope as the measurement variable. Parameters act out this role, too, according to the principle of homogeneity of scope.

ANOTHER LOOK AT SECTION 11.5

Write a regression equation similar to Equation 11.3a for two quantities of interest to you. Complete the calculation of spatial scope, as in Equation 11.3c.

11.6 The Scope of Natural Phenomena

Strictly speaking, scale and scope are characteristics of measurement activities, not characteristics of the natural phenomena that are the object of these activities. However, scale (and hence scope) are routinely considered characteristics of natural phenomena (e.g., Delcourt et al., 1983; Walker and Walker, 1991), an approach that is of utility in evaluating planned programs of measurement. This approach will work when continuously acting processes result in recognizable boundaries in space and time. Disease epidemics, for example, are phenomena that have an upper and lower limit on duration. They take a while to run their course but do not last forever, even though the processes that transmit disease continue to act. In the ecological literature, the scale of a phenomenon often refers to the upper limit in duration or extent of such phenomena. It can also refer to the lower limit. An equivalent pair of terms is the *minimum inner scale* and the *maximum outer scale* of a phenomenon. The terms *grain* and *extent* are also used (Wiens, 1989). "Scale" can also refer to a typical value of a phenomenon (Powell, 1989; Wiens, 1989; Rahel, 1990; Turner and Gardner, 1991). The spatial scale of a midlatitude storm system is on the order of 2000 km. Storm systems have a time scale (lifetime) on the order of a week. This usage, implying a characteristic value, applies to phenomena that fall within a fairly narrow range of values, such as storm durations or disease outbreaks.

Assigning a single characteristic value does not work for phenomena that range widely in duration or spatial extent, such as the spatial scale of bird migration. The ratio of the maximum relative to minimum duration or extent of a phenomenon is more informative than a single value. This ratio, called the *scope of a natural phenomenon,* is defined as the ratio of the outer to the inner scale, or equivalently, the ratio of extent to grain. For example, El Niño events occur at a frequency on the order of once every five years or so. The time between events can be as much as eight years but cannot be less than two years, the time typically taken to build up the global-scale pressure gradients that create these events. Thus El Niño return times have a scope of 8 years/2 years = 4. The statement of the scope is useful in avoiding the implication that El Niño events occur with a periodicity of five years.

If a phenomenon is episodic, with discernable starting and ending points, its temporal scope can be defined as the ratio of duration of longest and shortest cases. Examples are snowstorms, earthquakes, fires, mass extinctions, epidemics, foraging bouts, and reproductive seasons. If a phenomenon has definable boundaries in space, its spatial scope can be defined as the ratio of diameters of the largest and smallest cases. But not all phenomena have definable boundaries or clear start and end points. Phenomena without strong spatial gradients or rapid onsets and ends generally cannot be characterized as to scale or scope. Examples are changes in population size or the flow of energy and material in ecosystems.

Natural phenomena are routinely assigned a "scale" or scope, but any such assignment will depend on the words we happen to have. Most speakers of English get by with "snow" and "ice", but people who live under the cold thumb of the Labrador Current have a richer vocabulary of wintry terms. *The Dictionary of Newfoundland English* (Story et al., 1982) distinguishes several forms of snowfall (*dwy, scad, snowing by the reeves*), even more forms of ice forming on objects (*glitter, silver thaw, ballicatter, black ice*), and yet more forms of ice at the sea surface (*clumper, growler, lolly, pinnacle, sailing pan, sish, tabby*). Many of these phenomena have distinctive and limited scopes. For example, *ballicatter* (ice formed by seaspray and waves) is limited to the fringe along the coast. A *dwy* (snow squall) characteristically starts suddenly and lasts a short time, usually less than an hour, rarely more. We can assign a scope to phenomena in this vocabulary a lot more easily than to the general terms "snow" or "ice".

The scale of phenomena, as the term is used in the literature, usually refers to comparison of cases (e.g., typical hurricane diameter) and hence is based on noniterative scaling, as in Equations 2.1 and 2.2 in Chapter 2. Individual cases can be assigned a scope using iterative scaling (more than one measure of the same object), as in Equations 2.3 and 2.4 in Chapter 2. For example, a field of ripples in a sandbar can be assigned a scope, the ratio of the extent of the field to the size of an individual ripple. At smaller scales, the ripple field becomes the geometry of sand grains rather than ripples. A particular sandbar can be assigned a cross-shore scope, the ratio of the distance across the bar to the size of a sand grain, at about 0.2 mm. At smaller scales, the geometry of a sandbar becomes the geometry of sand grains and interstices.

Many phenomena seem at first to be infinitely divisible and hence with no lower limit and no scope. Further thought about the phenomenon often uncovers a lower practical limit. A theoretical tortoise takes infinitely small steps, but a living tortoise lurches forward a certain distance after it exerts enough force to overcome the drag of its shell against the ground. Thus the velocity of a tortoise is not infinitely divisible with an infinitely fine

resolution. At a sufficiently fine resolution, the velocity of a tortoise becomes a series of events. The scope of tortoise velocity then becomes the maximum velocity relative to the minimum velocity set by a single lurch forward. In general, phenomena are not infinitely divisible and so do not have infinite scopes.

ANOTHER LOOK AT SECTION 11.6

Name two ecological phenomena of interest to you. Try stating the scope (upper and lower limits) based on your knowledge of the dynamics of the phenomenon. If that's not possible, state why.

11.7 Scope Diagrams

A simple and effective way of comparing the scope of ecological phenomena is with a diagram. A *scope diagram* uses one, two, or even three logarithmic scales to display and compare scope calculations. On a logarithmic scale, the scope is the distance between two points on a graph. Ecological phenomena are typically displayed relative to both space and time axes. Many of the conceptual space/time diagrams in the literature are scope diagrams since they display both an outer and an inner scale with respect to space and time. However, not all space/time diagrams are scope diagrams. For example, the first conceptual space/time diagram (Figure 2.4) depicted the space and time scales for patchiness of phytoplankton, zooplankton, and fish as points (single values), not as scopes. The accompanying instrumental space diagram displayed scopes. Space/time diagrams have appeared with exponentially increasing frequency since 1978 (Figure 2.4), and many of these are scope diagrams. Figure 2.1 is a space/time diagram that shows the scope of the problem of the collapse of cod stocks in the western North Atlantic. Figure 2.2a displays the scope of the problem of monitoring the effects of chronic pollution on the fauna in a harbor in New Zealand.

Space/time diagrams have been constructed for a variety of phenomena, including earth system processes (NASA, 1988), landform generation (Swanson et al., 1992), ocean circulation structures (Steele, 1991b), atmospheric structures (Dickinson 1988; Shugart et al., 1988), climate change (Hobbs, 1998), aquatic ecosystem interactions (Schindler, 1988; Ray, 1992; Hobbs, 1998), population dynamics in groundfish (Langton et al., 1995), variance in species richness in marine and terrestrial systems (Marquet et al., 1993), patch dynamics in forests (Shugart and West, 1981; Deutschman et al., 1993), biomass dynamics of forests (King, 1991), disturbance regimes (Delcourt et al., 1983; Walker and Walker, 1991), vegetational patterns (Delcourt et al., 1983), human impacts on ecosystems (Powell, 1989; Walker and Walker, 1991; Hobbs, 1998), ecosystem responses to disturbance (Delcourt et al., 1983), organizational levels in ecology (Lugo, 1996), human land-use practice (Hobbs, 1998), and even vehicular movement (Gadgil, 1996).

The strength of these diagrams lies in their effectiveness in comparing phenomena of vastly different extents in space and time. The weakness of these diagrams lies in their inconsistent translation of numbers to graphical objects and frequent failure to convey

whether a scope is being displayed or whether the intention is to display only the extent of a phenomenon. Most of these diagrams use two-dimensional objects—circles, squares, and so on. The implication is that a scope is being displayed. In some cases it appears that a shape has been used to display a single number, such as an extent. The key to accurate representation is to display points where a single value is intended, lines where a scope is intended, and shapes encompassing the lines (scopes) used to construct the shape. These practices are illustrated by a narrative accompanying the construction of Figure 11.3.

The annual migrations of birds is astonishing relative to the distances that most people move every year, even with the aid of cars, trains, and planes. Surprisingly, some of the longest of all migrations are undertaken by species the size of one's hand. Some of the longest-distance migrants belong to the sandpiper family (*Scolopacidae*), a group that typically inhabits the open spaces of tundra and beaches. A 60 g Sanderling (*Calidris alba*), caught on a beach in Chile, will return to that beach year after year from its high Arctic breeding site. The sandpiper family also includes species that undertake little or no migration. This diversity in migratory extent is surprising in its own way, for the species in this

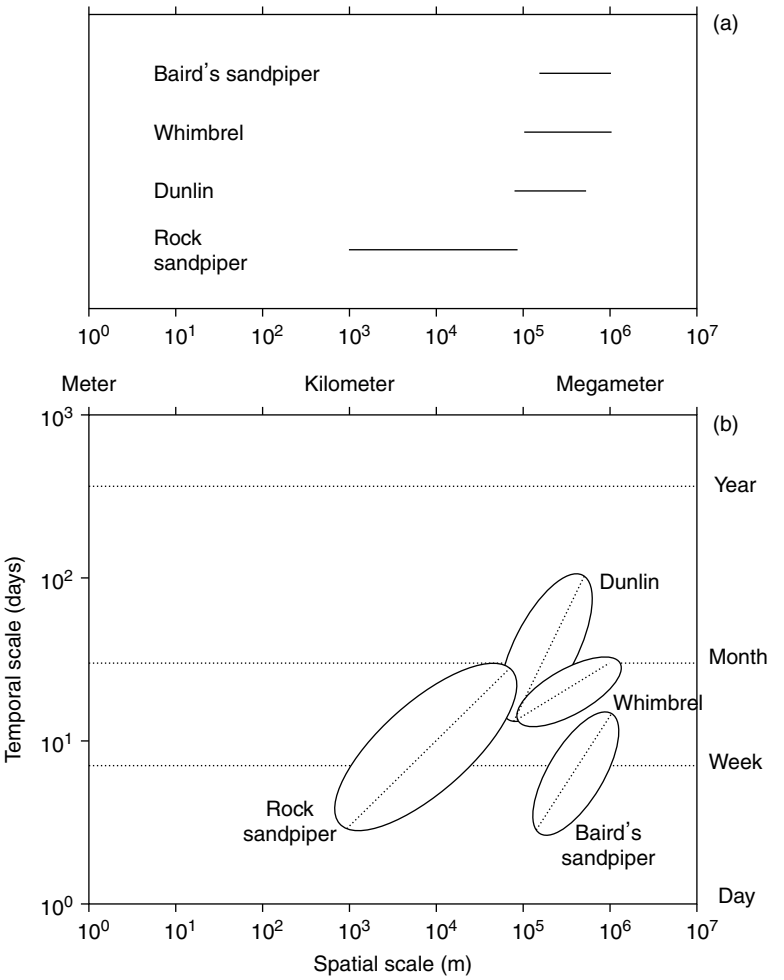


FIGURE 11.3 Scope of Migration in Four Species of Sandpiper; (a) Spatial scope of migration; (b) Space-time diagram for scope of migration.

family are so similar in shape and wing form as to be difficult to distinguish. Perhaps only the quirky nature of historical phenomena can explain why one species would migrate 10,000 km or more while a nearly identical species shifts slightly or not at all away from its breeding grounds in winter. This diversity in the scale of migration can be quantified in a diagram that lends itself to comparative analysis of the phenomenon.

Four species of sandpiper were chosen for graphical comparison. All breed in northern Alaska. All contrast strongly in migratory scope. The least migratory of the four is the rock sandpiper (*Calidris ptilocnemus*), which breeds on islands in the Bering Sea, then winters from the adjacent Alaska peninsula southward to Oregon. The next strongest migrant is the dunlin (*Calidris alpina*). Dunlins breed from Pt. Barrow south past the Yukon delta, wintering from Puget Sound to northern Baja California. The whimbrel (*Numenius phaeopus*), a still stronger migrant, breeds south to the Yukon delta, wintering from central California to southern South America. Another strong migrant, Baird's sandpiper (*Calidris bairdii*), breeds along the Alaskan coast southward to the Yukon delta, wintering in the Andes Mountains.

Distributional maps in the field guide by Robbins, Bruun, and Zim (1983) were used to construct a diagram showing the migratory scope of these four sandpiper species. Figure 11.3a shows the minimum distance between breeding and wintering grounds (left end of each line), together with the maximum distance between breeding and wintering grounds (marked by the right end of the line). The length of each line represents the scope of migration, that is, the ratio of the maximum to minimum migratory distance.

The duration of the migration period differs among these species. Some accomplish migration during short periods of time (whimbrel, Baird's sandpiper). Others, such as the dunlin, take several months to reach their wintering grounds. A space and time diagram displays these contrasting migratory patterns (Figure 11.3b). This diagram shows the scope for the species based on the distance from the northern edge of the breeding ground to the southern edge of the wintering grounds. The scope of individual migration would be less if birds from the northern part of the breeding range migrated only as far as the northern part of the wintering range. However, many sandpiper species migrate in a "leapfrog" pattern, where individuals at the northern limit of the breeding range migrate to points further south than birds from the central or southern part of the breeding range (Hale, 1980). The scope for individuals will thus lie along the rightward rising lines in Figures 11.3b, with some individual variation around the lines.

Scope diagrams need not be limited to just two logarithmic scales. A third logarithmic scale can bring out interesting contrasts. The four species depicted in Figure 11.3 differ in body size. The scope of body size (as a skeletal measurement) differs hardly at all within these four species. However, the scope of body size (as mass) can have a large scope in long-distance migrants, which are capable of storing substantial fractions of their normal weight as fat prior to long-distance migration. Extension of Figure 11.3 to a third axis, body mass, would result in four elliptical objects of similar thickness (with scope of body mass defined as the ratio of largest to smallest weight of an individual bird during a single year).

Most of the scope diagrams found in the ecological literature are space/time diagrams. However, any ratio scale quantity can be used in a scope diagram. For example, Swanson et al. (1992) used power (energy per unit time) to describe disturbance dynamics relative to spatial extent. In the case of bird migration, the power requirements of long-distance migrants are considerable, so a scope diagram with an axis of energy per unit time would bring out functionally significant differences not evident in a plot relative to time.

One of the values of a scope diagram is that it can be constructed from limited data and a small number of calculations. More detailed information can be added to the diagram as it becomes available. For example, a poorly known quantity such as seedling recruitment $[\dot{N}]$ of trees in tropical forests can be sketched into a scope diagram as an educated guess, perhaps on the order of $[\dot{N}] = 10$ seeds ha^{-1} day^{-1} with a temporal scope (1 month event^{-1})/(1 day event^{-1}) = 30. The spatial scope might be (1000 ha event^{-1})/(1 ha event^{-1}) = 1000. As data become available on several species, the scope diagram can be redrawn to show the scope of seed dispersal of a variety of species. The diagram, as it develops, indicates the current state of knowledge of the scope of seed dispersal.

Once a scope diagram has been constructed for phenomena, the next logical step is to compare it to the scope of research programs to investigate the phenomena. The scope of a research program is typically displayed relative to both space and time axes. This produces an instrumental space/time diagram such as the one in Figure 2.4. Comparison of the instrumental diagram with a conceptual diagram brings out mismatches between research programs and natural phenomena. Scope diagrams are thus a tool for the evaluation and revision of research programs.

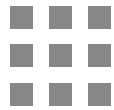
In this chapter we examined the scope of the components of research programs: measurements, the instruments used to make measurements, the quantities used to make calculations and draw conclusions about natural phenomena, and the equations (models) used in environmental studies. In Chapter 12 we will examine the scope of entire research programs.

ANOTHER LOOK AT SECTION 11.7

Sketch a space/time diagram for an ecological question concerning human effects on the environment.

Defined Concepts and Terms for Review and Future Reference

____ density gradient
____ homogeneity of scope
____ maximum outer scale
____ maximum inner scale
____ scope
____ scope diagram
____ scope of a parameter
____ scope of a quantity
____ scope of a single measurement
____ scope of an instrument
____ scope of natural phenomena



12 The Scope of Research Programs

The scale of resolution chosen by ecologists is perhaps the most important decision in the research program, because it largely predetermines the questions, the procedures, the observations, and the results. … Many ecologists … focus on their small scale questions amenable to experimental tests and remain oblivious to the larger scale processes which may largely account for the patterns they study.

—P. D. Dayton and M. J. Tegner, from *A New Ecology: Novel Approaches to Interactive Systems*; © 1984 John Wiley and Sons, Inc. Reprinted by permission.

12.1 Synopsis

Research programs have a spatial scope, which is the ratio of the extent to the spatial resolution. Similarly, the temporal scope is the ratio of the duration to the temporal resolution. Spatial and temporal scope are dimensionless ratios that can be partitioned into components that reflect the design of the program and the effort at each level in the program. Scope calculations quantify the scale-up at each level in a research program. These calculations are conveniently displayed as diagrams to evaluate research programs relative to one another and relative to the phenomenon being investigated. The magnitude of scale-up at each level differs substantially, as does the logical and evidential basis for scale-up.

Scope diagrams are relatively simple for descriptive studies, surveys, and monitoring programs. The scale-up from the directly measured area to the area of interest is often large. Descriptive studies rely more on judgment than on statistical inference to scale to larger areas. Surveys, which are important in applied ecology, notably wildlife management and fisheries, rely on statistical inference to estimate a quantity over a large area. Monitoring programs have the goal of measuring the degree of deviation from a norm, set by regulatory standards or derived from predictions in the impact statement. The magnitude of scale-up can be quantified for an entire survey, then partitioned among levels in a survey. The magnitude of scale-up at each level depends on the number of units that accumulate by replication at that level and on the ratio of unmeasured to measured units at that level. The basis for scale-up varies across studies

doi: 10.1016/B978-0-12-627865-1.00012-X

and can vary within a study. Scale-up can be by judgment that sites are representative, by informal inference based on haphazard sampling, by formal statistical inference in a finite spatial frame, by a descriptive spatial model, or by a spatial model with theoretical content.

Research programs based on manipulative experiments have become increasingly prominent in ecology. Scope diagrams display how experiments compare one to another and to the phenomenon being investigated. Fisherian experiments use randomization, replication, and local control to address problems of spatial heterogeneity. In agroecosystems, scale-up is by statistical inference and judgement based on knowledge of the experimental site. In other ecosystems, scale-up from Fisherian experiments is limited by several factors: lack of replicate units, substantial heterogeneity among units, and the inevitable increase in spatial heterogeneity with increase in area. Where Fisherian experiments are severely constrained or not possible, scope calculations provide a logical framework for the evaluation of uncertainty. By partitioning the scope, we can identify appropriate ways of addressing uncertainty at each of several levels in a study. The relevance of experiments to ecosystem scale processes depends on their being embedded in maps derived from surveys or computational models.

Computational models typically have a large temporal scope due to the small cost of high temporal resolution over long durations. At the same time they have small spatial scopes due to the cost of computing dynamics across a large number of spatial units.

Scope diagrams allow simultaneous display of the structure of integrated programs that combine experiments, surveys, time series, and computational models.

12.2 The Scope of a Set of Measurements

Measurements are taken within organized research programs that range from simple surveys and experiments to complex combinations of several methods. Measurements in the program will have several temporal attributes (Chapter 7.2), notably the time required to complete a single measurement (duration), the time between successive measurements (temporal lag), and the time between first and last measurement (temporal extent). The *temporal scope of a set of measurements* will be defined as the ratio of the temporal extent of each measurement relative to either the duration or the lag.

A set of measurements will also have spatial attributes (refer back to Section 7.3). In one dimension the attributes are the maximum distance between two measurements in the set (spatial extent or range), the distance between two measurements (spatial lag), and the linear extent of the completely measured unit (minimum spatial resolution or grain). In two dimensions (Section 7.4), the attributes are the entire area from which the set was taken (extent), the area around each measurement (lag), and the area for a single measurement (resolution expressed as an area). The *spatial scope of a set of measurements* will be defined as the ratio of the spatial extent relative to either the lag or the grain.

The duration and minimum spatial resolution are components of the spatial and temporal support, as defined by geographers. The *spatial support* is defined as the *n*-dimensional volume from which the values of a variable distributed in space may be computed. The complete specification of the support includes the geometrical shape, size, and orientation of the volume (Olea, 1999). *Temporal support* is analogous to spatial support.

Box 12.1 shows the calculation of the overall scope of a reconnaissance study, often the first step in ecological research. Eberhardt and Thomas (1991) classify such studies as sampling for pattern. Such studies typically serve as the basis for subsequent studies of the processes that generate patterns. An example of sampling for pattern is characterization of the flora of eccentric bogs in the state of Maine (Davis and Anderson, 1991). True bogs (including eccentric bogs) consist of peat-containing wetlands whose growing surfaces acquire minerals from the atmosphere rather than from underlying or adjacent mineral soils. Bogs cover extensive areas in the northern hemisphere, and so carbon sequestration (as peat) and release (as methane) can affect global carbon balance (Almquist-Jacobson and Foster, 1995). True bogs have slightly convex surfaces due to continued plant production on top of peat capable of holding water above the surrounding water table. Eccentric bogs are unusual in that the highest point is at the edge. These bogs slope downward from a high point resting against the side of a valley; they usually sit above a fen (a wetland with a flat or concave growing surface). They are numerous in Russia, Canada, and northern Europe, but only 35 eccentric bogs are known in the United States, all of which are in eastern Maine (from 45° to 46°40′N, from 67°30′ to 69°W). Of these, 15 were visited between June 30 and August 18, 1987 (Davis and Anderson, 1991), to ground-truth aerial photography, take samples of water and peat from chemical analyses, take cores for stratigraphic analysis, and make a standard vegetational reconnaissance (Mueller-Dombois and Ellenberg, 1974) using a 5 m by 5 m area called a *relevé*. Plant abundance by individual species was recorded in seven categories within four vertical strata ranging from ground cover (0.1 m) to tree (greater than 5 m). Analysis of the data uncovered a relation between vegetational types and chemical trophic gradient. No rare species were found. Based on this information, the authors of the study recommended 10 bogs for designation as critical areas based primarily on their unique geomorphic/hydrological characteristics.

The study was unusual in that all examples of the ecosystem in a defined area were identified. Nearly half (15 of 35 eccentric bogs) were visited and measured. The study was also unusual in reporting the duration as well as spatial extent of the measurement effort. Consequently, it was possible to calculate the spatial and temporal scope of this study from information in the published report.

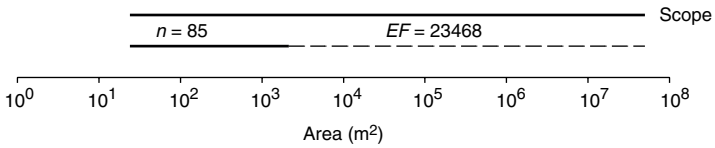
The overall spatial scope is computed relative to the extent of the completely measured or first-level unit, which is a component of the spatial support. In the bog study the first-level unit is the relevé, a square area 5 m on each side. The choice of spatial extent in computing the overall scope depends on the assumed generality of the study. In the bog study the generality certainly extends to the area of the 15 measured bogs and to the time from start to end of the study. The spatial scope of the 15 bogs relative to the relevé is $2 \cdot 10^6$. The overall scope could also be computed relative to the area of all 35 bogs if the 15 measured bogs are considered representative. Davis and Anderson (1991) did not report the area of all 35 bogs, so the area listed at this level (Box 12.1) is that of 15 bogs rather than all 35.

The overall temporal scope is computed relative to the duration of the first-level unit, which is the relevé. Each relevé required approximately half an hour to record vegetation cover. The choice of temporal extent from which to compute the scope depends again on assumptions concerning generality of the study. In this example the temporal extent was taken as 50 days, the time from start to end of the field work. The temporal

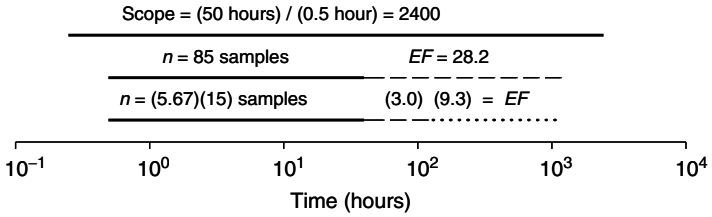
Box 12.1 Scope Calculations for Vegetational Reconnaissance of Eccentric Bogs in Maine (Davis and Anderson, 1991)

			A_o	A	Scope	T_o	T	Scope
Calculation of overall scope			$(5\,m)^2$	4987 ha	$2.0 \cdot 10^6$	30 min	50 days	2400
Level	Replication	Units	A_o	A	Scope	T_o	T	Scope
Relevé		1	$(5\,m)^2$			30 min		
Bog	5.67	5.67	$142\,m^2$	332.5 ha	23468	2.8 hr	8.6 hr	3.0
Inventory	15 bogs	85	$2125\,m^2$	4987 ha	23468	42.5 hr	50 days	28.2

Spatial scope $= 2.0 \cdot 10^6$ possible samples
Units $n = 85$ samples taken
Expansion factor $EF = 2 \cdot 10^6/85 = 2.35 \cdot 10^4$ possible samples/samples taken
Sampling fraction $SF = EF^{-1} = 4.26 \cdot 10^{-5}$ samples taken/possible samples



Temporal scope $= 2400$ possible samples
Units $n = (5.67)(15) = 85$ samples taken
Expansion factor $EF = 2400/85 = 28.2$ possible samples/sample taken
$EF_{bog} = (8.6\,hr)/(2.8\,hr) = 3.0$
$EF = EF_{bog}(T_{inventory}/T_{bog}) / (n_{inventory}/n_{bog})$
$= 3.0\ (50\ days/\ 8.6\,hr) / (85\ /5.67)$
$= 3.0\ (139)\ / (15)$
$= 3.0\ (9.3)$



extent could have been taken for a longer timeframe. This would assume that bogs change so slowly that the same result would have been obtained no matter when the study was completed within the longer time frame. Rates of vegetational change in undisturbed bogs are slow enough that the temporal extent might be taken as several years.

Most ecological studies have at least one intermediate level in addition to the level of the unit of measurement and the level of the study. In the bog study, there is one additional level, that of the bog. Spatial and temporal scope can be calculated for each level. The calculations begin with a listing of each level. For the example at hand, these were the relevé, the bog, and the inventory of all 15 bogs (Box 12.1). The analysis of scope table is then completed from left to right, under the headings shown in Box 12.1 (Level, Replication, Units, A_o, A, Spatial Scope, T_o, T, and Temporal Scope). The replication is the number of units at each level. In this example the replication was 5.67 relevés per bog and 15 bogs per complete inventory. The replication is multiplied by the number of units at the next lower level to obtain the number of first-level units taken at each level. In this example there were 5.67 relevés per bog and 85 relevés in the entire inventory (Box 12.1). The spatial effort at each level (A_o) is the product of first-level

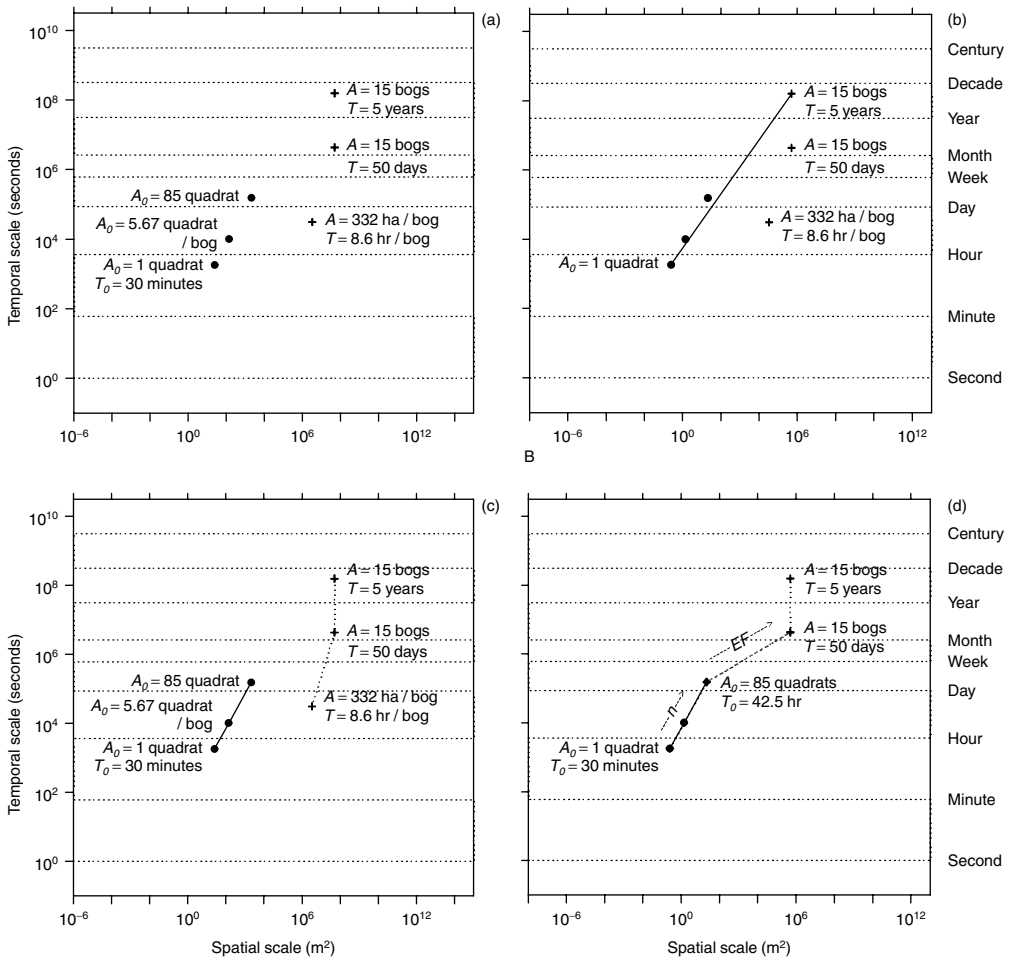


FIGURE 12.1 Scope Calculations for Relevé Analysis of Vegetation Structure. Data on eccentric bogs in Maine from Davis and Anderson (1991); (a) Spatial and temporal attributes of the study, from Box 12.1; (b) Overall spatial and temporal scope; (c) Scale-up by accumulation of samples shown as solid line connecting circles. Scale up by informal inference shown as crosses connected by dotted line; (d) Sampling effort *n* and expansion factor *EF*.

units and the area of the first-level unit. In this example the effort was $A_o = 142\,m^2$ per bog and $A_o = 2125\,m^2$ for the entire inventory. As in many ecological studies, the samples together cover an area A_o that is only a fraction of the total area at the same level. In this study the total areas were $A = 332\,ha$ per bog and $A = 4987\,ha$ for the entire inventory of 15 bogs. The spatial scope was A/A_o, the area relative to the area directly measured within A. The scope is $332/0.0142 = 2.3 \times 10^4$, which means that on average, bog area was over 23,000 times greater than the area covered by direct measurement within a bog. Because averages were used, the same scope applies at the level of the entire inventory: $4987/0.2125 = 2.3 \times 10^4$.

Sampling effort as a duration T_o takes place within a larger timeframe T. Temporal effort T_o is the product of duration of first-level units (relevés) and the number of first-level units. The temporal effort was $T_o = 5.67 \cdot 1/2$ hour $= 2.8$ hours at the level of each bog; it was $T_o = 85 \cdot 1/2$ hour $= 42.5$ hour at the level of the entire inventory. The temporal frame T at each level depends on the assumptions concerning generality of the results. In this study the temporal frames were taken as the time to complete the survey of a single bog (8.6 hours, on average) and the time to complete the inventory (50 days). The ratio T/T_o gives the temporal scope at each level. Because of the time intervening between each bog survey, the scope at the level of a bog differs from that at the level of the entire inventory.

Scope computations are conveniently displayed in a space/time diagram. Figure 12.1 displays the scope computations for the bog study. Figure 12.1a shows, as filled circles, the spatial effort A_o and temporal effort T_o at each level. A small cross marks the spatial frame A and temporal frame T to which these measurements apply. A line connecting any two points will be a scope. Figure 12.1b displays the overall scope as the distance from the circle at the lower left to the cross furthest above and to the right.

Scope diagrams such as Figure 12.1 are readily constructed in any graphics program. Spatial units can be added as labels along the top border of the diagram, but they're not as necessary for interpretation as are temporal periods. In general, standard multiples of length (10^1 km, 10^3 km) or area (km^2, km^4) will suffice. Other diagrammatic material can be added either as text pasted onto the graph or as lines plotted between points, using the same technique as the plot of time periods shown in Box 12.2. In Figure 12.1d, the arrow showing the expansion factor *EF* was added as a text box: —*EF*—> . The line connecting the area of 85 relevés to the area of 15 bogs was plotted as a line connecting two points. In general, if an approximation is desired (e.g., the scope of several models), a pasted shape adequately represents the graphical component. If exact values are available (e.g., the scope of a particular model; see Figure 12.4 later in this chapter), a plot more accurately represents a graphical component such as *EF*. Each model in Figure 12.4 was plotted as a line connecting four known values, where the four values were listed in such a way as to achieve a box.

Box 12.2 Constructing a Scope Diagram: Example Based on Calculations in Box 12.1

1. Calculate overall scope.

A_o	A	Scope	T_o	T	Scope
$(5\,m)^2$	4987 ha	$2.0 \cdot 10^6$	30 min	50 days	2400

2. Decide on base unit (e.g. seconds and m^2 in Figure 12.1). Length can be used instead of area.
3. Convert to base units if necessary and enter A_o, A, T_o, and T in four columns.

   Here are the calculations in Box 12.1, as they appear in a spreadsheet:

		Area (m^2)		Time (Sec)			
Level	**Repl**	A_o	A	T_o	T	**Scope**	**Scope**
		Support	Extent	Support	Extent	A/A_o	T/T_o
m^2/relevé		25		1800			
Relevé/bog	5.67	141.67	3324667	10080	30960	23468	3.1
Bog/inventory	15	2125	49870000	153000	4320000	23468	28.2
All bogs	35		49870000		116363333		

4. Decide on time and space ranges for the graph. In Figure 12.1 the spatial range is 10^{-6} to $10^{15}\,m^2$, the temporal range from 1 to 10^{10} seconds.
5. Using logarithmic axes, plot T_o against A_o as one series, T against A as another series.
6. Add interpretable time units (hours, days, etc.). These can be added as labels at appropriate points on the vertical axis, but a more accurate presentation is to plot these units as lines, as in Figure 12.1. A simple method is to plot the Area and Time as shown in the accompanying table, as a third data series in the graph. Note the alternation between the maximum and minimum spatial units compared to the progression in time units, to produce horizontal lines.

Horizontal Lines

Labels	Area	Time
Sec	0.000001	1
Sec	1E + 15	1
Min	1E + 15	60
Min	0.000001	60
Hour	0.000001	3600
Hour	1E + 15	3600
Day	1E + 15	86400
Day	0.000001	86400

ANOTHER LOOK AT SECTION 12.2

In a study familiar to you, list the unit area and duration (support), the number of samples at each level of the design, and the spatial and temporal extent of the study. Compute the overall spatial and temporal scope.

12.3 Partitioning the Scope

Scope calculations quantify the magnitude of spatial and temporal scale-up for the entire study and at each level in a study. The *overall scope* is the number of samples that could have been taken. This overall scope has two components: scale-up via accumulation of samples (Figure 7.3) and scale-up via inference from these samples to all potential samples. In the case of the bog reconnaissance, 85 units were sampled, each representing $2.35 \cdot 10^4$ times its own area. The factor that magnifies or scales up the results from sampled units to the area of interest is the *expansion factor*, defined as the ratio of the potential number of samples (overall scope) relative to the number of samples taken (Cochran, 1977). Box 12.1 shows the partitioning of the overall spatial scope into samples taken and an expansion factor. Another way of looking at the expansion factor is to compute its inverse, which tells us the fraction of the possible samples that were taken. The *sampling fraction* (Cochran, 1977) is defined as a ratio: the number of samples taken relative to the potential number of samples. The sampling fraction for the bog reconnaisance is $23{,}468^{-1} = 4.26 \cdot 10^{-5}$. This fraction is small but by no means unusual. It is roughly the same as taking 15 plots, each 10 m by 10 m, to represent forest production dynamics in the 3000 ha Hubbard Brook study (Bormann and Likens, 1979). Many ecosystem studies cover larger areas and have even smaller sampling fractions.

More generally, spatial and temporal scope can be *partitioned* according to any of several levels of a study. Box 12.1 shows the partitioning of the temporal expansion factor ($EF = 28.2$) into two components: the scale-up from relevé to bog ($EF_{bog} = 3.0$) and the scale-up from bog to inventory ($EF = 9.3$). The overall scope, a factor of 2400, is thus the product of two scale-ups via accumulation ($n = 5.67$, $n = 15$) and two expansion factors ($EF = 3.0$, $EF = 9.3$). These two scale-up factors are based on judgment and experience. The investigators assumed that these scale-up factors apply because past experience shows that little or no change in bog flora is to be expected during the 50-day duration of the study. The results based on 42.5 hours of effort apply to the 50-day duration of the study. The results could be taken as representative of several years based on the rates of vegetational change in bogs in the absence of human disturbance or change in the hydroregime. If the study were considered applicable to a five-year period, the scope would be (5 years) (365 days/year) (1440 minutes/day)/(30 minutes) = 87600. This is partitionable into two measured components ($n = 5.67$ relevés, $n = 15$ bogs) and three inferred components ($EF = 3.0$, $EF = 9.3$, and $EF = 5$ year/50 day = 36.5). The rationale for the first and second expansion factor is that vegetation did not change during the study, and hence the relevé samples could have been taken in any order. The rationale for the third expansion factor is that the vegetational change is slow enough that the study would produce the same results no matter when it was done in a five-year period.

The spatial scale-up has four components: accumulating relevés within a bog to increase the area measured (a factor of $n = 5.67$ relevés/bog, on average), accumulating bogs (a factor of $n = 15$ bogs), inference from relevés to bogs (a factor of 23468), and inference from the inventory to all bogs (a factor taken as 1). The basis for the spatial scale-up from relevé to bog was informal inference. The investigators exercised judgment in placing the relevés so as to be typical. Formal inference, which would require random placement of each relevé, was not used. There was no scale-up from inventory to all bogs and hence the factor is unity. If the area of all 35 bogs had been reported, the scale-up could have been calculated. The basis for this scale-up would be informal inference, based on judgment that the 15 bogs measured were representative of all 35 bogs.

The spatial and temporal scale-up at each level in the study can be displayed by connecting each value of effort (marked as a circle in Figure 12.1a) with its corresponding extent (marked with a cross). This display tends to be cluttered and hard to interpret. A simpler display results from connecting the effort (circles) with a solid line and connecting the study design (crosses) with another line (Figure 12.1c). The separation between the two lines shows the inferred scale-up. The values in the graph can also be connected in a fashion that displays the sequence of scale-ups encompassed in the overall scope of the study (Figure 12.1d). The solid lines show the scale-up achieved by accumulating samples (n); the dotted lines represents the scale-up (EF) from samples. This display is the two-dimensional version of the diagram in Box 12.1. These instrumental scope diagrams will prove useful in designing and comparing research programs.

ANOTHER LOOK AT SECTION 12.3

Recompute the spatial scopes in Box 12.1, assuming an area of 11636 ha (instead of 4987 ha) at the level of a survey of all 35 bogs. Show that the inferred spatial scale-up (expansion factor) is partitionable into two components, of which one is 35/15.

12.4 The Scope of Surveys

The purpose of a formal survey is to obtain an accurate estimate of a quantity, usually summarized as a mean or as a proportion. An *exhaustive survey* enumerates all cases, a procedure that guarantees an accurate value. An example is a census of all nests of a particular bird species in a nature reserve, to obtain local population size. A *probabilistic survey* uses randomly selected units to produce an accurate (unbiased) estimate. An example is the number of nests in a reserve, estimated from complete enumeration within randomly chosen blocks of 1 hectare each. A probabilistic survey uses statistical inference, thus differing from the informal inferences of a descriptive study, described in the previous section. The uncertainty of the estimate from a probabilistic survey can be reduced by skillful design. Texts on survey design describe many effective techniques. These include grouping of units into strata, efficient allocation of sampling effort to strata or clusters, efficient allocation of effort among levels, and introduction of covariates to control effects statistically.

The technical literature on survey design (e.g., Cochran, 1977) provides two key terms: the unit and the frame. The *unit* is the item on which a measurement is taken. The unit can have natural boundaries. Examples of natural units are individuals (such as a corn plant), parts of individuals (such as the leaf of a plant), islands, and lakes. Units can also be defined by artificial boundaries. Examples of artificial units are plots in an agricultural survey. The *frame* is the list of all possible units. The frame in a survey of agricultural production might be an entire field, an entire farm, or even an entire district.

The temporal scope of a survey depends on the temporal attributes of the data. The *temporal scope of a survey* is defined as the duration of the study, relative to the duration of measurement of a single unit. One common objective in a survey is to minimize the temporal scope of the study so as to reduce or eliminate any change in the quantity being estimated. The goal of surveys often is to produce an estimate for a single point in time, hence of limited temporal scope.

The *spatial scope of a survey* depends on the spatial attributes of the data. The spatial scope of a survey is defined as the extent of the frame relative to extent of the unit. Extent can be in units of length (as in a transect study), in units of area, or in units of volume (as in sampling a body of water). Several factors set the spatial scope. The type of instrument often sets unit size. A typical rain gauge resolves precipitation at a spatial scale on the order of $10\,cm^2$. An easily deployed plankton net resolves copepod concentrations at a spatial scale on the order of $10\,m^3$ to $100\,m^3$, the volume swept along distances on the order of 10 to 100 m. The frame often depends on the applied or social goal of the survey. The frame and hence the scope of the survey is often defined by human activities. An example is an investigation of the effects of logging practices on wildlife. The scope of these studies, including control sites, will be determined by the spatial scale of the logging practices. Because the scope of interaction with human activities is often important, deciding on the spatial and temporal scope of a survey is frequently one of the most difficult parts of applied ecology.

Scope calculations and scope diagrams are useful in designing and comparing surveys. First, such calculations quantify the magnitude of scale-up from sample to the target of inference. An example is a simple random (Section 4.1) or stratified random survey (Section 4.2) in which the target of inference is a completely enumerated frame. Scope calculations are useful in comparing alternative designs (Section 4.3). Partitioning the overall scale-up into components allows evaluation of the type of inference at each level in hierarchically structured surveys (Section 4.4). Scope calculations display the structure of monitoring programs and the type of inference at each level (Section 4.5).

12.4.1 Simple Random Surveys

The goal of a simple random survey is to produce an accurate estimate of quantity. To achieve this goal, samples are chosen randomly from a list (frame) of all possible units. All units have an equal probability of being sampled, resulting in an unbiased estimate for the entire frame. An example is the density of Icelandic scallops *Chlamys islandica* on St. Pierre Bank, located in the northwest Atlantic. An accurate estimate of population size is needed, along with an estimate of replacement rate, to set a sustainable harvest rate. A convenient symbol for scallop density is $[N] \equiv N/A$, where N is the count of scallops and A is the area of the count. The quantity $[N]$ can be defined at any of several spatial scales. It can be the number of scallops per 10 cm diameter area, the smallest area

occupied by an individual scallop of commercial interest. This minimum area is on the order of $5\,\text{cm} \cdot 5\,\text{cm} \cdot 3.14 = 78.5\,\text{cm}^2$. At this small scale *[N]* is a binomial quantity—it can either be 1 scallop $\cdot (78.5\,\text{cm}^2)^{-1}$, or it can be 0 scallop $\cdot (78.5\,\text{cm}^2)^{-1}$. The standard method for estimating scallop density is to use a dredge, which scrapes some fraction of scallops from a known area of seafloor. When defined at this scale, the quantity *[N]* is highly variable, ranging from zero to thousands per hectare. The quantity *[N]* can also be the total number of scallops on an entire offshore bank, which is the traditional spatial unit for assessing stock size and regulating harvest rates. At this scale *[N]* has a single value, which we are trying to estimate.

For the purpose of assessment of stock size at the scale of a bank, the spatial scope of the quantity *[N]* is the ratio of the area of the bank to the area occupied by a scallop of about commercial size. The two areas are expressed in different units, so a rigid conversion factor (number of square centimeters per square kilometer) appears in the calculation:

$$\frac{19{,}000\,\text{km}^2}{3.14\cdot(5\,\text{cm}^2)}\cdot\frac{10^{10}\,\text{cm}^2}{\text{km}^2} = 2.4\cdot10^{12}$$

The sampling unit for a dredge survey of the bank is a single haul by a 10 m wide dredge towed 1 nautical mile (1.852 km). The resulting dredge data has a support consisting of a strip with an area of $18520\,\text{m}^2$ and a duration of 6.7 minutes. The calculations are shown in Box 12.3. The extent of the survey is a 19,000 km bank; hence the spatial scope is 10^6. The frame of the survey is 10^6 possible sample locations on the entire bank. Randomly chosen sampling locations are separated by roughly two hours of steaming time, which allows the catch to be counted between hauls. On a research vessel with 12-hour workdays, this comes to six samples per day, or 180 samples during a 30-day survey of the bank. The sampling fraction is thus $1.8 \cdot 10^{-4}$. The expansion factor is 5700, which means that each sample will represent an area 5700 times as large (Box 12.3).

Box 12.3 Scope Calculations for Dredge and Acoustic Surveys on St. Pierre Bank, Northwest Atlantic

1. Dredge survey at a speed of 9 nautical miles/hour. Dredge is 10 m wide.
   Haul area $A_o = (10\,\text{m})(1\,\text{nmile})(1.852\,\text{km/nmile}) = 0.0185\,\text{km}^2$
   Haul time $T_o = (1\,\text{nmile})(60\,\text{min}/9\,\text{nmile}) = 6.7\,\text{min}$

Level	Replication	Units	A_o	A	Scope	T_o	T	Scope
Scallop			$\pi(5\,\text{cm})^2$					
Haul		1	$0.0185\,\text{km}^2$			6.7 min	2 hr	18
Survey day	6 hauls	6	$0.11\,\text{km}^2$			40 min	1 day	36
Survey	30 days	180	$3.33\,\text{km}^2$	$19{,}000\,\text{km}^2$	5700	20 hr	30 day	36

Spatial scope $= 19{,}000\,\text{km}^2 / 0.0185\,\text{km}^2 = 1.0 \cdot 10^6$ possible samples
Units $n = 6/\text{day} \cdot 30\,\text{days} = 180$
Expansion factor $EF = 10^6/180 = 5700$
Sampling fraction $SF = EF^{-1} = 1.8 \cdot 10^{-4}$

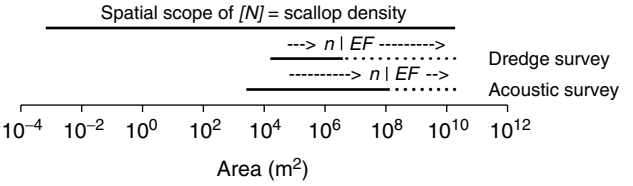
2. Acoustic survey at a speed of 9 nautical miles per hour. 1 sample/minute in 10 m wide swath.

Haul area A_o = (1 min)(9 nmile/60 min)(1852 m/nmile) = 2778 m^2

Level	Replication	Units	A_o	A	Scope	T_o	T	Scope
Sample		1	2778 m^2			1 min	2 hr	18
Survey day	1440	1440	4 km^2			1 day	1 day	1
Survey	30	43200	120 km^2	19,000 km^2	158	30 day	30 day	1

Spatial scope = 19,000 km^2/ 0.00278 km^2 = $6.8 \cdot 10^6$ possible samples
Units n = 1440/day · 30 days = 43200
Expansion factor EF = $6.8 \cdot 10^6$/43200 = 158
Sampling fraction SF = EF^{-1} = $6.3 \cdot 10^{-3}$

3. Comparison of dredge and acoustic survey.
Scallop density *[N]* on St. Pierre Bank at resolution of 1 scallop diameter
Spatial scope = (19,000 km^2)(10^5 cm/km)^2/ [(5 cm)^2(3.14)]
= $2.4 \cdot 10^{12}$ possible samples



Several sources of uncertainty beset these surveys. First, how efficient is the dredge? If the dredge scrapes up 50% of the scallops, the average catch per unit area scraped is divided by 0.5 to estimate the density. But if the dredge scrapes up only 10% of the scallops, the catch is divided by 0.1, resulting in a far larger estimate of population size. An even greater problem is the enormous spatial variability in scallop density, which generates considerable uncertainty regarding how close an estimate is to the true value. Increasing the sample size will reduce this uncertainty, but at the cost of more time at sea, which is expensive.

ANOTHER LOOK AT SECTION 12.4.1

State a quantity of interest to you, define units and frame for a simple random survey to estimate the mean value of that quantity, then make a rough estimate of the number of units in the frame and the expansion factor.

12.4.2 Stratified Random Surveys

The frame of a survey is often stratified into groups of units that share a common trait. A good stratification results in estimates of the parameter, that will differ among strata. If the units are natural (leaves, organisms), stratification will be by some trait measured on each unit. If the units are artificial within a finite spatial frame (relevés, quadrats, dredge hauls, etc.), stratification can be either by a measured trait or by contiguity. Stratification by contiguity results in estimates for subareas of the overall area. Stratification by contiguity is a form of coarse graining (Figure 7.3) and hence addresses questions of spatial scale.

One way of reducing uncertainty in an estimate is to allocate greater sampling effort to strata with highly variable catches. This reduces the variance in the estimate in those strata with high variance and hence will usually reduce the uncertainly in the estimate across the frame. This differential allocation of samples is called *stratified random sampling* (e.g., Cochran, 1977). The statistical literature on this technique is extensive. It is worth noting that when the variance increases with the mean (as it does for count data such as scallops), the effect of allocating more samples to areas of high variance will be to increase the effort in areas of high mean density. The scope of a stratified random survey will be the same as a simple random survey—the number of units in the frame.

ANOTHER LOOK AT SECTION 12.4.2

For the quantity, units, and frame you listed in the previous section, define strata and why you think these strata represent a major source of variance in the quantity of interest. Are your strata spatially contiguous?

12.4.3 Comparison of Survey Designs

Scope calculations, when displayed in a diagram, provide a quick and readily grasped comparison of alternative designs. The calculations in the diagram can be used to compare costs as well as gains in precision. The following example carries graphical comparison to quantitative comparison of two methods of surveying scallop density on St. Pierre Bank, in the western North Atlantic. The example is presented as what can be accomplished with scaling methods. It is well beyond the basics of spatial and temporal scaling in ecology.

Rapid development in acoustic technology now allows extensive and high-resolution delineation of scallop habitat (fine gravel for some species, coarser material for others). This information could be combined with local estimates of scallop density as a function of habitat, to make an estimate of stock size. The immediate advantage of acoustic measurement is its greater capacity—more data for the same amount of time at sea. This must be weighed against the unknown characteristics of the method. In particular, the relation of the acoustic signal to habitat type is unknown and would have to be determined during the study. Before undertaking a calibration study, it is worth calculating how much the instrument will increase the scope of the survey and, as important, how much it can potentially increase precision through greater sample size. To evaluate this, we compare the scope of the dredge based and acoustic-based survey to each other, within the scope of the quantity of interest.

The spatial support for the acoustic survey is a strip with less area than that covered by a dredge, because the acoustic device can integrate signals every minute in a 10 m wide swath. At a steaming speed of 9 nautical miles per hour, the area of each acoustic measurement is 2778 m^2 (Box 12.3). Because the unit area is smaller, the spatial scope (number of potential sampling sites) exceeds that of the dredge survey in the same area. Acoustic measurements can be taken continuously, day and night, resulting in over 43,000 measurements in a 30-day cruise. Each sample represents an area 158 times as large (Box 12.3). The sample fraction, $SF = 6.3 \cdot 10^{-3}$, exceeds that of the dredge survey.

Diagrammatic display of these calculations (Box 12.3) makes it evident at a glance that the increase in the scope of the acoustic survey is small, whereas the increase in sampling effort n is substantial. Consequently, the reduction in the inferred component EF is substantial, from a factor of 5700 for the dredge survey to only 158 for the acoustic survey. The calculations indicate that a further look at the acoustic survey is warranted.

How much of an increase in precision can we expect from the acoustic survey? We can expect a considerable gain, since the sampling effort increases by a factor of 43,200/180 = 240. Set against this is the change in variance accompanying the change to a unit with different spatial support (Dungan, 2001). We can apply scaling arguments to make a rough estimate of the change in variance (and change in standard error) due to change in support. Even though scallop counts will not be normally distributed around the estimate of the mean, it is reasonable to assume that means computed from 180 samples (dredge survey) or even more samples (acoustic survey) will be normally distributed around the estimate, with a standard error proportional to the square root of the variance in count var*(N)* and inversely proportional to the square root of the sample size n. Because our interest is in the standard error of the dredge survey at the extent of the acoustic survey, we resort to an estimate of change in variance (and hence standard error) with scale from the acoustic survey.

The change in standard error (or relative standard error) can be estimated from the measurement relation between variance and the resolution (grain) of the spatial unit in the acoustic survey as measured by coarse graining, provided that the spatial extent of the dredge survey falls within the scope of the acoustic survey. Acoustic surveys typically fall within the range of spatial resolutions possible with a dredge survey. This allows us to use the measurement relation between variance and resolution within the acoustic survey to estimate the standard deviation of the dredge survey at any of several resolutions, including the rather large scale (extent) of an acoustic survey. The computations (Box 12.4) show that the reduction in standard error due to increase in sampling effort will more than offset the potential increase in variance due to smaller measurement area (support). Scope calculations combined with measurement relations suggest that it would be worth putting effort into an acoustic survey.

The next step is a calibration study to measure how strongly scallop density is related to acoustically defined habitat. We would also need to consider how to scale from transects to the entire surface of St. Pierre Bank. Random placement guarantees that transect measurements are representative of St. Pierre Bank, at least in the long run. An adequately randomized survey is expensive and so unlikely to be run more than once. If it is run only once, one cannot escape the fact that one survey might not be representative temporally. It would be prudent to consider additional information, such as sand/gravel/cobble substrate in relation to water depth (available for the entire bank),

Box 12.4 Calculating the Relative Error from Sample Size and Unit Area

1. The formula for the standard error of the mean value of N requires an estimate of the variance var*(N)* in addition to sample size n.

$$\mathrm{StErr}(N) = \left(\frac{\mathrm{var}(N)}{n}\right)^{1/2}$$

2. Lacking an estimate of var*(N)*, we rewrite the formula as a scaling relation to obtain the relative error.

$$\frac{\mathrm{StErr}(N)_{acoustic}}{\mathrm{StErr}(N)_{dredge}} = \left(\frac{\mathrm{var}(N)_{acoustic}}{\mathrm{var}(N)_{dredge}}\right)^{1/2}\left(\frac{n_{acoustic}}{n_{dredge}}\right)^{-1/2}$$

3. The spatial variance var*(N)* of highly clumped marine organisms scales with spatial frequency f according to an exponent β that is usually less than 1 and sometimes near 0, as estimated by spectral analysis (Horne and Schneider, 1997).

$$\frac{\mathrm{var}(N)}{\mathrm{var}(N_{ref})}\left(\frac{f}{f_{ref}}\right)^{-1} = \left(\frac{f}{f_{ref}}\right)^{-\beta}$$

4. In the worst case, $\beta = 0$, which means that the variance increases with increasing frequency (i.e., at smaller unit sizes). The measurement frequency is the inverse of the area of the sampling unit. The acoustic device samples the sea floor 6.7 times more frequently.

$$\frac{f_{acoustic}}{f_{dredge}} = \frac{1/2778\ \mathrm{m}^2}{1/18520\ \mathrm{m}^2} = 6.7$$

5. Due to a higher frequency of measurement, the variance of the acoustic measurements could be 2.6 times greater than the dredge measurements.

$$\frac{\mathrm{StErr}(N)_{acoustic}}{\mathrm{StErr}(N)_{dredge}} = \left(\frac{18520}{2778}\right)^{1/2}\left(\frac{n_{acoustic}}{n_{dredge}}\right)^{-1/2} = 2.6\left(\frac{n_{acoustic}}{n_{dredge}}\right)^{-1/2}$$

6. The reduction in standard error due to increase in sampling effort more than offsets the potential increase in variance due to smaller measurement area (support).

in designing a study to estimate scallop abundance across all of St. Pierre Bank. The topic of scaling from point, transect, experimental plot, and survey measurements to an ecosystem (such as St. Pierre Bank) was addressed in a special issue of the *Journal of Experimental Marine Biology and Ecology* (Volume 216).

ANOTHER LOOK AT SECTION 12.4.3

If you have experience with a research project in which both surveys and field experiments were conducted, draw a diagram comparing the spatial and temporal scales of the survey(s) and experiment(s).

12.4.4 Monitoring Programs

The goal of a monitoring program is to measure the degree of deviation from a defined norm for an activity that has an expected environmental impact. Norms established by regulatory requirements result in compliance monitoring, which resembles an audit or performance review. An example of a regulatory norm is: No discharges beyond a defined value. Norms also arise from the predictions of an environmental impact statement. These predictions result in effects monitoring, a field with vigorous development in the last two decades, since the landmark book by Green (1979). The principles of experimental design are often advocated for effects monitoring, even though it is usually impossible to assign treatments randomly to replicate units, an important condition of classical experimental design (Fisher, 1954). The development of the topic here takes the approach that Fisher's principles apply to experiments, not surveys. Thus, the rigor of a monitoring program follows from the application of the principles of good survey design to the predictions from an environmental impact statement. Effects monitoring becomes a matter of repeating a formal survey often enough to detect predicted changes.

The scope of a monitoring program is defined in the same way as the scope of a survey. The *spatial scope of a monitoring program* is defined as the area of the frame relative to the diameter, area, or volume of the unit that is completely measured. The frame is defined by the predicted impact. If the predicted impact is distributed over an entire watershed, the program will require an extensive frame, one that covers the watershed. For many activities the impact occurs at a point and so is expected to attenuate with distance. The frame is relatively limited, being just larger than the spatial scale of the impact. The *temporal scope of a monitoring program* is defined as the duration of the predicted impact relative to the duration of measurement of a single unit. The temporal scale of a monitoring program will thus be relatively large compared to a single survey. Well-designed monitoring programs have a defined stopping rule, set either by the life of the activity or by some consideration of cost relative to the benefit from more information.

Because effects monitoring is based on predictions, the spatial structure of the program is usually more complex than a survey. An example is the effects monitoring program developed for the Hibernia oil production platform, located in 100 m of water near the edge of the continental shelf in the northwest Atlantic, about 300 km east of St. John's, Newfoundland. Over a 20- to 30-year lifetime, the platform is expected to discharge drilling fluid and tailings from at least 60 wells, each on the order of 4 km deep. Discharge is predicted to alter the sediment chemistry and abundance of benthic organisms. Changes are expected to be greatest near the platform, attenuate with distance, and fall to background levels at 2–5 km from the platform. To monitor relative to these predictions, the monitoring program defined stations at roughly logarithmic intervals (100 m, 200 m, 500 m, 1 km, 2 km, 5 km, 10 km) along 4 radii, one running into the

prevailing current, one running downstream, and 2 perpendicular to the current to form a cross-shaped layout. A survey was carried out before the platform was set on the seafloor. Surveys were carried out 3 years and 5 years later. Scope calculations (Box 12.5) are broken out by survey because change at this time scale is the focus of the study.

Box 12.5 Scope Calculations for a Program to Monitor the Impact of Discharges and Tailings from the Hibernia Drilling Platform on the Benthos of the Outer Grand Bank, Newfoundland

Station: $A = \pi(5\,\text{m})^2 = 78.5\,\text{m}^2$
Radius: $A = (10\,\text{m})(10\,\text{km}) = 100\,\text{m}^2$
Survey: $A = \pi(10\,\text{km})^2 = 314\,\text{km}^2$

Level	Replication	Units	A_o	A	Scope	T_o	T	Scope
Pre-impact								
Core		1	$(25\,\text{cm})^2$			1 sec		
Station	3 cores	3	$0.19\,\text{m}^2$	$78.5\,\text{m}^2$	419	3 sec	10 hr	$1.2 \cdot 10^4$
Radius	6 stations	18	$1.125\,\text{m}^2$	$100\,\text{m}^2$	88.9	18 sec	2.5 day	$1.2 \cdot 10^4$
Survey	4 radii	72	$4.5\,\text{m}^2$	$314\,\text{km}^2$	$7 \cdot 10^7$	72 sec	10 day	$1.2 \cdot 10^4$
Post-impact 1		144	$9\,\text{m}^2$	$314\,\text{km}^2$	$3.5 \cdot 10^7$	144 sec	3 year	$6.6 \cdot 10^5$
Post-impact 2		216	$13.5\,\text{m}^2$	$314\,\text{km}^2$	$2.3 \cdot 10^7$	216 sec	5 year	$7.3 \cdot 10^5$

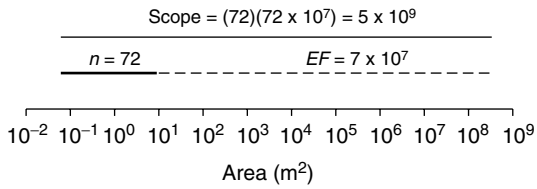


Figure 12.2 displays the scope calculations in graphical form. As before, the effort (n) at each level in the study is plotted as a series of circles, connected by a line. The study design is shown as another line connecting a series of crosses, one at each level. The scale-up (EF) is shown for the pre-impact survey alone. In Figure 12.2a the scale-up or expansion factor EF has been resolved into spatial and temporal components. Figure 12.2b shows the scale-up (EF) for the pre-impact survey and for the entire monitoring program to detect change. Either expansion factor EF can be partitioned into components according to the levels in the design.

The inferential basis for scale-up in this survey differs from the frame-based survey described for scallops. The frame-based survey estimates a parameter (usually, a mean) from samples taken randomly (hence, with known probability) within a fixed list of units. Such a design could have been used in this study by fixing an area around the

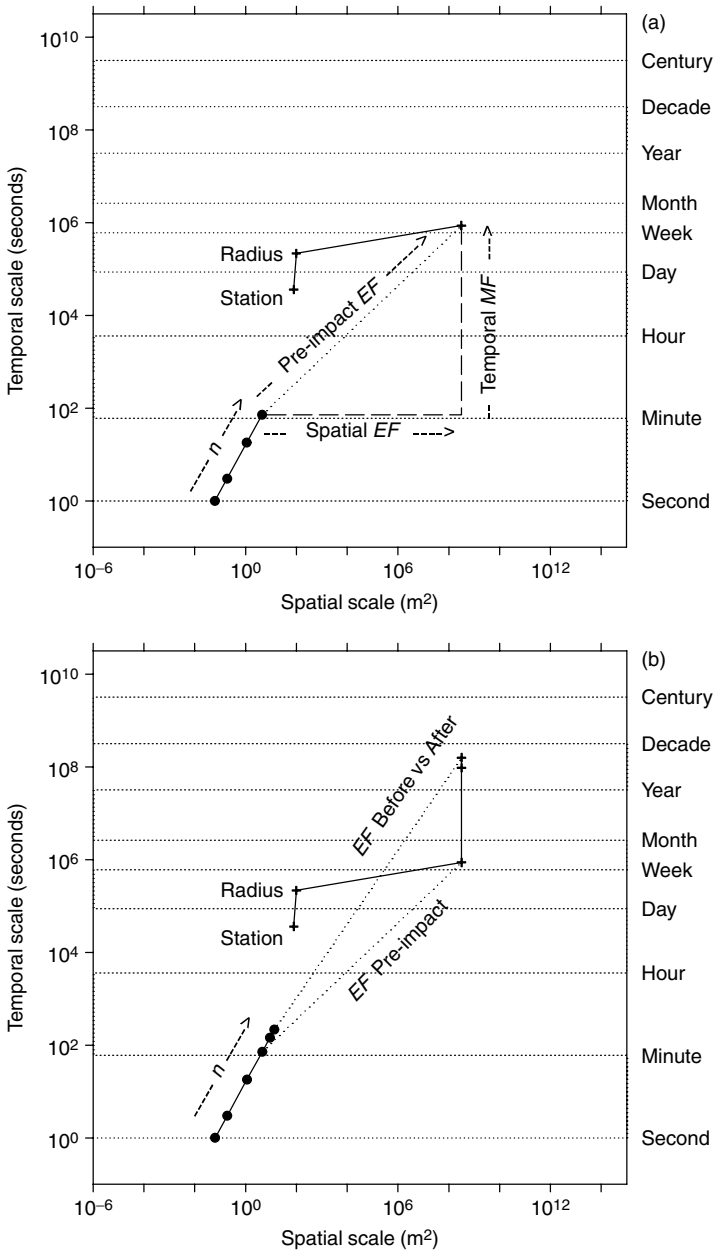


FIGURE 12.2 Scope Diagram for a Monitoring Program. A radial design was chosen to detect predicted effects of point release of contaminants from oil platform at the Hibernia site in the northwest Atlantic. Sampling effort at each level in the study is plotted as a series of circles, connected by a line; (a) The scale-up factor *EF* for the pre-impact study, resolved into spatial and temporal components; (b) The study design is shown as a line connecting a series of crosses, one at each level. Dotted line shows scale-up factor *EF* for the pre-impact survey and for the entire monitoring program to detect change.

drilling platform, then sampling randomly from within the frame of all sites before and after the impact. Though this is a valid design, it does not make effective use of our knowledge of the expected effect of drilling discharges, which is a decrease in impact with increasing distance from the discharge site. To monitor relative to the predicted

impact, a gradient design was adopted (Ellis and Schneider, 1997). The design estimates the gradient in some quantity Q with increase in distance from the point of discharge. Q might be, for example, the concentration of a contaminant such as barium, a component of drilling fluids that is released into the environment. A likely model, based on diffusion due to eddy mixing, is an exponential function of distance X.

$$Q = Q_o e^{(\beta_x X)} \tag{12.1a}$$

Here, Q_o is the value of Q immediately adjacent to the platform. The parameter β_x is the gradient in Q on an exponential scale. The same model in different format is:

$$Q = e^{(\beta_o + \beta_x X)} \tag{12.1b}$$

Here, $\ln(Q_o) = \beta_o$.

The number of possible radii along which to estimate β_x is substantial. At the scale of a station (radius = 5 m, diameter = 10 m) there are $n = (2\pi\ 10\,\text{km})/(10\,\text{m}) = 6284$ endpoints (nonoverlapping stations) along the periphery of a circular area A with a radius of 10 km. In the absence of knowledge of patterns of dispersion of contaminants around a release point, we would use a random sampling design to select radii. However, the Hibernia site sits within the influence of the southward-flowing Labrador current; hence we expect the gradient to be attenuated downstream (to the south), stronger upstream (to the north), and somewhere in between at right angles to the current (east and west). Consequently, the gradient design was oriented according to the Labrador Current with one transect running upstream, one downstream, and two at right angles to the prevailing current.

Because the stations were fixed in two spatial dimensions (fixed positions along each transect at logarithmic intervals, fixed direction relative to the prevailing current), inference from the scale of the station to that of the entire area is no longer that of statistical inference from unit (grab) to frame of 5×10^9 possible grab sites in a 10 km radius area. More realistically, the statistical inference from unit to frame is by a factor of 10.5×10^3 , computed as:

$$\frac{(24\text{ stations})(78.54\text{m}^2/\text{station})\left(\frac{100\text{cm}}{m}\right)^2}{(72\text{ samples})(25\text{cm}^2/\text{sample})} = 10472$$

The burden placed on statistical inference has been reduced from a factor of 7×10^7 (expansion factor in Box 12.5) to 10.5×10^3. The reduction is substantial but by no means free of cost. It is based on knowledge of diffusion in the ocean in general and at the Hibernia site in particular, information that is expensive to obtain.

The program at Hibernia monitors relative to both regulatory norms and predictions from knowledge of marine systems. The predicted effect, based on knowledge summarized in the impact statement, was for a post-impact gradient in contaminants and in population density to develop out to distances of less than 5 km. The gradient is estimated via the research program diagrammed in Figure 12.2. The gradient design, oriented relative to the Labrador Current, reduced the burden on statistical inference by introducing different forms of inference at different levels within the overall scope of the

monitoring program. Statistical inference (samples from a defined frame) was confined to within stations of approximately 5 m in radius. At larger scales, stations were fixed along transects, which were in turn fixed relative to prevailing currents. At this scale, statistical inference is to an infinite population of all possible measurements that could be made, given the fixed station protocol. Statistical inference no longer addresses spatial uncertainty. The scale-up from station to transect rests on the assumption that logarithmically spaced stations will adequately represent the gradient in effects from a point source. Scale-up from transects to the entire area around the point source rests on the assumption that four transects oriented either parallel or perpendicular to the prevailing current will represent the extremes in gradient strength, from high (upstream) to intermediate (perpendicular to the flow) to least gradient (downstream).

ANOTHER LOOK AT SECTION 12.4.4

State two predictions of the impact of building a road into a 25 km^2 area devoid of roads. State quantities you would measure, draw a spatial layout to detect effects, and construct a table (as in Box 12.5) showing levels, replication, and units.

12.4.5 Hierarchically Structured Surveys

Another purpose of a survey is quantifying variance (and hence contrasting estimates of the mean among units) at multiple spatial scales. Such surveys are particularly useful at the start of a research program because they allow efficient allocation of effort in ways that yield the most information per sample. An example is a set of measurements to characterize spatial differences in the density and rate of loss of invertebrate prey during a period of intense predation by migratory shorebirds. Each year migratory shorebirds congregate in large numbers at a few midlatitude coastal sites along their southward migratory route. At these locations they put on substantial fat stores, which are required to complete transoceanic migration to wintering grounds in the tropics or southern hemisphere. Because of human pressures on the coastal zone at the midlatitude sites, a key conservation question was whether the prey base remains stable during periods of high use at coastal sites with major concentrations of birds.

Systematic studies of invertebrate prey were begun in 1975 on White Flat, located in a coastal lagoon at Plymouth, Massachusetts (42°N), between Boston and New York City. The lagoon has an area of 40.7 km^2, of which about two-thirds is exposed at low tide. During July and August it is used by thousands of migratory birds, which were known to stay for several weeks to feed and fatten before continuing their southward migration. The quantities of interest were the densities $[N_i]$ of all infaunal invertebrates, the densities $[N_j]$ of prey species at sizes taken by shorebirds, and rates of loss of infaunal prey $[\dot{N}]$ during the period of high predation by migratory birds. The spatial variance in these quantities was estimated by a hierarchically structured survey on White Flat, an intertidal flat with high habitat heterogeneity. Five plots, each 1 hectare in area, were chosen to represent the range of habitat and shorebird species composition on the flat. Each plot was marked with stakes at each corner, feeding birds were counted at

weekly intervals in each plot, and the first round of invertebrate sampling was completed July 22–24 at randomly chosen sites in each plot.

A two-stage process was used to locate sites randomly within each plot. A 10 m by 10 m area within each plot was located via random numbers from 0 to 9. Within this subplot two samples were taken, each by a random number from 0 to 9. This process was repeated three times, resulting in four subplots and eight sample sites per plot. Once a site was located (to an accuracy of 1 m), a core was taken within a 1 m^2 area by driving a plastic coring tube into the sediment. This was done while the researcher was looking up, to eliminate bias caused by looking at the sediment surface when placing cores.

This study had six levels. The first (lowest) level is that of the coring device, with a diameter of 10 cm, hence an area of 0.007854 m^2. The diameter was chosen to give counts on the order of 0–100 organisms per core in the size range of invertebrates (1 mm to 3 mm diameter) taken by birds. The second level is that of the 1 m^2 area in which each core was placed. The third level was that of the $(10\,m)^2$ subplot. This level was of interest because the distribution of birds and sediment features in the plot suggested there was important spatial variation at this scale. The fourth level was that of the 1 hectare plot. This unit was chosen to be large enough to make quick counts of all birds foraging in a known area. The fifth level was that of the habitat. Three plots were set in silty sand on the west side of the flat, representing about half the flat. These three plots ranged from light silt (representing about one-quarter of the flat), moderate silt (about one-eighth of the flat), and silt with mud (about one-eighth of the flat). Two plots (representing about half the flat) were set in sand on the east side or north end of the flat, where ebb-tide currents winnow silt from the sediment. The sixth level was a single flat of 1.1 km^2, surrounded on all sides by water at low tide. Box 12.6 shows the replication at each level, the overall scope of the study, and the scope at each level.

Box 12.6 Scope Calculations for a Stratified Survey of Infaunal Invertebrate Density in Areas Used by a Large Number of Migratory Shorebirds on White Flat, Plymouth Harbor, 1975

Core area: $A_o = \pi(5\,\text{cm})^2 = 78.5\,\text{cm}^2$

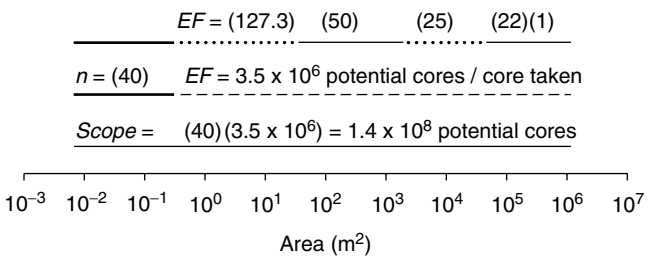
Level	Replication	Units	A_o	A	Scope	
Core			$\pi(5\,\text{cm})^2$			
Site	1 core	1	0.007854 m^2	1 m^2	127.3	core/core taken
Subplot	2 sites	2	0.01571 m^2	100 m^2	6366	core/core taken
Plot	4 subplots	8	0.06283 m^2	10000 m^2	159200	core/core taken
Habitat	(2 + 3)/2 plots	20	0.1257 m^2	550000 m^2	3501000	core/core taken
Flat	2 habitats	40	0.3142 m^2	1100000 m^2	3501000	core/core taken

Overall spatial scope	$SpSc$	=	$1.1\,km^2/0.007854\,m^2$ =	$1.401 \cdot 10^8$	potential cores
Units	n	=	8 cores/plot · 5 plots =	40	cores taken
Expansion factor	EF	=	$1.1 \cdot 10^8/40$ =	$3.501 \cdot 10^6$	potential core/ core taken
Sampling fraction	SF	=	EF^{-1} =	$2.856 \cdot 10^{-7}$	core taken/ potential core

Partitioned spatial scope

$$SpSc = 40\left(\frac{127.3}{\mathrm{m}^2}\right)\left(\frac{10^2\ \mathrm{m}^2}{1\ \mathrm{m}^2}\cdot\frac{1}{2}\right)\left(\frac{100^2\ \mathrm{m}^2}{10^2\ \mathrm{m}^2}\cdot\frac{2}{8}\right)\left(\frac{55\mathrm{ha}}{1\mathrm{ha}}\cdot\frac{8}{20}\right)\left(\frac{110\mathrm{ha}}{55\mathrm{ha}}\cdot\frac{20}{40}\right)$$

$$SpSc = 40(127.3)(50)(25)(22)(1)$$



Box 12.6 shows the partitioning of the overall scope into a measured (n) and an inferred (EF) component. The box further shows the partitioning of the inferred component into scale-up factors for each level of the stratified survey. In this example the overall expansion factor was partitioned in five components (core to site, site to subplot, subplot to plot, plot to habitat, habitat to flat). Each component is the ratio of two areas, adjusted for the replication at that level. For example, the site-to-subplot component is the ratio of the site-to-subplot area ($100\,\mathrm{m}^2/1\,\mathrm{m}^2$) adjusted by the replication at this level (two sites per subplot). The expansion factor at this level is 50, the expansion factor at the next level is 25, and so on (Box 12.6). Each expansion component can be calculated from the ratio of scopes at two levels. The expansion factor from site to subplot is thus 6366/127.3 = 50. The expansion factor at the next level is 159200/6366 = 25.

ANOTHER LOOK AT SECTION 12.4.5

Pick a species for which you have some sense of spatial variation (e.g., pigeons in a city). Define survey levels that you think capture major components of spatial variance. Make rough scope calculations for the survey, as in Box 12.6, using reasonable guesses of sample number and area.

12.4.6 Magnitude and Basis of Scale-up in Surveys

The magnitude of scale-up can be quantified for an entire survey, then partitioned among levels in a survey. The magnitude of scale-up at each level depends on the number of units

that accumulate by replication at that level and on the ratio of unmeasured to measured units at that level. The basis for scale-up varies across studies and can vary within a study. Scale-up can be by judgment that sites are representative (eccentric bog study), by informal inference based on haphazard sampling (small-scale sample placement in bog, scallop, and Hibernia studies), by formal statistical inference in a finite spatial frame (in the scallop survey), by a descriptive spatial model (habitat area assignable to plots at Plymouth), or by spatial model with theoretical content (Equation 12.1 for Hibernia monitoring).

The basis for spatial scaling from measured to unmeasured units was the same (informal inference based on judgment) at both levels in the bog reconnaissance (Box 12.1). The basis for spatial scaling in the random and stratified random survey (scallop survey, Box 12.3) was formal statistical inference in a finite frame because all units in the frame had known probabilities of being sampled. In the Hibernia monitoring program (Box 12.5), the basis for spatial scale-up depended on level. At the level of core to station, the basis was informal inference based on haphazard sampling, because the mechanics of taking box cores over the side of a ship put the chance of a particular core site being hit beyond the control of the investigator. Thus it is reasonable to assume that core samples are independent of variation in density and hence produce a representative and unbiased estimate of the density at each station. The basis for scaling from station to radius was a model based on attenuation of contaminants with distance from a point source (Equation 12.1). The basis for scaling from radius to circular area around the drilling platform was the same model. It is expected that the key parameter in the model (rate of attenuation with distance) will vary among the 4 radii. The parameter is expected to be lower (weaker gradient) in the direction of net current (southward at Hibernia) than the upstream direction.

The scale-up at each level in the Plymouth invertebrate density study, quantified in Box 12.6, rests on different forms of inference. The scale-up from core to site (by a factor of 127.3) is based on haphazardly sampling. Because cores were taken blindly, the probability of a site being sampled was assumed to be independent of variation in density within a site and hence assumed to give representative and unbiased estimates of the mean within a site. The scale-up from site to subplot (factor of 50) and from subplot to plot (factor of 25) was by probabilistic sampling, where random sampling guarantees representative and unbiased estimates. The scale-up from plot to habitat (factor of 22) was by systematic selection of plots within habitats on the flat. This was assumed to represent the range of variability within the two habitats. There was no scale-up from habitat to flat (factor of one) because the entire flat was divided into either sand or silty sand habitat, which were both measured. Partitioning of the overall scale-up revealed a surprising fact, which is that simple or stratified random surveys rest on haphazard sampling at small scales, which can account for substantial fractions of the overall scale-up addressed by these surveys.

The basis for temporal scaling was informal inference at all levels in both the bog reconnaissance and the scallop survey. The basis for the judgment that a few points in time represent the timeframe of the study was the knowledge that the organisms sampled are sedentary, and hence densities do not change as a result of movement in or out of the measurement unit (relevé, dredge haul, box core). Temporal scale-up was explicit in the monitoring (Hibernia) study. The time between surveys was chosen to detect change predicted to result from release of contaminants. The time to complete each survey was assumed to be small relative to the time for changes in benthic faunal density to occur. The temporal

scope of the hierarchically structured survey was adjusted to the expected rates of change in the density of invertebrates due to shorebird predation. The first round of sampling was completed as quickly as possible, to minimize change due to predation during the sampling. The second round of sampling (not reported in Box 12.6) was completed as soon as shorebird numbers begin declining in late August, to maximize the length of time of exposure to shorebird predation while minimizing the period of time after the departure of shorebirds.

12.5 The Scope of Experiments

An *experiment* is a test or a trial to discover something unknown. The principal advantage of an experiment is that it compels an investigator to consider and control confounding factors before undertaking structured observations. Explanations easily form around confounding factors, which are more numerous and sometimes more dramatic than causal factors. In a descriptive study, confounding factors are considered after an explanation is formed. In an experiment, confounding factors are considered before taking observations. Within this broad definition of experiment there exists considerable diversity in approach. An explanation is tested by altering the factor thought to be causal, holding other factors constant. An explanation may be challenged with a set of observations chosen to separate two or more competing hypotheses (Platt, 1964). An explanation may be tested in the presence of a confounding factor that is measured rather than held constant. These measurements are incorporated into the analysis to discover whether a proposed explanation remains tenable after removing the effects of a confounding factor. This approach, statistical control, is a key feature of experimental design (Cochran and Cox, 1980).

Ecological experiments are undertaken to understand pattern and process in a larger context, usually that of an ecosystem. Scope diagrams reveal the enormous scale-up from experiment to the system of interest. These diagrams are useful in comparing and evaluating experimental results. They are also useful in evaluating the generality of an experiment. The *spatial scope of an experiment* is defined as the maximum distance between measurements divided by the resolution (support) of a single measurement. Areas or volumes can be used instead of distances in making scope calculations. The *temporal scope of an experiment* is defined as the time between the start and end of the experiment divided by the duration of a single measurement.

The following sections demonstrate the construction of scope diagrams for a series of increasingly complex experiments. Section 5.1 demonstrates the procedure for two experiments to estimate a rate parameter, one in the lab and one in the field. Section 5.2 demonstrates the procedure for classical Fisherian experiments in agriculture. Section 5.3 shows scope calculations for a Fisherian experiment in ecology. Section 5.4 considers the limits of Fisherian experiments in ecology. It then describes the scope of inference at multiple levels in experiments that do not conform to Fisherian designs with just two levels. Section 5.5 describes embedded experiments, which make explicit the magnitude and basis for scale-up of the experimental components of a study.

12.5.1 Experimental Estimate of a Parameter

Experiments are often conducted to estimate a parameter under controlled conditions. For example, an experimental analysis of plant growth in relation to light clearly has the aim of estimating a parameter: the growth rate as a function of light. A statistical test of

whether growth depends on light is somewhat beside the point since there is no question that reduced light reduces growth. A measure of uncertainty on the estimate of the parameter is of more interest.

One use of parameter estimates from controlled conditions in the laboratory is to calibrate field studies. An example (Ogilvie, 2000) is an experiment to determine the mortality rate due to handling and marking prior to undertaking a mark-recapture experiment in the field. Juvenile cod *Gadus morhua* were marked with flourescent dye, released into a tank, then recaptured and counted 24 hours later. The design held mortality to zero from all sources except marking.

The scope diagram is constructed from a table that begins with the resolution (or support) of the sampling unit. For the tank experiment (Box 12.7), the lateral extent is $A_o = (3\,\text{m})^2$. The measured duration (T_o) is one day for the trial. At the next level (next row in the table, Box 12.7), the replication is listed. The experiment was repeated six times in a three-month period in the same tank. The area sampled remains the same $A_o = (3\,\text{m})^2$. The area within which samples are taken (A) is then listed. Sampling did not represent some larger area than the tank, so the overall area A at the level of the study is the same as the area sampled (A_o) at this level. The effort at the level of the study is calculated as the product of the replication and the duration of the unit ($T_o = 6$ days). This effort occurred within an overall period of $T = 30$ days. These calculations (Box 12.7) are relatively simple compared to more complex designs that will be encountered later.

Once the components for scope calculations are obtained, they can be displayed in a diagram. Figure 12.3 shows the measured area and duration (A_o, T_o) as circles connected by a solid line. The total extent and duration (A, T) is shown as a cross. Temporal scale-up to the full three-month duration of the experiment is shown as a dotted line.

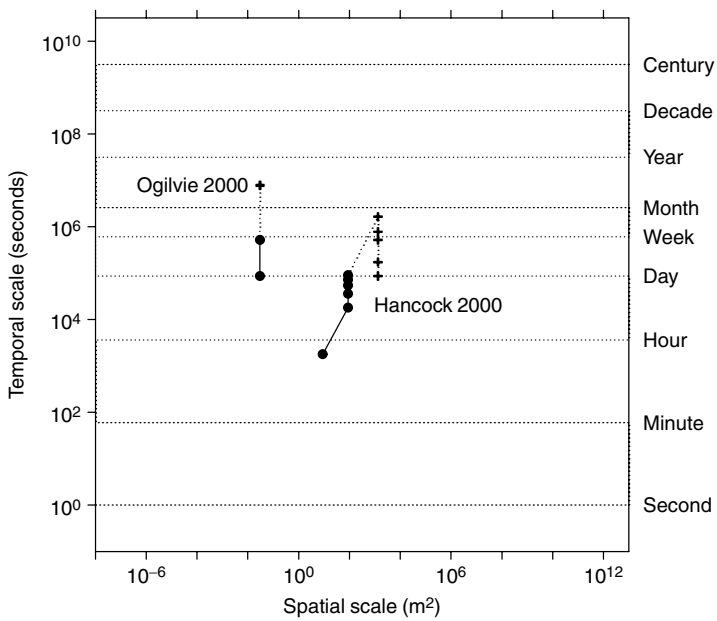


FIGURE 12.3 Scope Diagram for Field Experiment and Related Laboratory Experiment. Mortality due to handling and marking of fish was estimated in the lab by Ogilvie (2000). Box 12.7 shows scope calculations. Dispersal rate of juvenile cod in the field was estimated by Hancock (2000) using mark-recapture. Box 12.8 shows scope calculations. Solid circles connected by lines show effort at different levels of the experiment. Crosses show target of inference. Dotted line shows scale factor *EF*.

Box 12.7 Scope Calculations for a Laboratory Experiment (Ogilvie, 2000) to Estimate Mortality Due to Handling and Marking of Fish

Level	Replication	Units	A_o	A	Scope	T_o	T	Scope
Trial		1	$3\,m^2$			1 day		
Study	6 trials	6	$3\,m^2$	$3\,m^2$	1	6 days	3 months	15

Field experiments can also have the aim of estimating a parameter. An example is a mark-recapture experiment (Hancock, 2000) to estimate the rate of juvenile cod dispersal. The handling protocol in this experiment was the same as for the complementary laboratory experiment by Ogilvie (2000). Marked fish were released at a point on a beach. Seine hauls were then taken at 10 locations spaced along more than 200 meters of beach. Each haul collects nearly all fish in a 16 m swath running 55 m seaward from the beach. Hauls were made at all 10 locations at roughly logarithmic intervals in time: 1, 2, 6, 9, and 16 days after release of fish. The degree of dispersion was plotted against the cumulative time, so cumulative time was used in the temporal scope calculations (Box 12.8).

Box 12.8 Scope Calculation for a Mark-Recapture Experiment (Hancock, 2000) to Estimate Dispersal Rate of Juvenile Cod

Site = (240 m)(55 m)
Haul = (16 m)(55 m)

Level	Replication	Units	A_o	A	T_o	T	Scope
Haul		1	$880\,m^2$		30 min		
Visit	10 sites	10	$8800\,m^2$	$13200\,m^2$	2.5 hr	1 day	9.6
Visit	10 sites	20	$8800\,m^2$	$13200\,m^2$	5.0 hr	2 day	9.6
Visit	10 sites	30	$8800\,m^2$	$13200\,m^2$	7.5 hr	6 day	19.2
Visit	10 sites	40	$8800\,m^2$	$13200\,m^2$	10.0 hr	9 day	21.6
Visit	10 sites	50	$8800\,m^2$	$13200\,m^2$	12.5 hr	16 day	30.7
Experiment	5 visits	50	$8800\,m^2$	$13200\,m^2$	12.5 hr	16 day	30.7

Spatial scope	$SpSc$	$= 13200\,m^2/\,880\,m^2$	$= 15$ possible sites
Units	n	$=$	$= 10$
Expansion factor	EF	$= 16.5/10$	$= 1.5$
Sampling fraction	SF	$= EF^{-1}$	$= 0.67$

The scopes of the field and laboratory experiment were plotted together for comparison (Figure 12.3). The two experiments have similar temporal scales. The spatial scale-up from lab to field assumes the same mortality rate for fish in a $3\,m^2$ tank as fish ranging over larger areas. Scale-up beyond the scope shown in the figure relies on a knowledge of the biology of cod. One would expect the same dispersal rate to apply to other coastal areas and other years but not to other seasons.

ANOTHER LOOK AT SECTION 12.5.1

Compare the scope diagram for the two experiments (Figure 12.3) to the diagram for the descriptive study of eccentric bogs (Figure 12.1). Do you think the differences are typical of experiments versus descriptive studies? Why or why not?

12.5.2 Fisherian Experiments

In ecology, some investigators constrain the meaning of *experiment* to the prescription developed by R. A. Fisher (1954) for agricultural experiments. The prescription requires the manipulation of an explanatory variable at fixed treatment levels with repeated trials, including controls (trials with no manipulation), and random assignment of treatments and controls to units. This prescription defines a *Fisherian experiment*, which has become increasingly prominent in ecological research (Hairston, 1989; Resetarits and Bernardo, 1998). The Fisherian program of randomization, replication, and local control was developed to address problems of spatial heterogeneity and confounding within agroecosystems. Heterogeneity among units in agroecosystems cannot be eliminated entirely, but it can be reduced substantially before the experiment. The classic Fisherian solution to the problem of spatial variation is adequate replication of units to estimate variance due to random (uncontrolled) factors. Experimental units in one area may differ from those in another. If treatments are assigned to one location and controls to another location, then even after chance effects are excluded by the logic of inferential statistics, differences due to treatment cannot be separated from differences due to location, even if the locations are adjacent and resemble each other closely. The classic Fisherian solution to this problem is randomized assignment of treatments to a defined set of spatial units so that treatment is independent of spatial heterogeneity among units.

In agricultural work it has long been recognized that variance among experimental units will increase as they become more widely separated (e.g., Smith, 1938). As the treatment levels or number of factors increase, the spacing between treatments and control necessarily increases in a completely random design. The result is an inefficient design due to naturally large differences between widely separated treated and control units. The classic Fisherian solution to the problem of increase in spatial heterogeneity with increase in separation is local control. In experiments in which several factors are manipulated, achieving local control along with randomization and replication requires sophisticated designs to analyze the results of layouts that reduce spatial separation of treatment and control units (Cochran and Cox, 1957). Local control and sophisticated spatial designs are solutions to the problem of the increase in heterogeneity with increasing spatial scale or separation between experimental units.

Fisherian experiments focus on defined spatial units, to which treatment and control are assigned randomly. Measurement is usually made across the entire unit, resulting in one value per unit. Thus the *spatial scope of a Fisherian experiment* is the maximum distance between two experimental units in a study, divided by the resolution (area or support) of a single unit. The *temporal scope of a Fisherian experiment* is the time between the start and end of the experiment divided by the duration (support) of a single measurement.

An example of a Fisherian design is the *Latin square*. This design assigns treatments to plots in random order within both rows and columns of a regular (usually square) array. It guarantees the principal features of sound design in agriculture: randomization, replication, and local control. In particular, the design guarantees the proximity of control to treatment at the scale of the row and column rather than at the larger scale of the entire array. Snedecor and Cochran (1980, p. 289) present the results of a Latin square design to compare the effects of four types of soil fumigants on wireworms. A fumigant or control (no treatment) was assigned to plots measuring 22.86 cm on a side and 12.87 cm deep. Local control was achieved by random dispersal of untreated plots within each row and column of a 5 by 5 array. Random assignment within both rows and columns allows larger-scale (row and column) effects to be removed during statistical analysis. Adding row and column factors to the analysis reduces the error term and increases the ability to detect an effect. Random assignment of each plot to either a treatment or a control eliminates confounding of treatment effects with unknown sources of variation among plots.

Box 12.9 Scope Calculation for a Latin Square Design Reported by Snedecor and Cochran (1980)

Level	Replication	Units	A_o	A	Scope	T_o	T	Scope
Plot		1	$(22.86\,\text{cm})^2$			1 month?		
Array	25	25	$1.31\,\text{m}^2$	$4.23\,\text{m}^2$	3.2	1 month	1 month	1
Repeat	2	50	$2.61\,\text{m}^2$	$8.47\,\text{m}^2$	3.2	2 month	1 year	6
Farm			$2.61\,\text{m}^2$	$1\,\text{km}^2$	$3.8 \cdot 10^5$			

Spatial scope $SpSc = 1\,\text{km}^2/(22.86\,\text{cm})^2 = 1.9 \cdot 10^7$ possible samples

Units $n = 25/\text{array} \cdot 2\ \text{arrays} = 50$

Expansion factor $EF = 1.9 \cdot 10^7/50 = 3.8 \cdot 10^5$

Sampling fraction $SF = EF^{-1} = 2.6 \cdot 10^{-6}$

Spatial scale-up: $n \cdot EF = n \cdot EF_{array} \cdot EF_{farm}$

$$= 50 \cdot \frac{162}{50} \cdot \frac{1.9 \cdot 10^7}{162}$$

$$= (50)\,(3.2)\,(1.2 \cdot 10^5)$$

The calculation in Box 12.9 begins with the extent (A_o) of the sampling unit. For the wireworm example, the lateral extent is $(22.86\,\text{cm})^2$. The measured duration is time from day of treatment to day of collection (1 month?). At the next level (next row in the table), the replication is listed. From this the measured extent (A_o) is calculated as $5^2(22.86\,\text{cm})^2$. The area sampled (A) is then listed. If the samples are dispersed rather than contiguous, A exceeds A_o. Spatial separation was not reported, but it is reasonable to assume that treated plots were separated from each other by at least 1 plot width to avoid lateral contamination of one plot by another. Under this assumption the extent of the array is $(5 + 4)^2(22.86\,\text{cm})^2$.

The experiment was repeated the following year (Snedecor and Cochran, 1980, p. 289), which adds a new level to the experiment. The replication at this level was 25 additional plots. It was not clear whether the experiment was carried out on the same plots. The calculations in Box 12.9 assume that the array in the second year had the same spacing as in the first year. This brings the replication to 50, increasing the measured extent (A_o) to $2 \cdot 5^2(22.86\,\text{cm})^2$. These 50 plots occur in an area $A = 2\ (5 + 4)^2(22.86\,\text{cm})^2$, assuming separate arrays with one plot width between each experimental plot. The area assumed available at the station for this type of experiment was taken as $1\,\text{km}^2$.

The overall scope can then be partitioned into two components, one due to accumulation of measured units (n) and one due to inference (EF). In this example, the latter component can be further partitioned into the scale-up within the array and the scale-up from the array to the farm. Box 12.9 shows the partitioning of the overall spatial scope into three scale-up factors.

This partitioning allows us to evaluate the basis for scale-up at each level, in relation to the magnitude of scale-up. At the first level scale-up is by a factor of 25. The basis for scale-up is accumulation of units with randomly assigned treatments. These are representative as long as the units are accumulated independently of the measured values in other units. At the next level, scale-up is by a factor of two, the result of repeating the experiment. The basis for scale-up is again accumulation of units. The next scale-up is inferred by a factor of $162/50 = 3.2$. The basis for scale-up at this level can be statistical inference, depending on how we define the population to which we are inferring. At the final level, scale-up is by a larger factor, 10^5 (Box 12.9). The basis for scale-up here can also be statistical inference, again depending on how we define the population.

In a classic Fisherian experiment the inference is usually from the sample to an infinite list: all possible difference in means that could have arisen if we repeat the experiment again and again. The definition of the population thus turns on the experimental design and measurement protocol, which includes the preparation of the experimental units. In particular we must decide whether the 50 measured units represent the 162 possible nonoverlapping units in the two arrays. If both arrays ($A = 8.47\,\text{m}^2$) were prepared in the same way before the experiment, we would likely judge that the population includes the means obtained at all 162 possible sample locations. By this definition of the population, statistical inference justifies the spatial scale-up (by a factor of 3.2). If all fields were prepared in the same way, we might further judge that the population includes the experimental outcome from any array prepared the same way on an experimental farm. By this definition statistical inference supports the scale-up to the farm (by a factor on the order of 10^5 or more). However, it is also possible to judge that the preparation of a field cannot eliminate variation at this scale, that the results of the experiment will depend on field within a farm, and hence that statistical inference does not

include the scale-up to the level of the farm. If we are not prepared to state that the list of all possible outcomes will include those from plots measured outside the two arrays, we cannot use statistical inference. We could not use the Fisherian design to eliminate confounding effects at scales larger than the array of plots.

The Fisherian design of this experiment allows us to eliminate chance as an explanation of the results; it allows us to eliminate confounding effects at the scale of the array. It does not necessarily allow us to eliminate confounding effects at scales larger than the array. It is possible that the experimental result will depend on some peculiar feature of the field in which it was conducted. Judgment based on knowledge of practice on the farm is needed to infer beyond the array and so justify the expectation that applying the same soil fumigants at the same concentration with the same procedure would have the same effect (within error) at other sites on the same farm and at other farms with the similar agricultural practices. As in any agricultural experiment, knowledge of the system and of the history of farming practice at the site is the basis for judging that statistical inference allows scale-up to areas beyond the perimeter of the experimental array.

ANOTHER LOOK AT SECTION 12.5.2

Try to design a rigorous Fisherian experiment (randomization, replication, local control) to address the problem of monitoring the predicted impacts of the Hibernia project described previously. List the problems you encounter with regard to the three principles of Fisherian experiments.

12.5.3 Fisherian Experiments in Ecology

The ecological literature contains many examples of well executed Fisherian experiments. An example is an experiment to discover the effects of competitor density on the water balance of shrubs in the Mojave Desert of California (Fonteyn and Mahall, 1981). Xylem pressure potential was measured in the two species *Larrea tridentata* and *Ambrosia dumosa*. Predawn xylem pressure is thought to reflect soil moisture availability and thus the competitive effect of neighboring plants. Measurements were taken one hour before dawn in a plant centered within a $100\,m^2$ plot having known density of competitors of the same species, of the other species, or both. The experiment was particularly well executed in regard to minimization of variance and control of confounding effects. Twenty plots were chosen, of which half were centered on a *Larrea* plant and half centered on an *Ambrosia* plant. To reduce variance among units, the investigators chose plots that contained at least 3 *Larrea* and 11 *Ambrosia* plants. To eliminate confounding with locational effects, they assigned treatments and controls randomly to units. If logistics allowed, local control could have been increased by assigning treatment and control randomly within blocks of eight plots. This would place controls in greater proximity to treatments. The replication of units allowed the investigators to eliminate chance as an explanation for the observed lower xylem pressure potential in the more crowded plots. The random assignment of treatments to units allows the investigators to eliminate locational effects as an explanation of differences between controls and the several types of treatment.

Box 12.10 shows the scope calculations for this experiment. The computational flow reflects the fact that measurements were taken at the same locations over a 14-week period. Thus the measured area A_o on the first visit remains the same for the entire study. Measurements were taken periodically at one hour before dawn, before the onset of evapotranspirative stress. The measurements taken during a single visit extend over roughly a minute during the hour before dawn, but there is only one such hour in the day and seven such hours in the week. The temporal scale-up from a visit to the entire study is by a factor of 12.25, the ratio of predawn hours in 14 weeks (98) to the number of hours (8) in which measurements were taken.

The results presumably apply to the Mojave Desert ecosystem. Otherwise they would be of little interest. A study based on $100\,m^2$ plots in the Mojave Desert will have an overall scope of $6 \cdot 10^8$ (Box 12.10). Within this scope the scale-up from plot to study area is by a factor of 6.25, compared to a scale-up of $2.4 \cdot 10^6$ from study area to ecosystem. This large scale-up factor can be reduced somewhat by noting that not all habitat in the Mojave will support shrubs. If we assume that 10% of the ecosystem is desert floor, and we further assume that 10% of the desert floor is suitable for shrubs, the scale-up factor from study area to suitable habitat becomes $2.4 \cdot 10^4$ (Box 12.10).

Box 12.10 Spatial Scope Calculation for a Field Experiment to Discover the Effects of Competitor Density on Water Balance of Desert Shrubs (Fonteyn and Mahall, 1981)

Level	Replication	Units	A_o	A	Scope	T_o	T	Scope
Plant			$1\,m^2$			1 s		
Plot	1	1	$1\,m^2$	$100\,m^2$	100	1 s	1 hour	
Site	40	40	0.4 ha	2.5 ha	6.25	40 s	1 hour	90
Study	8 visits	320	0.4 ha	2.5 ha	6.25	8 hour	14 · 7 hour	12.25

Spatial scope	$SpSc$ =	$2.5\,ha/100\,m^2$	= 250
Units	n =	40/visit · 8 visits	= 320
Expansion factor	EF =	250/40	= 6.25
Sampling fraction	SF =	EF^{-1}	= 0.16

Mojave Desert
Tehachapi to Mojave Mountains: 300 km
Death Valley to San Bernardino Mountains: 200 km
Area $= 6 \cdot 10^4\,km^2$

Desert floor (10% ?) Area $= 6 \cdot 10^3\,km^2$

Shrub habitat (10% ?) Area $= 6 \cdot 10^2\,km^2$

$EF_{habitat} = 6 \cdot 10^2\,km^2/2.5\,ha = 2.4 \cdot 10^4$

Overall spatial scope $= (6 \cdot 10^4\,km^2)/(100\,m^2) = 6 \cdot 10^8$

Partitioned scope $= (n)(EF_{site})(EF_{ecosystem})$
$= (40)\ (6.25\ )\ (2.4 \cdot 10^6) = 6 \cdot 10^8$

Partitioned scope $= (n)(EF_{site})(EF_{habitat})(EF_{ecosystem})$
$= (40)\ (6.25\ )(2.4 \cdot 10^4)(100) = 6 \cdot 10^8$

The scale-up supported by rigorous statistical inference (a factor of 6.25) is far smaller than the scale-up based on judgment that sites were typical (a factor of at least $2.4 \cdot 10^4$). The limited scale-up achieved by this Fisherian experiment is typical in ecology, where logistics constrain experimental manipulation to small areas.

The limitations of Fisherian experiments in ecology are by now well known. Explanatory variables cannot always be manipulated. Confounding factors cannot be held constant by manipulation. Experimental units have not been homogenized by agricultural practice, so substantial replication is required to overcome variance among units. The rate of increase in heterogeneity with separation (scale) is poorly known, in contrast to agroecosystems, where estimates are available from the literature (e.g., Smith, 1938).

Logistics limit the number of experimental units, especially if units must be large to be realistic. In the wireworm experiment (Box 12.9), a relatively small plot size of $(22.86\,\text{cm})^2$ was judged representative. In the desert shrub experiment (Box 12.10), a far larger unit was needed—by a factor of $100\text{m}^2/(22.86\,\text{cm})^2 = 1900$. The size of the unit judged to be relevant to the dynamics of an ecosystem can be far larger. An example is deforestation (Likens et al., 1970) and reforestation (Schneider and Ayer, 1961), which alter the hydroregime at the scale of the watershed. In a forest, the effects of change in forest cover cannot be evaluated solely by small experimental units such as a single tree or even 1 hectare plots. In lakes the dynamics of phosphate loading, acid precipitation, or fishing mortality are appropriately studied at the scale of the lake, becoming the natural unit for intervention (Schindler, 1987). When logistics and realism are both considered, the result is often a limited number of experimental units, each of which is measured repeatedly. A repeated-measures design makes effective use of multiple measurements, even though it cannot address the problem of few experimental units. Dutilleul (1998) covers the topic of repeated measures in detail.

Scope calculations quantify the problem of scale in experimental ecology. The scale-up achievable by rigorous experiments that adhere to Fisherian principles can only be small compared to the scale-up from the experimental area to the ecosystem. This is less a problem in an agroecosystem because we have considerable knowledge of the past history of the system and of the spatial context in which an experiment occurs. From this we can judge whether the experimental results apply to the rest of the farm and, beyond that, to other farms. In statistical terms, we can judge whether the population of all possible outcomes, given the experimental protocol, includes other sites on the farm or other farms. In ecology, we have far less knowledge of the history of a site, its spatial context, and degree of heterogeneity among sites in the same ecosystem. Consequently, we have far less basis for using statistical inference based on the judgment that the infinite list of all possible outcomes for one experiment includes other sites in the ecosystem. To make effective use of ecological experiments, we need to bring larger-scale information to bear, as in a survey or monitoring program.

ANOTHER LOOK AT SECTION 12.5.3

Oceanographers use physical models to predict the value of a variable (e.g., width of upwelling zone along a coast), then design a measurement program to verify whether the value is correct. Fisherian experiments, based on the logic of the null hypotheses, result in a yes/no decision. Which do you find more convincing: testing whether an effect is present or testing a predicted level of effect?

12.5.4 Embedded Experiments in Ecology

Experimental programs typically have limited spatial and temporal extents (Kareiva and Anderson, 1988; Lodge et al., 1998) in exchange for greater control over confounding variables. The 1990s saw widespread recognition of the limited spatial scale of experiments relative to the ecosystems being studied. In a book summarizing the state of experimental ecology at the end of the 20th century (Resetarits and Bernardo, 1998), nearly every chapter mentions the problem. A subsequent book reviews the problem in freshwater, terrestrial, coastal, and marine ecosystems (Gardner et al., 2001), with an emphasis on mesocosm experiments. The classic solution to the problem is an *embedded experiment*, where experimental units are placed within some form of survey (Ellis and Schneider 2008).

Experiments can be embedded within surveys in several ways. A common way is to use an informal survey to identifying strata in which to place experiments. An early example is Menge (1976), who examined the role of physical factors in modifying the outcome of competition for space on rock surfaces by sessile intertidal invertebrates. Menge established transects downward through the intertidal zone at six locations differing in exposure to wave action in Maine, in the northwest Atlantic. Randomly placed quadrats were used to estimate percent cover along the transect. Menge scraped visible organisms from areas of rock to examine the rate at which barnacles and mussels filled space. Predator exclosures and competitor removal were used to examine the effects of wave exposure, tide height, the presence of a predator, and the presence of competitors. Percent cover on transects at each site was observed during the course of the experiment. This served as a form of control as well as providing a larger-scale context in which to interpret the results in the experimentally altered areas. The approach is classic in that surveys are often undertaken to choose sites, even if this effort is informal and unreported. This information allows the investigator to embed the experimental results into the larger-scale context of the informally surveyed area.

Another approach is to embed experiments into a survey designed to address a question. An early example (Schneider, 1978, 1985) is the use of predator exclosure experiments embedded into a systematic survey to estimate rates of prey depletion and prey community reorganization at coastal sites where shorebirds store energy as fat before undertaking nonstop migration across thousands of kilometers of inhospitable ocean. The motivating conservation question was whether shorebirds faced a declining prey base at the scale of the lagoon; so a hierarchically structured survey was conducted within a single flat in 1975 (refer back to Section 12.4.5), then extended to the scale of the entire lagoon by surveys on five flats in 1976 (575 cores), 1977 (590 cores), and 1978 (240 cores). The scope of the study (lagoon area/core area) was $34.6\,km^2/78.5\,cm = 4.4 \times 10^9$. The scale-up, from area measured, ranged from $EF = 34.6\,km^2/(240 \times 78.5\,cm^2/core) = 14.4 \times 10^4$ in 1978 to $EF = 34.6\,km^2/(590 \times 78.5\,cm^2/core) = 5.9 \times 10^4$ in 1977. Exclosures (half meter by half meter) were placed within or adjacent to 1 hectare study plots with known densities of invertebrates, known usage by shorebirds, and known rates of prey depletion. Exclosure experiments were conducted on the same July–September schedule as surveys. Several exclosures were run each year. No attempt was made to place exclosures randomly, since this would have been less informative than pairing the exclosures with survey sites. These experiments, paired with surveyed areas, showed that depletion measured by the survey was due to predatory removal (Schneider, 1978).

Embedding can be accomplished by weighting the results of each experiment by the relative frequency of comparable sites in the ecosystem. This approach was adopted in a

Table 12.1 Use of Surveys to Address the Problem of Scale in the Experimental Analysis of Ecosystems

1. Define the question. Define the boundaries of the system in space and time.
2. Assemble and inspect available information at all scales, including information at the scale of the system, including from remote sensing data.
3. Undertake pilot experiments embedded within the available survey information.
4. Define experimental units and variables needed to scale to the ecosystem.
5. Conduct a survey of the variables. Describe the result as a spatial model consisting of strata or gradients.
6. Assign experimental units to the range of conditions by strata or along the full range of the gradient.
7. Undertake scope calculations based on logistics and available resources to define levels and projected replication at each level. If experimental units are too large to measure completely, use survey design principles to choose sample locations within units. Examine the basis for inference at each level. Adjust the effort among levels as needed.
8. Complete scope calculations based on realized effort. Evaluate results in light of the magnitude of scale-up and inferential basis at each level.
9. Repeat the experiment if temporal scale-up is needed.

study of adult-juvenile interactions on intertidal flats in the lagoon at Manukau, adjacent to the city of Auckland (Thrush et al., 1997). These investigators used a formal survey to measure bivalve densities within a $(500\,\mathrm{m})^2$ experimental site. They then assigned experimental units to the full range of densities in the area, with emphasis on extremely high- or low-density sites. Not only did this embed experimental units into a well-designed survey; it also allocated effort in a balanced fashion to surveys and experiments. This contrasts with the highly skewed allocation of effort to surveys in the Plymouth study (Schneider, 1978), or to experiments in the barnacle study (Menge, 1976).

The embedding of experiments along density gradients, as at Plymouth (Schneider, 1978) and Manukau (Legendre et al., 1997; Thrush et al., 1997), provides a template for addressing the problem of scale in experimental ecology. Surveys can be used to address scale-up from measurements to experimental unit, in situations in which realistically sized units are too large to measure completely. Surveys can be used to place experimental units according to a spatial model that becomes the basis for scale-up beyond the limits of the experimental site. Table 12.1 provides a generic recipe for addressing the problem of scale in experimental ecology based on experience gained from the Plymouth and Manukau invertebrate study.

The examples in this section illustrate the range of ways that surveys and experiments can be combined to address problems of scale identified by experimental ecologists (Resetarits and Bernardo, 1998). The examples demonstrate a history of increasingly effective use of surveys in conjunction with experiments.

ANOTHER LOOK AT SECTION 12.5.4

Experimental ecologists allocate little or no effort to survey design, whereas fisheries biologists allocate nearly all their efforts to survey design. After offering some speculation on the reasons for this difference, make a list of criteria for deciding on how to allocate effort between experiments and surveys within a study.

12.5.5 Levels of Inference in Ecological Experiments

Fisher's innovations addressed the problems encountered in undertaking experimental research in agroecosystems. Fisher's prescription (randomization, replication, local control) introduces statistical control where confounding variables are beyond manipulative control. As we move from agroecosystems to ecosystems that are less altered by human activity, spatial heterogeneity increases and statistical control becomes more difficult. In agroecosystems, farming practice homogenizes the landscape at a small scale; an agricultural field is less heterogeneous than similarly sized plots of untilled land. Similar practices across farms flatten away larger-scale variance. Consequently, the results from one farm in Manitoba can be extended to other farms in Manitoba. In contrast, the results from one prairie site in Manitoba cannot be applied with such confidence to other prairie sites in the province. Stated more abstractly, variance grows with increasing spatial scale in any ecosystem; intervention flattens away the increase in variance with spatial scale as well as reducing the absolute amount of variance at the scale of fields or less.

The effect of an unflattened variance spectrum is that small-scale experiments rapidly lose their relevance at the scale of the ecosystem, compared to those with flattened variance spectra, such as agroecosystems. The solution is to quantify the magnitude of scale-up at several levels, introduce Fisherian experiments where possible, then use other forms of scale-up at those levels where Fisherian experiments are not possible.

Scope calculations provide a logical framework in which to evaluate the magnitude of scale-up and mode of inference at multiple levels. At each level defined by the experimental protocol we can quantify the scale-up and judge whether statistical inference is warranted. At appropriate levels we can scale up from limited measurements via inference based on the measurement protocol (as in Fisherian experiments) or via inference based on a finite frame (as in a survey). At other levels, where Fisherian experiments do not apply, we can apply appropriate forms of statistical inference to constrain interpretation (Carpenter, 1990), control for some sources of variation (Jassby and Powell, 1990; Dutilleul, 1993), or put a probability level on an outcome (Reckhow, 1990) or estimate of risk (Sutor, 1996).

This multilevel approach allows statistical inference to be restricted to the scale at which it is appropriate. An example is clear-cutting a 15.6 ha watershed in the 3000 ha Hubbard Brook Forest (Likens et al., 1970). This informative and influential experiment clearly did not conform to Fisherian design. Only one experimental unit was altered. We cannot undertake the Fisherian statistical apparatus to rule out chance effects among experimental units, including the confounding effects of variables that vary between watersheds.

We can, however, use the principles of survey design to rule out chance effects at the smaller scale of measurement units within the watershed. If we record soil water content at several points in Hubbard Brook watershed #2 and at several points in its neighbor, we expect the difference to increase substantially as soon as we deforest. The difference over time may be peculiar to just watershed #2 and its neighbor. To overcome this problem, we record soil water content at 10 randomly selected plots in watershed #2 and its neighbor. With this information we can compute the mean difference and the probability of obtaining this difference by chance. With a p-value less than 5%, we rule out chance (at the scale of plots within the watershed) as an explanation of the observed difference. This p-value cannot be used to rule out confounding factors at the larger scale of the watershed. To address confounding effects at the scale of the watershed, we would

need other sources of information. This might include time series of soil moisture levels in both watersheds, to which intervention analysis can be applied (Stewart-Oaten and Murdoch, 1986). We could also embed small-scale experiments into a watershed scale survey to establish the mechanisms leading to change in soil moisture due to removal of trees. We cannot use statistical inference to eliminate chance variation at the scale of watershed as an explanation of the observe difference, but we can use it to eliminate chance at the scale of plots within the watershed, something we cannot do with only one site in the watershed.

Scope diagrams are useful in evaluating the magnitude of scale-up and mode of inference at multiple levels. An example is the Mojave shrub study. In Box 12.10 the overall scope of 6×10^8 was partitioned into four components or scale-up factors. At the first level, scale-up is by a factor of 40. The basis for the scale-up is accumulation of units with randomly assigned treatments. At this same level the inferred scale-up to all possible plots within the perimeter of the study site is 6.25. The basis is statistical inference because we know the procedure to identify the plots. From this we conclude that the list of all possible experimental outcomes will include those from any plot in the study area. Beyond the perimeter of the site we have a much larger scale-up factor, estimated roughly at 24,000 possible plots in similar habitat in the same ecosystem. Knowing little about the study site, we would hesitate to judge that the list of all possible experimental outcomes resulting from the experiment extends to any 2.5 ha site with shrubs at these densities. To support the judgement we could turn to a survey, preferably one in which each experimental unit is linked to a variable that is measured at the scale of sites within the ecosystems.

Another example serves to illustrate the use of statistical and nonstatistical inference at different levels in an ecological experiment. The study was conducted at 21 sites scattered over four inlets on the Pacific coast of Panama, to examine whether migratory shorebirds deplete benthic invertebrate prey in the tropics (Schneider, 1985). The design was shaped by several constraints and by knowledge of benthic infaunal patchiness gained over five previous field seasons. The primary constraint was time, that of one investigator on a three-month fellowship. Sample size was limited by processing rate, which is slowed by the time-consuming activity of sorting and identifying benthic organisms to the lowest possible taxonomic level. Identification becomes especially time consuming in the tropics, where the number of species is high and there are no local guides to key out an organism to species.

Core samples were taken inside and outside a 1 m by 1 m roped exclosure at the beginning, middle, and end of a three-month period prior to departure of migratory shorebirds to arctic breeding grounds. Complete exclosures, even those with coarse wire mesh, disrupt flow and alter benthic densities (Nowell and Jumars, 1984). To eliminate these, shiny plastic rope was set along the top of stakes set 1 m apart in a square. Shorebirds were observed walking near but not through the exclosures. However, shorebird footprints were found inside one roped area, so a wire canopy, with sides open, was placed over the roped exclosure at this and one other site in March 1978.

This canopy successfully excluded birds, based on footprints found right up to the canopy but not under it. The canopies were placed at the two sites with the heaviest use, rather than randomly. Core samples 10 cm in diameter were taken haphazardly in each caged quadrat and in two or more quadrats immediately adjacent to the cage. Cores were washed on a sieve with a mesh of 0.5 mm, all organisms retained on this mesh were

counted. A total of 40 cores were collected in January, another 76 in late February and early March, and 142 in April. An additional 30 cores were collected and examined in the field at inlets ranging from Panama City eastward 150 km to the Azuero Peninsula. Box 12.11 shows the effort at each spatial scale in the study. As with previous examples, the effort at each level is used to compute areas and scopes.

The experiment had six levels with differing forms of inferential support. To evaluate the inferential basis for scale-up, we begin by partitioning the overall scope into components. The overall scope can be compute over any extent, but it is convenient to set this at the largest scale with a defined sampling protocol. In the scallop survey this was the 19,000 km^2 bank (Box 12.3). In the desert shrub experiment this was the 2.5 hectare study site (Box 12.10). For the prey depletion study this will be the 5 km^2 zone in which the samples were taken systematically from 21 sites. Box 12.11 shows the partitioning of effort by level. The partitioning of the expansion factor *EF* was guided by unit cancellation using the ratio of sizes of units (A/A_o), as shown in Box 12.11.

Box 12.11 Scope Calculations for a Predator Exclosure Experiment at Twenty-One Sites in Four Inlets on the Pacific Coast, Panama

			Area	Effort (*n*)			Replication
				Jan	March	April	
Core	$\pi(10\,cm/2)^2$	=	78.5 cm^2	40	76	142	258/180 = 1.43 core/quadrat
Quadrat	$(1\,m)^2$	=	127 cores	40	40	100	180/21 = 8.57 quadrat/block
Block	$(3\,m)^2$	=	9 m^2	21	21	21	21/21 = 1/site
Site	$(20\,m)^2$	=	400 m^2	21	21	21	21/4 = 5.25/inlet
Inlet	$(200\,m)^2$	=	40,000 m^2	4	4	4	4/1 = 4/zone
Zone	5 km · 1 km	=	5 km^2	1	1	1	
Coast	150 km · 1 km	=	150 km^2		1		

Level	Replication	Units	A_o	A	Scope	Type of Inference
Core		1	78.5 cm^2			
Quadrat	1.31 cores	1.31	103 cm^2	$(1\,m)^2$	$9.7 \cdot 10^1$	Statistical (finite frame)
Block	3.19 quadrats	4.19	329 cm^2	$(3\,m)^2$	$2.7 \cdot 10^2$	Statistical (design)
Site	1 block	4.19	329 cm^2	$(20\,m)^2$	$1.2 \cdot 10^4$	Statistical (finite frame)
Inlet	5.25 sites	22.0	1728 cm^2	$(200\,m)^2$	$2.3 \cdot 10^5$	Statistical (design)
Zone	4 inlets					
(March)		88	6912 cm^2	5 km^2	$7.2 \cdot 10^6$	Informal survey
(April)		176	13823 cm^2	5km^2	$3.6 \cdot 10^6$	Statistical (design)
Coast	3 zones	176 + 30	16179 cm^2	150 km^2	$? \cdot 10^8$	Partial survey

Spatial scope $SpSc = 5\,km^2/78.5\,cm^2 = 6.4 \cdot 10^8$ possible samples
Units $n = = 176$
Expansion factor $EF = 6.4 \cdot 10^8/176 = 3.6 \cdot 10^6$
Sampling fraction $SF = EF^{-1} = 0.28 \cdot 10^{-6}$

Spatial scale-up:

$$
\begin{aligned}
n \cdot EF &= n \cdot EF_{quadrat} \cdot EF_{block} \cdot EF_{site} \cdot EF_{inlet} \cdot EF_{coast} \\
&= 176 \cdot \frac{1 zone}{176 cores} \cdot \frac{127 cores}{quadrat} \cdot \frac{3^2 quadrat}{block} \cdot \frac{20^2 site}{3^2 site} \cdot \frac{200^2 site}{20^2 inlet} \cdot \frac{5 inlet}{0.2^2 zone} \\
&= 176 \cdot 0.72 \cdot 9 \cdot 44.44 \cdot 100 \cdot 125 \\
&= 176 \cdot 3.6 \cdot 10^6
\end{aligned}
$$

The basis for inference differs among levels in the study. At the level of cores within $1\,m^2$ quadrats. the inference is statistical, from the sample to a finite frame or list of all 127 possible coring sites within a quadrat. Cores were taken haphazardly within the quadrat so that each possible site has roughly the same probability of being cored. At the level of quadrats within blocks, inference is again statistical. Inference is from a sample to the infinite list of all possible outcomes, given the protocol for placement of quadrats within a block. The protocol was to select a quadrat haphazardly at a site, assign treatment to the quadrat, and define a $(3\,m)^2$ block as a treated quadrat surrounded by adjacent control quadrats. The assumption is that quadrats within such a block on an intertidal flat with little vertical gradient will have on average the same faunal composition. Placing controls around a haphazardly chosen quadrat is taken to be equivalent to defining a block of quadrats, then selecting one at random for treatment, leaving the rest as controls.

At one inlet, Culebra Beach, a block was defined around each of 3 quadrats at the same tide level because of the greater slope at this beach. At the level of blocks within sites, inference is statistical, again based on the design. Blocks were placed within sites having areas on the order of 20 m by 20 m with similar substrate throughout the site. Inference is from a haphazardly chosen block to the list or frame of all $(20/3)^2 = 44$ possible blocks that can be fit within a site that was visually homogeneous. At the next level, inference from the site to the inlet was based on placement of sites across the tidal gradient, thus capturing the strongest source of spatial variance in intertidal habitats. Sites were placed at roughly even intervals along a line running from high to low tide. Spacing depended on distance from high to low water and on the number of sites (six at three inlets, three at Culebra Beach). Inference to the scale of the inlet is thus to the infinite list of outcomes of an experiment where sites are placed evenly along the intertidal height gradient within an inlet.

At the next level, inference from inlet to zone was informal, based on judgment that the four inlets were representative of shorebird feeding areas within about 20 km of the Smithsonian marine station on Naos Island. The next level is temporal, with sampling in April repeated at haphazardly chosen locations within the same quadrats sampled in March. Inference at this level, like the previous, is based on an infinite list defined by the sampling protocol.

At the next level, the choice of zone along the coast was based on an informal survey of shorebirds and potential prey at coastal locations accessible from the Pan-American highway running southwest of the marine lab to Aguadulce, 150 km from the city of Panama. After the March sampling round, six beaches were visited between Punta Chame and Aguadulce to examine whether the invertebrates and shorebirds within the stretch of coast were comparable to those observed in the study zone.

At each beach the shorebirds present were noted. Cores were taken and sieved in the field to gauge whether invertebrate density, size, and taxonomic composition at the phylum/class level (family, in the case of polychaetes) differed among inlets within the stretch of coastline. Inference at this level was based on qualitative characterization of inlets within the entire stretch of coastline compared to the range of conditions seen in the zone where the study was conducted.

This example illustrates several benefits of multilevel analysis of experiments in ecology. *Multilevel analysis* quantifies the magnitude of scale-up, allowing comparison to other studies. In this study it is evident that the magnitude of scale-up increased at each successive level (Box 12.11). This was because the time to process samples limited the ability to increase effort at large scales. Multilevel analysis makes full use of accessory information. Information at scales other than that of the experimental unit is lost when the focus is restricted by the limits of Fisherian experimental design. Multilevel analysis leads to a better assessment of study strength and weakness, which will vary with level. In the Panama study, scale-up from core to quadrat was based on the statistical inference from sample to a finite frame of spatial units. The weakness at this level was use of haphazard sampling instead of a random number table to choose coring locations. At several other levels, inference was based on an infinite list of possible outcomes, given the design and sampling protocol.

Multilevel analysis shows that Fisherian criteria of randomization, replication, and local control were met unevenly. Local control was strong, since treatment and control were assigned to all 21 units. Replication to estimate variance among units was uneven across levels. Replication was adequate to estimate variance at the scale of cores within quadrats, quadrats within blocks, and blocks. Variance among blocks within sites could not be estimated. There were too few inlets to estimate variance at this level. Random (haphazard) selection of units addressed problems of confounding at the level of quadrat within block and block within site but not at larger scales. At the scale of quadrats within blocks, a simulation with typical core data suggested that surrounding a randomly placed quadrat with 8 control quadrats was equivalent to randomly selecting a treated site within a 9 quadrat block, but this was not investigated thoroughly. Randomized assignment of experimental units was not attempted at the level of sites within an inlet because of the known faunal heterogeneity from low tide line to high. At this level there were too few units to use statistical inference to eliminate the vagaries of spatial heterogeneity as the source of observed change in density of prey. Similar limitations applied at the scale of inlets within a zone, and zones within the Bay of Panama.

Confounding was a potential problem because only 2 of 21 sites had a canopy. Rapid growth and recruitment to edible size classes offsets by removal by shorebirds toward the end of their stay at this tropical location could have been due to the peculiarities of the two blocks that were assigned canopies. The conclusion would have been stronger if there were more canopy sites in the study. At the larger scales, the basis for inference was necessarily informal based on judgment and accessory evidence rather than inferential statistics.

ANOTHER LOOK AT SECTION 12.5.5

Find an experimental field study with treated and untreated units. Identify levels, identify the scale at which the statistical analysis applies, and evaluate the study at each you identified.

12.6 The Scope of Computational Models

Computational models are a well-known means of addressing questions at the scale of ecosystems. Computational models can be used to integrate available knowledge relative to the question at hand. They are employed to identify inconsistencies among disparate sources of information and to reveal gaps in knowledge. They aid conceptual simplification.

Computational models have a *spatial and temporal scope*, although this is not usually explicit. They have a minimum temporal resolution, set by the balance between realism and computational limits. Considerations of realism often turn on cyclic components: Does daily or seasonal variation need to be included in the model? Models have a spatial resolution that turns out to be severely constrained by computational limits. Though a large number of spatial units are desirable to capture spatial detail, no more than several hundred units are feasible without considerable computational resources, programmer skill, and programming time. Models have a potentially infinite temporal extent, but the extent of interest is usually set by the question, typically at time scales of years to several decades. The spatial extent is set by the question, usually at the scale of the ecosystem or extent of environmental impact.

Scope computations were undertaken for 13 models listed by Auble et al. (1995), who list temporal resolution and duration for each model. They provide enough spatial information to determine scope (number of spatial units). The spatial extent was taken from maps for those models with a named location. The temporal resolution ranged from days (two models) to weeks (one model), months (five models), seasons (two models), or years (three models). The temporal extent ranged from 5 to 70 years. The number of spatial units was remarkably small, ranging from one to just nine (river reaches in two connected river basins).

In diagrammatic form, the scope of any one model will consist of a rectangle within which dynamics can be computed. The upper-right corner is anchored to the time and space scale of the question. Figure 12.4 shows the scope diagrams for two of the models listed by Auble et al. (1995). The other models had smaller spatial scopes. All the models had about the same temporal scopes. The diagram displays the characteristic strengths and weaknesses of computational models in addressing ecological questions. Computations can be made over a wide temporal scope at space and time scales relevant to any ecological question. However, the spatial scope is quite limited. Models do not include the scale at which data are typically gathered (compare Figure 12.4 to Figures 2.2, 12.1, and 12.2). Thus an unexamined scale-up usually occurs when data are used in these models.

There are several ways to bridge the gap between computational models and the spatial scale of field data. One is to use remotely sensed data, obtained from aerial photogrammetry or from satellites (Innes, 1998; VandeCastle, 1998). This data will have a spatial extent comparable to that of a computational model. Depending on the type of sensor, data can have a resolution fine enough (on the order of meters or tens of meters) to match data gathered on the ground or at the sea surface. Cloud cover and lack of ground truthing of the image are problems. Another way to bridge the gap is to embed a smaller-scale model (Wessman, 1992) into the larger model that is anchored to the scale of the problem. A promising but relatively unexplored way to bridge the gap is to apply power laws such as species-area relations.

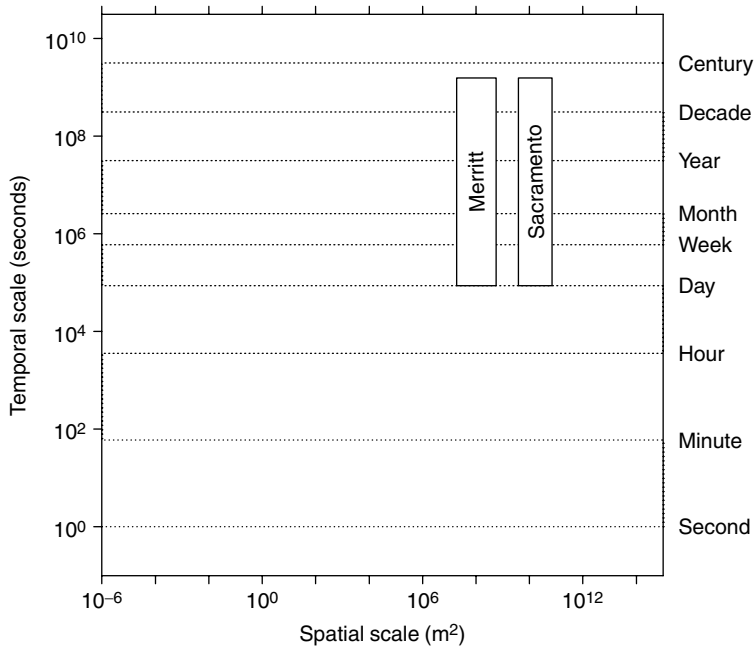


FIGURE 12.4 Scope Diagram for 2 of the 15 Models Listed by Auble et al. 1995.

ANOTHER LOOK AT SECTION 12.6

Open an issue of *Ecological Modelling* or similar journal and pick an article from the index. Is there enough information in the article to draw a scope diagram? If there's not, continue to the next article until you find one with enough information or until you reach a limit of five. Comment on the utility of model studies for which the spatial and temporal scope cannot be determined.

12.7 The Scope of Integrated Research Programs

Descriptive studies, experiments, surveys, and modeling efforts often take place within integrated research programs that combine several types of studies. Scope diagrams are a useful way of summarizing the activity within these integrated programs. These diagrams display much that is not otherwise apparent: the relation of the components, their degree of complementarity, and gaps in the overall program.

In the late 1980s and early 1990s, investigators at the New Zealand Institute of Water and Atmospheric Research (NIWA) undertook a series of studies to investigate and monitor the effects of release of toxic materials into estuaries (see Figure 2.2). Early in the program an experiment was conducted to estimate the effects of chlordane release on benthic invertebrates in the lagoon at Manukau, adjacent to the city of Auckland (Pridmore et al., 1991). The study was limited by logistics to two large experimental units, measured repeatedly. The scope diagram (Figure 12.5) shows the support and extent of the experiment. Multiple measurement of the experimental units allows the authors to exclude chance variation within the treated and untreated units as an

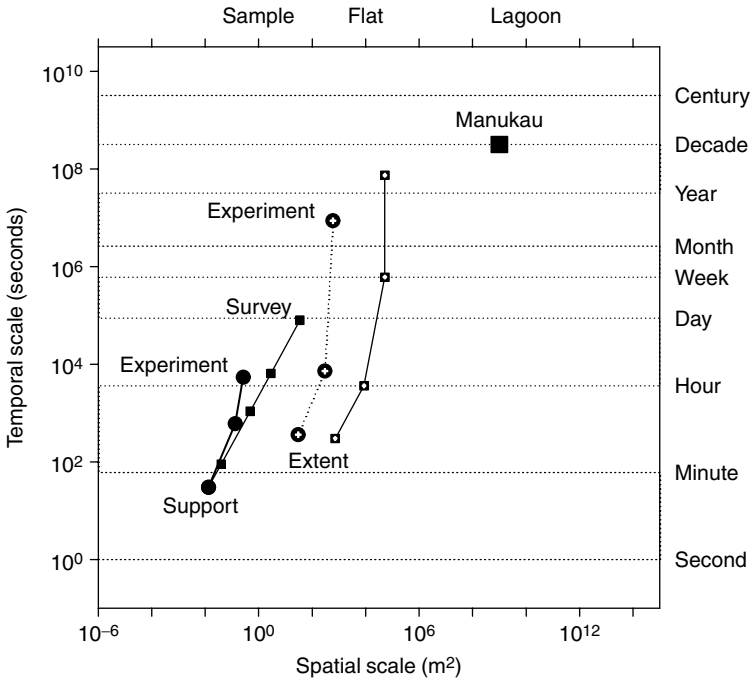


FIGURE 12.5 Scope Diagram for a Field Experiment and Related Survey. Purpose of experiment was to measure the effects of chlordane on benthic invertebrates (Pridmore et al. 1991). Purpose of repeated survey was to characterize spatial variance of lagoon ecosystem in which chlordane experiment was conducted (Roper et al. 1992).

explanation for the decline in bivalve abundance. The spatial layout (one treated and one control area) did not permit random effects at the scale of the units to be addressed statistically. Several lines of evidence (including reversal of decline as chlordane disappeared) supported the conclusion that the significant drop in bivalve number in the treated plot was not due to random effects at the scale of the experimental unit. Fisherian replication to address the problem would have been prohibitively expensive and would have required releasing at least 10 times as much chlordane into the environment. The NIWA group addressed the problem by quantifying natural variability at the time and space scales of the experimental units and by undertaking process-oriented studies of this variation. A survey that quantified spatial variability at multiple scales (Thrush et al., 1989) was repeated every two months for just over two years (Roper et al., 1992). The support of the survey exceeded that of the experiment (Figure 12.5). The support and extent of the survey bracketed the space and time scales of the experiment. Thus, the experimental results can be compared to natural variability as measured by the survey in much the same way that treatment effects are compared to uncontrolled process error when forming an F-ratio in a Fisherian experiment. The changes in the experiment exceeded those observed in the survey, evidence that the change observed in the experiment was not due to chance variation at the scale of the experimental unit. The spatial scale of the experiment and the survey was substantially smaller than the entire lagoon (Figure 12.5). The experimental site was not chosen randomly, so probabilistic sampling cannot be used to address whether the results apply to the entire lagoon. Scale-up was qualitative in the sense that the direction of the experimental result was

considered applicable to the entire lagoon, even though the magnitude of the decrease might not have been representative.

In this program as in others to isolate the effects of human activity on ecosystems, high spatial variance in density is one of the largest sources of uncertainty. Most of this variance results from natural processes rather than measurement error. A series of studies were undertaken by the NIWA group to isolate processes responsible for spatial variation in benthic invertebrates on intertidal flats in northern New Zealand. At the Manukau lagoon, one obvious source of spatial variation was the digging of feeding pits by rays, which concentrate their activity in areas on the order of hectares. Pits have a characteristic size (on the order of 30 cm) and lifetime (filling rapidly in the first week), so a directed survey (Figure 12.6a) was undertaken, consisting of six visits over 12 days to two sites, each with five pits measured by four core samples (Thrush et al., 1991). Ray pits differ in degree of initial defaunation, so an experiment was conducted to estimate recolonization rate of completely defaunated plugs of sediment similar in size to pits (Thrush et al., 1992). The spatial scope of the survey exceeded that of the experiment (Figure 12.6a), as was the case with the chlordane experiment. The effects of rays were then compared to those by migratory shorebirds, which also feed in invertebrates but do not excavate pits. Thrush et al. (1994) used canopies to exclude birds only, cages to exclude both, and control sites. This study had a spatial scope similar to the recolonization studies, but a greater temporal scope (Figure 12.6b). Studies in other benthic habitats suggest that spatial variation in juvenile settlement arises from interactions with adults, which interfere with or consume juveniles. Thrush et al. (1996) investigated this by examining rates of juvenile colonization of buckets with known densities of adults. The scope of this experiment (Figure 12.6c) was similar to the previous experiments.

In these experiments and surveys, the level of effort was set by constraints of time, notably the time taken to sort and count benthic organisms. The limitation on number of units, together with the need for replication to estimate a variance at each level, limits the spatial and temporal scope. These limits, together with allocation decisions at each level, create similar scope diagrams (Figures 12.5 and 12.6). The scope of this body of work is conveniently summarized by drawing a polygon that encloses both the support and the extent of all the studies (Figure 2.2). This diagram displays the gap from experimental and survey results to the system of interest, the Manukau lagoon.

To address the problem, a collaborative project was undertaken, with the goal of designing and executing an experiment to bridge the gap from data to the scale of the problem. Known factors governing the bivalve density in the lagoon included predation by rays, predation by birds, response of juveniles to presence of adults, and several physical factors (sediment composition, eelgrass presence, duration of exposure by the tides, wind mixing of the water) that were in turn a function of height above mean low water. A seven-factor experiment was clearly not possible: just two levels per factor results in $2^7 = 128$ treatments in a full factorial design, before replication. Attenuation to two factors per block results in 2^2 treatments per block; there are $7*6/2 = 21$ pairs of factors, which results in 84 treatments before replication.

The solution was to embed an experiment into a survey-based estimate of the two-dimensional density-scape with hills of high density and valleys of low density across a 250 m by 500 m area. The density-scape was estimated from cores place randomly in 25 m by 25 m sectors; the estimate was then used to assign a density value to each sector rather than treating sectors as blocks. Adult bivalve density was altered experimentally

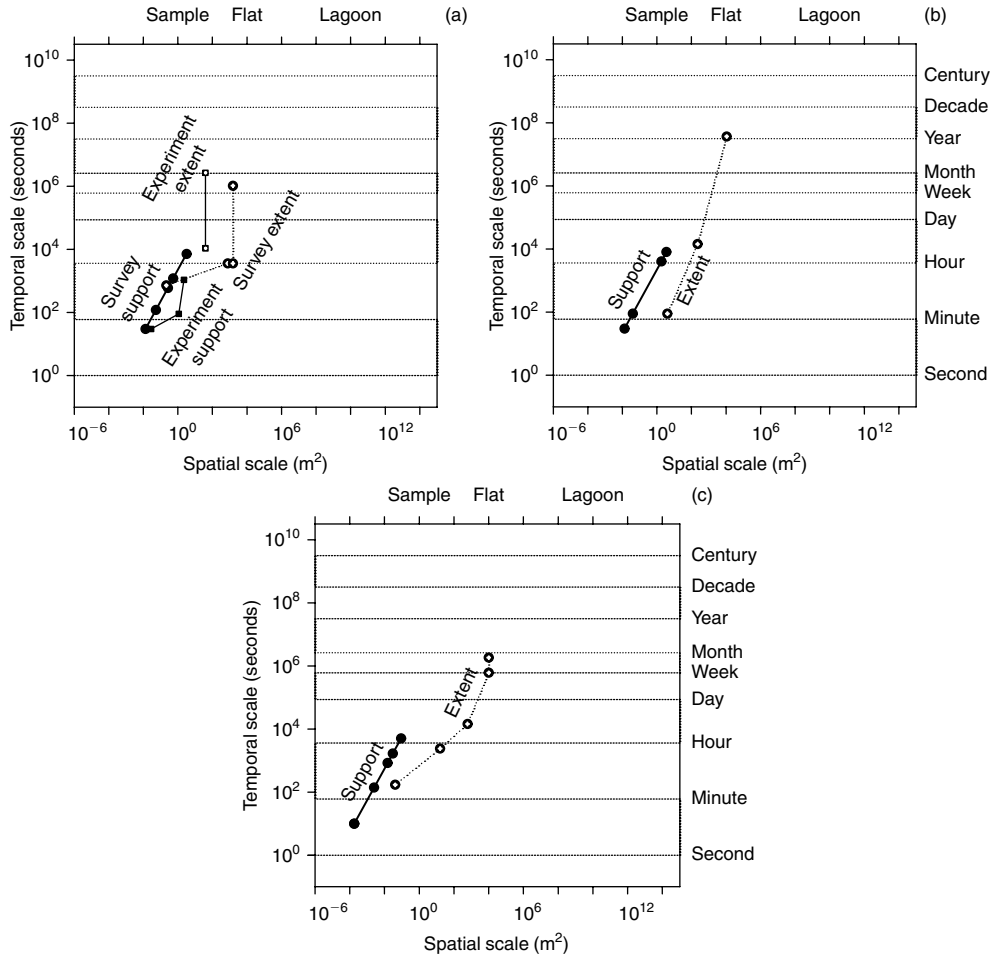


FIGURE 12.6 Scope Diagrams for Experimental Investigation of the Spatial Dynamics of Benthic Fauna at Manukau Lagoon, New Zealand; (a) Survey estimate of recolonization rate of pits dug by rays feeding on intertidal flats (Thrush et al. 1991) compared to experimental estimate of recolonization rate of completely defaunated plugs of sediment (Thrush et al. 1992); (b) Exclosure study to compare the predatory effects of rays to migratory shorebirds (Thrush et al. 1994); (c) Experimental estimate of rates of juvenile colonization of buckets with fixed densities of adults (Thrush et al. 1996).

in 22 sectors, which were chosen randomly from three categories: high density, medium density, and low density. This forced greater representation of the peaks and valleys without sacrificing the advantages of probability sampling of sectors (each sector had a known probability of being selected).

Note that sector selection was based on a value fixed by the density-scape, not on the (random) value for each sector. Fixed categories on a continuum were used in the seminal paper on linear regression (Pearson and Lee, 1903). The advantages of the design were: (1) results can be scaled up to the entire area via the density-scape; (2) there were no border effects through location of treatment at the edge of patches; (3) the experimental results were integrated with a substantial suite of physical variables for which the field or density-scape could be computed across the area; (4) results can be scaled up via physical variables that can be computed across the entire lagoon; and (5) survey and experimental efforts reinforced each other through quantitative integration, moving beyond qualitative

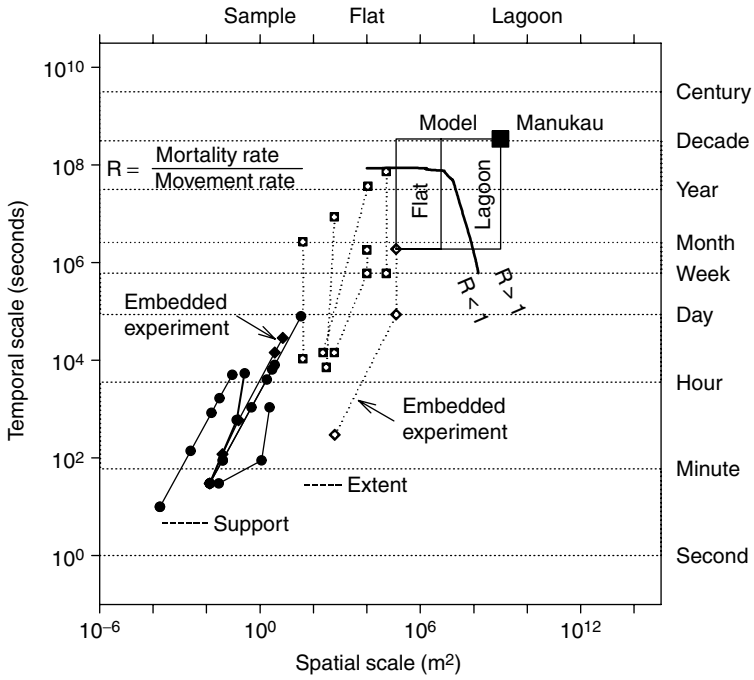


FIGURE 12.7 Scope Diagram for an Integrated Research Program. See Figure 2.2c for program rationale. Support and extent from Figures 12.5 and 12.6 have been redrawn for comparison with each other and with the scope of the embedded experiment (Thrush et al. 1997). Heavy line separates domain where bivalve mortality exceeds rates of movement ($R > 1$) from domain where movement exceeds mortality ($R < 1$).

integration, as in Figure 12.5. The number of sectors (22) for a random survey was the same as the number of pairs used to generate fixed categories on a continuum, but the latter generated results that could be scaled to the entire area via the density-scape.

Figure 12.7 shows the support and extent of this experiment, compared to previous studies (Figures 12.5 and 12.6). The summary diagram in Figure 2.2b proved particularly useful in a workshop setting, making evident the scope of previous studies (width of the gap from the support to the extent) and the scope of research to bridge the gap from previous work to the scale of the lagoon. The spatial support of the embedded experiment was comparable in area to previous studies. The spatial extent was greater, and hence the scope (distance from support to extent in Figure 12.7) was larger than previous studies. Statistical inference underpins all, not just part, of the spatial expansion factor of $EF = (250\,\text{m})(500\,\text{m})/7\,\text{m}^2 = 17835$ (Figure 12.7). In this calculation the support is $(0.0133\,\text{m}^2$ /core) (3 cores/quadrat) (4 quadrats/sector) (22 sectors/plot) (2 visits/plot) = 7 m^2. The embedded experiment increased the scope of statistical inference relative to previous work but did not close the gap between the extent of field work and the system of interest, the entire lagoon (Figure 12.7). The experiment was designed with an eye toward scale-up via physical variables (e.g., bedload transport) that can be computed at the scale of the flat (approximately 2 km by 3 km in area) or the entire lagoon. Consequently, the extent of the embedded experiment is shown (Figure 12.7) as connected to a model of the flat and the lagoon. The temporal scope required of the model is readily accomplished, but the spatial scope required $(3.4 \times 10^8\,\text{m}^2/(1.25 \times 10^5\,\text{m}^2) = 2.7 \times 10^3)$ is large compared to that created by computational limits (refer back to Figure 12.4).

Can a computational model constructed for a single flat be connected to a model for the entire lagoon? The experiment was carried out on Wiroa flat, for which the spatial scope of a model relative to the experimental area would be $EF = (2\,\text{km})(3\,\text{km})/((250\,\text{m})(500\,\text{m})) = 48$. The scope of a lagoon model relative to the flat model would then be $3.4 \times 10^8\,\text{m}^2/6\,\text{km}^2 = 57$. Both scopes are consistent with computational limits on spatially explicit dynamic models.

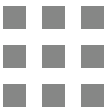
Figure 12.7 illustrates an additional application of scope diagrams, which is the display of the relative importance of competing rates based on a conservation equation (Schneider et al., 1997). For bivalves in the Manukau, the ratio of a demographic rate (mortality) relative to a kinematic rate (lateral movement with bedload flux) was of interest. The ratio of the rates was computed from available data by T. Bult (unpublished) using a numerical method described elsewhere (Schneider et al., 1999). The line at which the two rates were equal (refer back to Figure 2.2) was plotted across the scope of the model (Figure 12.7). The plot shows at a glance that flux rates due to bedload transport need to be considered, even at the relatively large scales of the flat and the lagoon model. Figure 12.7 shows that contrary rates (flux rates and demographic rates) are roughly equal in strength. If these act with a lag, complex dynamics can be expected (Schneider 2001b).

ANOTHER LOOK AT SECTION 12.8

The integrated research program displayed in Figure 12.7 included data from remote sensing but did not include mesocosm experiments. Make a copy of Figure 12.7 and draw in the support for a typical mesocosm experiment from the literature (e.g., Chen et al., 1997, Petersen et al, 2009). Then add the support and, if possible, the extent of an example of remotely sensed data. Do these components strengthen the information along the diagonal in Figure 12.7 or do they increase the information at space and time scales not shown in the figure?

Defined Concepts and Terms for Review and Future Reference

____ embedded experiment
____ exhaustive versus probabilistic survey
____ expansion factor and sampling fraction
____ experiment
____ multilevel analysis of experiments
____ partitioned scope
____ spatial and temporal support
____ spatial and temporal scope of a computational model
____ spatial and temporal scope of an experiment
____ spatial and temporal scale of a Fisherian experiment
____ spatial and temporal scope of a monitoring program
____ spatial and temporal scope of a set of measurements
____ spatial and temporal scope of a survey
____ unit and frame



PART

IV

Models

13 Equations

For the sake of persons of these different types, scientific truth should be presented in different forms, and should be regarded as equally scientific, whether it appears in the robust form and vivid colouring of a physical illustration, or in the tenuity and paleness of a symbolical expression.

—J. C. Maxwell, Presidential address on "Mathematics and Physics" at the Liverpool meeting of the British Association, 1870

13.1 Synopsis

An equation expresses an idea that can be used to make calculations about scaled quantities. The use of equations to make calculations from ideas differs from the analysis of equations to develop theory. The former is easier than the latter, it requires little mathematical training, and it is guided by reasoning about measurable quantities. This chapter is about calculations based on ideas. There will be no treatment of equations divorced from calculation or units.

Equations, like any foreign language, are unintelligible on first encounter. Practice, together with the use of graphs, increases facility in understanding the ideas expressed by equations. Consistent use of symbols, adeptly chosen, contributes to the ready comprehension of mathematically expressed ideas about quantities. Other aids to comprehension are stating the idea in words, making a typical calculation, graphing the equation, and identifying dimensions.

Equations that express ideas about quantities must be dimensionally homogeneous: The sum of 2 cabbages and 3 kings cannot be calculated. Inconsistent units and dimensions guarantee an incorrect calculation.

Writing equations resembles the writing of sentences. The goal of both forms of writing is clear expression within the rules of syntax. Equations use a larger set of symbols than does prose, with no rules governing the meaning of each symbol. This is a source of confusion that makes it difficult to communicate using equations. Solutions to this problem include clear notation, complete listing of symbols with units, and adherence to the rules for units. Together these solutions contribute to better communication of quantitative ideas.

Equations that relate quantities to one another arise from several sources. One source is exploratory analysis of data, resulting in empirical equations. Another source is direct reasoning about quantities, although this is less common in ecology. The procedure is to state response and explanatory variables, develop a simple relation, check for internal consistency, check calculations against data, and revise the equation as needed. Writing equations is a natural extension of reasoning with scaled quantities; like many activities, it grows easier with practice.

doi: 10.1016/B978-0-12-627865-1.00013-1

13.2 The Use of Equations

Equations have a number of uses in ecology. One is to show precisely how one quantity was calculated from another. An example is the calculation of feeding rate (r_F = grams food per kg of body mass per day) from metabolic water turnover (r_W = milliliters per kg of body mass per day) measured from change in concentration of stable isotopes. Kooyman et al. (1982) used the following equation to show, in an economical form, exactly how they calculated the feeding rate of penguins from isotopic measurements of water turnover:

$$r_F = \frac{r_W}{P_W + E_F E_M M_W} \tag{13.1}$$

The components of this equation are:

P_W = 4.0 ml preformed water per gram of food
E_F = 17.6 kJ per gram of food
E_M = 0.8 kJ metabolized per kJ ingested
M_W = ml water produced per kJ metabolized

With this equation, another person could take any measurement of water turnover and calculate a feeding rate that could be compared to those reported in Kooyman et al. (1982). The equation makes the study reproducible in this sense of the word. The equation opens the result to scrutiny by allowing comparable numbers to be calculated. Its absence makes the study irreproducible.

A second use of equations is to work from premises to conclusions using the rules of mathematics. This use rests on one of the principal strengths of mathematics, which is that its logical structure guarantees that conclusions will be consistent with premises, if the rules are applied correctly. This use, to develop theoretical conclusions, differs in several ways from the first, or demonstrative, use. Theoretical use is confined to a few specialized journals; demonstrative use should be found in any research report where one quantity is calculated from another. Theoretical use aims at general conclusions; demonstrative use aims at clear definition of a particular situation. Theoretical use relies heavily on mathematical analysis, sometimes in highly sophisticated forms. Demonstrative use rarely requires sophisticated mathematics. Theoretical and demonstrative use both require units. Unfortunately, theoretical use as it is practiced in ecology often fails to include units. Equations that lack units describe nature in a metaphorical way that precludes making calculations for a given situation. Such equations cannot be tested against measurements.

A third use of equations is statistical testing of ideas against data. Almost all the commonly used statistical procedures in ecology are model-based, even though many (e.g., Chi-square tests) are presented in texts as though no model were involved. Chapter 15 describes the statistical evaluation of equations relative to measured quantities.

A fourth use of equations is to develop scaling functions. These have many uses, including converting a functional relation to a different spatial or temporal scale computing results at large space and time scales from experimental results or local observations, and developing theory. Scaling functions are a recurring theme of Part IV of this book.

ANOTHER LOOK AT SECTION 13.2

Of the four uses listed for equations, how many have you encountered or had occasion to use?

13.3 Verbal, Diagrammatic, and Formal Expression of Ideas

Biological concepts can be expressed in words (an informal or verbal model), or in graphs (a diagrammatic model), or in equations (a formal or mathematical model). These three forms of expression are related to one another (Figure 13.1).

In ecology, the most common mode of expressing an idea is verbal. Sometimes a verbally expressed idea is accompanied by a diagram. Ecological ideas are also expressed in symbolic form as equations, but this form of expression is not always accompanied by the diagrammatic or verbal expression that would make it more comprehensible. Often the equation is left for the reader to decipher, with the unfortunate result that the reader simply skips right over the equation (and the idea). One solution to this problem is to annex the math to an appendix, then explain the ideas in words. Another solution is to state the idea in words and pictures accompanying the equation. This latter solution is becoming increasingly prevalent in ecology texts (e.g., Ricklefs and Miller, 2000; Case, 2000).

Each of the three forms of expression (Figure 13.1) has its own advantages and disadvantages. We use words every day, and so verbal models are the most readily used and easily understood form of expression. The disadvantage is that verbal models can be wonderfully fuzzy. The fuzziness of verbal models becomes apparent as soon as one tries to depict a verbally stated idea in the form of a graph. Verbal fuzziness also becomes apparent when the units and procedural statements for a quantity are examined. Two quantities that appear to be no more than different names for the same thing may in fact be completely different. And two quantities that appear to be different may in fact be alternate names for the same thing.

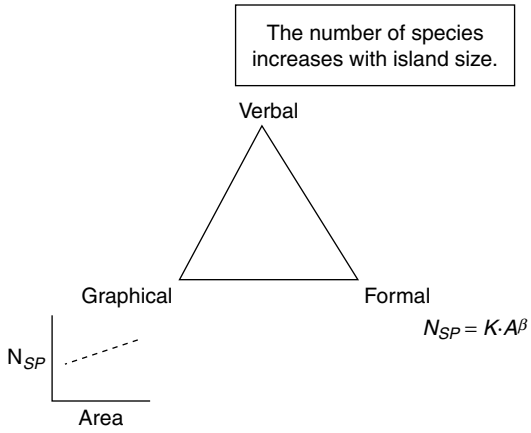


FIGURE 13.1 Verbal, Graphical, and Formal Model of a Biological Concept, the Species-Area Relation (SAR).

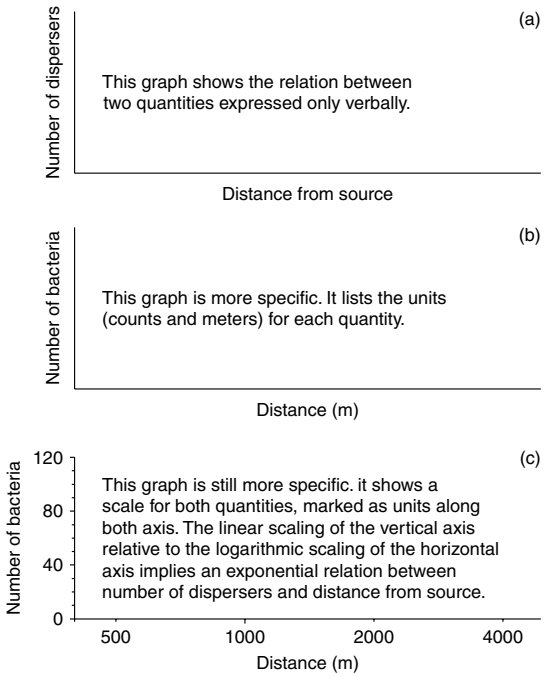


FIGURE 13.2 Range in Information Content of Graphs; (a) Axes for verbal model only; (b) Axes for graphical model; (c) Axes for formal model and data.

The advantage of a graph is that, like a verbal model, it is quickly grasped. Relative to verbal models, graphs convey more information about how one quantity is related to another. But important features of the relation may still be omitted, and the definition of quantities can remain vague. The degree of vagueness is assessed by asking: How thoroughly have the quantities on each axis been defined? Figure 13.2 shows three graphs, each more informative than the previous graph. The first graph (Figure 13.2a) is useful for general discussion of concepts. But graphs that lack units must be replaced when the discussion proceeds to knowledge based on measured quantities. The graph in Figure 13.2b takes this step by displaying the units of measured variables. The next step (Figure 13.2c) is to display the units on a scale. For example, are we considering the dynamics of local dispersal (e.g., distance in cm from a tree releasing seeds), or are we considering (as in Figure 1.3) the dynamics of dispersal at a larger scale? The display of units needs to be consistent with the dynamics that we want to understand. A further advantage of graphs such as Figure 13.2c is that they allow us to plot measurements against theory. Graphs such as Figure 13.2c are central to the practice of ecological science because they connect measurement to formal models expressed as equations.

ANOTHER LOOK AT SECTION 13.3

An accomplished theoretical ecologist, Simon Levin, noted that one can look at an equation for a long time until one suddenly "gets it." How does this statement compare to your experience?

13.4 Reading Equations

Equations, like any foreign language, make no sense upon first encounter. Comprehension and eventual facility result from repeated encounter and practice in using equations. There are several ways of hastening this process of learning and eventually "speaking" the language of scaled quantities. One important technique is translating equations into words. Some effort may be required to do this with complex equations, but the entire process of translation into words always goes more quickly each time an equation is encountered. Time put into translation on first encounter reduces the time required on subsequent encounters. Some equations or terms in an equation are so common in a field of research that the symbols or group of symbols can eventually be read into words with no effort.

The way to translate equations into words is to write down each symbol, state its name in words, and state its units, if it has such. Equations consist of two or more groups of symbols called *terms*, so the next step is to write out the name of each term, together with its units. The terms of an equation are connected to one another by addition, subtraction, or equivalence. One of three signs (+, −, or =) separates the terms in an equation. Each term consists of a symbol or several symbols connected by multiplication, division, or exponentiation. Here is an example of the use of substitution of words and units to read an equation expressing the idea that populations grow exponentially:

$$N_t = (N_{t=0})\, e^{rt} \tag{13.2}$$

This equation has two terms separated by the equal sign. The term on the left consists of a single quantity, represented by the symbol N_t, read as "the number at time t." The term on the right is the product of two quantities, $N_{t=0}$ and e^{rt}. The quantity $N_{t=0}$ is read "the number at time zero." At this point it helps to associate the symbol $N_{t=0}$ with an image of a group of, say, protozoa. Then associate the symbol N_t with an image of the same group and its progeny at a later time. The quantity e^{rt} is read "the crude rate of increase." An equivalent symbol for this quantity is R, used in many ecology texts. It is the percent increase after time t. If a population increases by 20% from time $t = 0$ to time t, then $N_t/N_{t=0}$ has a value of 1.2, e^{rt} is 1.2 or 120%, and hence N_t is 1.2 times greater than $N_{t=0}$. So we can now translate this equation into words: "The number of organisms at time t depends on the initial number and on the crude rate of increase, expressed as a percentage of the initial number."

The quantity e^{rt} consists of three symbols: the pure number e and two quantities. Time t is a *variable quantity;* it can take on many values. The intrinsic rate of increase r is a *parametric quantity*: It takes on a single value, one that holds across the many values of the variable (*parameter* means "across measured values"). This latter can be visualized as the rate that would apply at any instant in a population so enormous that new individuals appear to recruit continuously. The expression e^{rt} can now be translated into words. It is "the crude rate of increase, expressed as the product of time elapsed and the instantaneous rate, taken as a power of e, the base of natural logarithms." This is cumbersome, but it helps to use full translation on the first two or three encounters. Eventually, the entire expression, e^{rt}, can be shortened to something like "the crude rate of increase as a function of the instantaneous rate." Or even "the percent increase based on the instantaneous rate."

One of the most useful methods for learning to read and understand equations is to use a calculator. An example was shown in Box 4.1. This method is more effective if one retains the units when making the calculations, rather than stripping away the units and just using numbers. Retaining the units preserves the sense of working with visualizable quantities and thus contributes to comprehension of the idea expressed by the equation.

Another aid in understanding equations is to graph the equation by plotting the term on the left side against one of the quantities on the right side of the equation. Either terms or symbols from the right side can be used. In the example of exponential increase in population size, N_t can be plotted against r, against t, or against e^{rt}. Some readers may be puzzled about why one would want to plot N_t against e^{rt}, but it must be remembered that "obvious" depends on experience. To draw an analogy, it might seem obvious that verbs must have tenses, but anyone acquainted with an Oriental language knows that conjugating verbs is by no means necessary to convey temporal relations. The use of tenses seems obvious only to someone thoroughly familiar with an inflected language. Similarly, the relation between N_t and e^{rt} seems obvious only after some practice in visualizing equations as graphs. The relation between N_t and t is harder to picture, but with practice this too becomes so familiar that it can be visualized immediately and does not need to be graphed.

Another method that aids in reading and understanding equations is to write out the dimensions of each symbol immediately below the equation. This adds another level of understanding by displaying the role that mass, time, length, area, energy, electrical charge, and so on play in the concept expressed by the equation. I found this technique useful in trying to grasp equations from geophysical fluid dynamics, where the equations look formidable but tend to be about readily visualized quantities such as mass, energy, or spin (called vorticity). Eventually the association of a group of symbols with a group of dimensions (such as mass flux = $\mathrm{M\ L^{-2}\ T^{-1}}$) strengthens to the point where it becomes "obvious" that the equation is about energy. But this only becomes apparent after some familiarity is gained by writing out the dimensions underneath the symbols on the first few encounters.

Reading an equation is like translating a foreign language. Each symbol is like a word that must be translated. I find that it is better to directly associate the symbol with an image than to associate the symbol with a word only. Next, which words have the same function? (Which are nouns? Which are verbs?) Similarly, which symbols belong to the same dimension? Once each word (symbol) has been grasped and the grammatical relations (dimensions of symbols) have been identified, the next step is to make sense of the whole sentence (equation). Even after all the words (symbols) are understood, it might not be immediately clear what the sentence (equation) is saying until an effort is made to understand it at this level.

Another step in understanding an equation about quantities is to identify the time and space scales at which the equation applies. If the symbol ζ (zeta) represents the efficiency of lions in capturing zebras, at what space and time scales are we to measure this efficiency? It seems natural to measure the fraction of zebras caught per lion-day, repeating this for a month or so to obtain an average value. But will we be comfortable in using the value we obtain at the longer time scale of, say, a year? And if our measurements were made in a limited area, are we willing to use the same efficiency at the larger spatial scale of the entire zebra population? It seems quite likely that the capture efficiency measured in an area with many zebras may overestimate the efficiency by those same lions at the spatial scale of the entire zebra population, which includes areas with few zebras.

Analyzing an equation as though it were an idea about quantities rather than as though it were a mathematical abstraction brings us face to face with the question of the choice of appropriate time and space scales for the equation. If we are simply using the equation as a recipe for calculating one number from other numbers, without using units or dimensions, then we easily miss something important, which is the space and time scale at which the equation applies. If we treat the equation as expressing an idea about measurable quantities with a particular scope (range relative to resolution), we are aware of the limits of the equation. It is a curious fact that ecology throughout most of the 20th century paid no attention to time and space scales (Figure 2.3). The tradition has been to apply equations as though they had neither units nor dimensions (Table 1.2). The tradition was to treat equations as though they had no scale, even though measured quantities clearly do have a scale set by the measurement protocol that determines their scope (range/resolution). This tradition will eventually disappear. Its demise can be hastened by adopting the convention that all quantities in a concept expressed as an equation must have a stated resolution and range, with clear statement of spatial and temporal attributes.

ANOTHER LOOK AT SECTION 13.4

Try putting Equation 13.2 into your own words. Then try Equation 13.1.

13.5 Dimensional Homogeneity

The principle of dimensional homogeneity, introduced in Chapter 6, is an important tool in working with equations. The principle is used to check the validity of an equation, work out units and dimensions of unknown quantities, and gain a better understanding of the idea expressed by an equation. The principle of *dimensional homogeneity of an equation* states that an equation about measured quantities must be dimensionally homogeneous. Terms that are added (or subtracted) must have the same units and dimensions (Rule 6, Table 6.7), an extension of the apple/orange principle from Chapter 4. All the terms on one side of an equal sign must have the same units and the same dimensions as all the terms on the other side (Rule 7, Table 6.7).

One application of the principle is to check the dimensional homogeneity of an equation before using it to make a calculation. If the terms in an equation have different dimensions, the results of computation cannot be trusted. If the equation is not homogeneous, it needs to be repaired before it can be used. Dimensional homogeneity is routinely used in engineering applications, where computational errors have real consequences. Dimensional homogeneity is just as appropriate in applied ecology. Dimensional homogeneity does not guarantee that a calculation is correct, but heterogeneity will guarantee that most calculations are incorrect.

To check for dimensional homogeneity, one writes the equation, places the dimensions beneath each symbol, then applies the rules in Table 6.7 to make sure that all terms have the same dimension. Riggs (1963) developed a convenient list of rules for checking the dimensional consistency of equations. These are paraphrased in brief form in Table 13.1. This list is shorter and easier to use than the more complete list in Table 6.7.

Table 13.1 Rules for Checking Dimensional Homogeneity of Equations

1. All terms in an equation must have the same dimensions.
2. Multiplication and division must be consistent with Rule 1.
3. Dimensions are independent of magnitude;
   dx/dt is the ratio of infinitesimals but still has dimensions of Length Time^{-1}.
4. Pure numbers (e, π) have no dimensions.
5. Multiplication by a dimensionless number does not change dimensions.

Note: Adapted from Riggs (1963).

Box 13.1 gives two examples of checking for dimensional homogeneity using the brief list of rules in Table 13.1. The first example is for a commonly occurring model, the Lotka-Volterra equation for prey numbers with density-dependent loss to predators. In this equation the interaction coefficient ζ measures the kill rate per unit of predator effort. Typical units for ζ might be %prey per predator-day, as shown in Box 13.1. Most texts treat this interaction coefficient as a percentage rather than a percent per unit of predator effort. As a result, most texts report the interaction coefficient as a unitless number between 0 and 1 rather than as a scaled quantity with units of predator-hours, predator-days, and so on. Clearly, the results of using the Lotka-Volterra equation will depend on the value of the interaction coefficient, which will in turn depend on its units. The interaction coefficient must be expressed in the same time units as growth rate of the prey population, whether this is hours, days, or years.

Box 13.1 Checking Equations for Dimensional Consistency

1. Lotka-Volterra model of change in prey number in the presence of a predator:

N	= prey number	prey #
t	= time	day
r	= per capita rate of prey increase	% day^{-1}
P	= predator number	predator #
ζ	= percent prey killed per predator effort	(%prey)/(predator-day)

$$dN/dt = r\,N - \zeta\,N\,P$$
$$\#\ \mathrm{T}^{-1} = \mathrm{T}^{-1}\ \# - (\#\cdot\mathrm{T})^{-1}\ \#\ \#$$

2. Time to extinction (Lande, 1993):

T	= time to extinction	$[\mathrm{N}]^0\,[\mathrm{T}]^{+1}$
r	= intrinsic rate of increase = mean($\Delta N/N$)/time	$[\mathrm{N}]/[\mathrm{N}]\,[\mathrm{T}]^{-1}$
V_e	= var($\Delta N/N$)/time	$[\mathrm{N}]/[\mathrm{N}]\,[\mathrm{T}]^{-1}$
K	= carrying capacity/extinction threshold (Lande, pers com.)	$[\mathrm{N}]/[\mathrm{N}]\,[\mathrm{T}]^{0}$
c	$= 2\,r\,V_e^{-1} - 1$	$[\mathrm{N}]^0[\mathrm{T}]^0$

$$T = 2\ V_e^{-1}\,c^{-1}\ (\ c^{-1}\ (K^c - 1) - \ln K\ )$$
$$\mathrm{T}^1 = \mathrm{T}^1\ \mathrm{T}^0\ (\ (\mathrm{N}^0\,\mathrm{T}^0)^{-1})\ (\mathrm{N}^0\,\mathrm{T}^0)^{\mathrm{c}} - \ln(\mathrm{N}^0\mathrm{T}^0)\ )$$
$$\mathrm{T}^1 = \mathrm{T}^1\ (\ \mathrm{N}^0\,\mathrm{T}^0 - \mathrm{N}^0\,\mathrm{T}^0\ )$$

The second example in Box 13.1, like the first, comes from a list of 34 analytical models in an introductory textbook on ecology (Ricklefs and Miller, 2000). In both examples the dimensions were consistent. The example from Lande (1993) was a good deal more challenging, since it was necessary to use other equations in the same article to work out the dimensions of several symbols.

The principle of dimensional homogeneity also allows one to work out the units and biological meaning of model components that are not defined (Box 13.2). An example is Ivlev's equation for rate of prey ingestion (Ivlev, 1961), which contains a symbol ζ (zeta) that is not defined (Ivlev, 1961).

$$I = I_{\max}(1 - e^{-\zeta(N-N')}) \tag{13.3}$$

Box 13.2 Use of Dimensional Homogeneity to Work Out the Units and Biological Interpretation of an Undefined Symbol in an Equation

$$I = I_{\max}(1 - e^{-\zeta(N-N')})$$

I	≡ Ingestion, prey/hour	$[N]\,[T]^{-1}$
I_{max}	≡ Maximum ingestion, prey/hour	$[N]\,[T]^{-1}$
N	≡ Prey concentration, count/ml	$[N]\,[L]^{-3}$
N'	≡ Threshold prey concentration, count/ml	$[N]\,[L]^{-3}$
ζ	≡ ? (undefined)	

The product $\zeta \cdot (N - N')$ must have no dimensions (Table 13.1).

Consequently, ζ must have dimensions that are the inverse of $(N - N')$. The dimensions are $[L]^3[N]^{-1}$, which suggests an interpretation, that ζ is the volume swept per prey capture.

ζ	≡ Volume swept, ml/prey	$[L]^3[N]^{-1}$

Note: Ingestion equation from Ivlev (1961).

The parameter ζ appears in the exponent along with the difference between prey concentration and maximum prey concentration ($N - N'$ = prey/ml). For the entire exponent to be dimensionless and unitless, the mystery symbol ζ must have units of the inverse of prey concentration: ζ = ml/prey. Aha! Dimensional analysis shows that ζ is a volume per individual prey. Ivlev's equation concerns predation by fish, so ζ has to do with the volume searched by a fish, on average, to capture one prey item.

The principle of dimensional homogeneity is the door to interpretation of models. It permits the concepts expressed in symbolic form to be decoded by reference to interpretable units and dimensions.

ANOTHER LOOK AT SECTION 13.5

Write out dimensions below each symbol in the ingestion equation (Box 13.2), then check for dimensional homogeneity.

13.6 Calculations Based on Ideas

To an increasing degree, equations in ecology are used to make calculations about real problems. This style of quantitative ecology requires dimensionally homogeneous equations, for which each symbol has biologically or physically interpretable units. This style of quantitative biology is far easier to learn than analytic use of equations to develop theory, because reasoning is guided by calculations based on visualizable quantities. Accurate statement of ideas in mathematical form and subsequent calculations based on those ideas are a matter of practice rather than mathematical training. Little or no facility in solving equations is required. Algebra suffices and an understanding of the principles of calculus helps. One can go a long way on a little algebra combined with the ability to reason about quantities and to read and comprehend an equation.

Box 13.3 shows the use of an equation to make calculations based on a biological idea rather than to draw analytical conclusions. First, the idea: Food consumption must balance metabolic rate at the time scale of weeks to years, and hence consumption on these time scales can be calculated from metabolic rate measured in free-living organisms. The idea, expressed in functional form, is that food consumption ($\dot{M}$ = kg/day) depends on field metabolic rate ($\dot{E}$ = Watts), on assimilation efficiency (M_{assim} = % assimilated of food ingested), and on the energy density of prey ($E_{/M}$ = Joules/kg). Here is exactly the same sentence, cast in symbolic form:

$$\dot{M} = \mathrm{f}(M_{assim}, E_{/M}, \dot{E}) \tag{13.4}$$

A Watt is equal to a Joule per second so another conversion factor $k_{s/day}$ is needed to convert seconds to days to arrive at an equation for calculating daily food consumption:

$$\begin{array}{lcccccccc}
\dfrac{\text{kg}}{\text{day}} & = & \dfrac{18600\,\text{s}}{\text{day}} & \cdot & \dfrac{\text{kg ingested}}{\text{kg used}} & \cdot & \dfrac{\text{kg used}}{\text{Joule}} & \cdot & \dfrac{\text{Joules}}{\text{s}} \\
\dot{M} & = & k_{s/day} & \cdot & {M_{assim}}^{-1} & \cdot & {E_{/M}}^{-1} & \cdot & \dot{E} \\
\mathrm{M\,T^{-1}} & = & \mathrm{T^0} & \cdot & \mathrm{M^0} & \cdot & \mathrm{E^{-1}M} & \cdot & \mathrm{E\,T^{-1}}
\end{array}$$

This equation expresses the idea that food consumption depends on field metabolic rate $\dot{E}$, on energy density of prey $E_{/M}$, and on assimilation efficiency M_{assim}. It is here written with the highly specific units to show that units cancel correctly. It is also written with the dimensions below to show that the equation is dimensionally homogeneous.

Box 13.3 Food Intake $\dot{M}$ Calculated from Metabolic Rate $\dot{E}$

1. Write a dimensionally homogeneous equation:

$$\dot{M} = k_{s/day} \cdot E_{/M}{}^{-1} \cdot \dot{E}$$

2. Symbols and units:

$\dot{M}$	kg day^{-1}
$k_{s/day}$	s day^{-1}
$E_{/M}$	Joules kg^{-1}
M_{assim}	80%
$\dot{E}$	Watts = Joules s^{-1}

3. The idea in words: Food intake $\dot{M}$ is directly proportional to metabolic rate $\dot{E}$ at the time scale of weeks to lifetimes. Food intake is related to metabolic rate via the energy density of prey: ($E_{/M} = 7{\cdot}10^6$ Joule kg^{-1}) and assimilation efficiency ($M_{assim} = 80\%$).

4. Check that units on left equal those on right:
(s day^{-1})(Joule kg^{-1})$^{-1}$ (80%)$^{-1}$ (Joule s^{-1})
= s^0 Joule^0 kg^{1}day^{-1}

5. Substitute and calculate:

$$\dot{M} = \frac{86400\text{s}}{\text{day}} \cdot \left(7 \cdot 10^6 \frac{\text{Joule}}{\text{kg}}\right)^{-1} \cdot 0.80^{-1} \cdot \dot{E}$$

$$= 0.0154 \frac{\text{kg} \cdot \text{s}}{\text{Joule} \cdot \text{day}} \cdot \dot{E}$$

$\dot{E}$ = 5.5 Watt (as might be measured for a half kg rodent)
$\dot{M}$ = 0.0154 kg Watt^{-1} day^{-1} · 5.5 Watt
=> 0.085 kg day^{-1}

At the time scale of weeks to lifetimes, a rodent with a metabolic rate of 5.5 Watt has an expected intake of 0.085 kg day^{-1} of food.

Once we have arrived at a formal expression, we can make calculations based on the idea, as in Box 13.3. Table 13.2 lists a general sequence of steps for using ideas, formally expressed, to make calculations.

Table 13.2 Calculations from Equations Expressing Biological Ideas

1. Write the equation.
2. Write each symbol, with units.
3. State in words the idea expressed by the equation.
4. Make sure units on the left side equal those on the right.
5. Substitute values of parameters and variable quantities to calculate the quantity of interest from the equation.

ANOTHER LOOK AT SECTION 13.6

Rewrite Expression 1.4 for ingestion per unit area $[\dot{M}]$ instead of ingestion $\dot{M}$. Write an expression to calculate $[\dot{M}]$ from $\dot{E}$.

13.7 Writing Equations

Many people think that writing equations is an esoteric activity restricted to a few "modelers" who have great facility in mathematics. In fact, writing an equation is a readily developed skill founded on unit cancellation, dimensional homogeneity, algebra, and practice in visualizing the relation between quantities expressed as symbols. These are skills that can be gained with relatively little practice. Skill in algebraic manipulation is necessary, and some knowledge of calculus helps. However, it is remarkable how much one can express with no knowledge of calculus. Skill in working with units and dimensions is necessary. Unit cancellation not only guides the development of symbolic expressions, it is used to check whether an idea has been expressed adequately enough to permit calculations. A related skill lies in imagining the physical or biological interpretation of the operations of addition, subtraction, multiplication, division, and exponentiation in a given situation.

The next two subsections cover the principles of good notation (Section 13.7.1) and parsimony (Section 13.7.2). Section 13.7.3 outlines a generic procedure for writing ideas in symbolic form; Section 13.7.4 then demonstrates the procedure. Section 13.7.5 extends the procedure to writing equations from component equations. Mathematically correct operations, together with good notation and appropriate choice of symbols, constitute what might be called *quantitative grammar*.

13.7.1 Notation

Clear notation is as important in quantitative work as grammar and word usage are in writing. Notation and choice of symbols receive too little attention, perhaps because mathematical relations hold regardless of the symbols used. Any symbol can serve, but it does not follow that all symbols serve equally well in representing scaled quantities. Conventional symbols, such as t for time, function far more effectively than unconventional symbols. Mnemonic symbols also contribute to clarity, although this desirable quality must be balanced against the advantages of conventional notation. Another characteristic of good notation is focus on the primary quantity. A symbol such as M_{egg} or M_e will prove easier to read within an equation than a symbol such as E for egg mass. The latter symbol draws attention to a secondary feature rather than to the measured quantity, which is a mass.

Well-chosen diacritical marks contribute to clarity. An example of a diacritical mark is an overbar to represent an average of the quantity under the bar. Diacritical marks often help in representing quantities that result from an operation that's frequently repeated, such as taking an average. Clarification occurs when the diacritical mark simplifies the notation. An example is placing a dot over a symbol to represent the

Table 13.3 Symbols for Equality

≡	Defined as or equal to by definition
~	Approximately equal to
:=	Thought to be equal to or conditionally equal to
≅	Scales as
$\stackrel{\text{O}}{=}$	True by observation
⇒	Calculated as
⇐	Calculated from

time rate of change in a quantity. Another useful diacritical mark is placing a "hat" over a symbol to signify an estimate of a parameter. Thus, an estimate of the true value of egg mass M_{egg} is represented as $\hat{M}_{egg}$. In statistics, an old-fashioned convention is to use Greek symbols for a parameter, Roman symbols for its estimate. The convention places a greater burden on memory than the use of a hat. Happily, the convention is doomed: There are relatively few pairs such as ρ and r.

Well-chosen changes in typeface also contribute to clarity. The most common convention is the use of boldface type to distinguish a vector (which has direction) from a scalar quantity (which has no direction). The major difficulty in using bold typeface to distinguish symbols is that the distinctions are lost when the symbols are written by hand. A tilde (~) is conventionally written by hand beneath a symbol to represent a vector such as $\boldsymbol{i}$, the unit vector in the x direction. Italics help set off symbols from surrounding text, making it easier to follow a text description of a symbol.

Clear expression results from distinguishing the several types of equalities that can occur. The simple equality sign = is pressed into service when any of several forms of equality are meant. Table 13.3 lists several different signs, each of which communicates more than the simple equality sign.

Moving from good notation for symbols to good notation for equations, an important characteristic of a well-written equation is consistent usage of each symbol. A symbol must always stand for the same quantity. If it does not, confusion and erroneous calculation inevitably result. The device that contributes most to consistent notation is a listing of symbols, or a dictionary, showing each symbol with its name, such as the list recommended in Chapter 3.3, for construction by the reader. A dictionary of symbols will not guarantee consistent notation, but it does help greatly in working toward this goal.

A device that contributes to clear writing of equations is functional notation to represent a computational recipe. If an operation occurs again and again, functional notation increases clarity by standing for the set of operations, leaving aside the details. This is like reading a word as a unit with meaning, without having to read each letter. For example, in Chapter 10 the symbol var(Q), which is easier to grasp, was used instead of the harder to grasp computational formula for taking the variance of a quantity.

Another device that contributes to clarity is the use of italics, which has become nearly universal in setting mathematical expressions into type. It is the default option in equation editors in word processing packages, and it is the house style in journals in which mathematical expressions appear. Italics are particularly effective in setting off a symbol t within a line of text, rendering the line easier to read than a Roman symbol t in the line. Subscripts are usually set in italics for consistency, whereas functional

expressions var(t) remain nonitalic. Dimensional expressions [M] remain nonitalic along with their brackets. This leaves an italic bracket $[M] \equiv MV^{-1}$ free to represent a concentration of the quantity (in italics) within the bracket.

Another example of a quantity that can be written with functional notation is the formula for the number of pairs of N objects. The formula for the number of pairs is:

$$\text{duo}(N) \equiv \sum \frac{(N)(N-1)}{2} \tag{13.5}$$

The symbol duo(N) has the advantage of being directly tied to the concept and hence is more easily recognized than the complex term on the right side of the equation. Once defined, the shorter symbol duo(N) can stand in for the more complex form on the right side of Expression 13.5. Ecologically interpretable quantities such as mean crowding $M^*(N)$ or potential contact $PC(\boldsymbol{i})$, deserve a simple and easily remembered symbol of their own in place of the full expression for calculating that quantity.

A good notational system brings out the relationship between quantities, whether expected or observed. An example is a system with the following elements: a letter for a quantity, a simple dot over the symbol for the time rate of change, and brackets around a quantity to represent the concentration per unit area or volume, as in chemistry. This system can be applied to any quantity: number of species s, number of individuals N, biomass of individuals M, biomass of populations B, or numbers of a particular gene G in a population. The system reduces the number of different letter symbols that must be remembered. The system also brings out clearly the relation of quantities, facilitating insight and biological reasoning about measurable quantities.

To demonstrate this idea, systematic notation is applied to a quantity that is often of interest in applied ecology: the total biomass of a population. The static quantity is population biomass (B = grams), which is the product of population numbers (N = individuals) and biomass per individual (M = grams/individual):

$$B = N \cdot M \tag{13.6}$$

The time rate of change in biomass is $\dot{B}$ with units of grams per unit time ($\dot{B} = \text{g}\cdot\text{yr}^{-1}$). This is the biomass production, which traditionally has been assigned a new symbol P. The percent rate of change in biomass is $\dot{B}/B$, also called the P/B ratio, with units of % yr^{-1}. The symbols B, $\dot{B}$, and $\dot{B}/B$ display the relation among the quantities.

What is the relation of the percent production $\dot{B}/B$ to population growth, either as a simple rate $\dot{N}$ or as a per capita rate $\dot{N}/N$? The notation makes this evident, beginning with a mathematical equivalence for a per capita rate:

$$\frac{\dot{B}}{B} = \frac{d\ \ln(B)}{d\ t} \tag{13.7}$$

Now substitute $N \cdot M$ for B, then replace the derivatives with equivalent ratio symbols to make the relations among quantities more evident:

$$\begin{aligned} \frac{d\ \ln(B)}{dt} &= \frac{d\ln(N)}{dt} + \frac{d\ln(M)}{dt} \\ \frac{\dot{B}}{B} &= \frac{\dot{N}}{N} + \frac{\dot{M}}{M} \end{aligned} \tag{13.8}$$

The percent rate of production $\dot{B}/B$ is simply the sum of two familiar quantities. The first is the per capita rate of change in population size $\dot{N}/N$, a familiar quantity from population biology. The second is $\dot{M}/M$, the growth rate of individuals as a percentage of body mass. Intuitively one might not expect a quantity such as the production-to-biomass ratio to be the sum of two widely studied quantities. The per capita rate of change in number $\dot{N}/N$ and the individual growth rate as a percent $\dot{M}/M$ are both related to body mass M by allometric scaling. Consequently, the percent change in biomass $\dot{B}/B$ should also be related to body size, which proves to be the case (Banse and Mosher, 1981).

Percent production $\dot{B}/B$ can be investigated for fixed areas, such as a lake. The production per unit area of lake is $[B] = B/A$. The italicized square brackets here are used to denote the density, rather than concentration per unit volume. The percent rate of change in $[B]$ is $[\dot{B}]/[B]$. As before, this is mathematically equivalent to the derivative of the logarithm of the quantity:

$$\frac{[\dot{B}]}{[B]} = \frac{d\ \ln([B])}{d\ t} \tag{13.9}$$

The production per unit area, as a percentage, is the difference of two percent rates of change:

$$\begin{aligned} \frac{d\ln([B])}{dt} &= \frac{d\ln(B)}{dt} - \frac{d\ln(A)}{dt} \\ \frac{[\dot{B}]}{[B]} &= \frac{\dot{B}}{B} - \frac{\dot{A}}{A} \end{aligned} \tag{13.10}$$

This was obtained by substituting B/A for $[B]$. The derivatives were again replaced with simpler symbols to bring out the relation between quantities. The notation makes it clear that the production density $[\dot{B}]/B$ depends in part on population production $\dot{B}/B$ and in part on rate of change in area occupied $\dot{A}/A$. Equation 13.8 then leads to further analysis of percent production per unit area $[\dot{B}]/B$ into components of demographic change $\dot{N}/N$, growth rates of individuals $\dot{M}/M$, and change in area occupied $\dot{A}/A$:

$$\begin{array}{ccccccc} \text{Percent Production} & & \text{Demography} & & \text{Somatic Growth} & & \text{Kinematics} \\ \dfrac{[\dot{B}]}{[B]} & = & \dfrac{\dot{N}}{N} & + & \dfrac{\dot{M}}{M} & - & \dfrac{\dot{A}}{A} \end{array} \tag{13.11}$$

The production density $[\dot{B}]/[B]$ is the sum of three ratios. The first summarizes demographic processes (births and deaths), the second summarizes somatic growth, and the third summarizes kinematic processes (movements and distribution) as a divergence in two dimensions (see Boxes 8.7 and 8.8).

This analysis demonstrates how a consistent system of notation facilitates quantitative reasoning. Display of the quantities in this format immediately leads to a key question: What is the relative contribution of growth, demographic, and kinematic processes to the percent rate of production? The answer to this question will depend on the ecological characteristics of populations. Somatic growth will be important in small-bodied species, whereas demographics will likely be more important than somatic growth in larger,

longer-lived organisms. Kinematics will be less important in sessile populations than in mobile populations that coalesce during certain life stages, such as the coalescence of pelagic fish into spawning areas. The analytic strategy is to break a quantity down into major components, then ask which are important, before moving on to the processes responsible for each component of production. Systematic symbolic notation opens the door to this analytic strategy, made possible by displaying the relation of the components.

ANOTHER LOOK AT SECTION 13.7.2

Write out Expression 13.11, then beneath it write out dimensions. Is the expression dimensionally homogeneous?

13.7.2 Parsimony

An important consideration in writing an equation is *parsimony*, which is expressing an idea as simply as possible. This usually means writing the equation with as few quantities as are necessary to the purpose. The procedure is to start with the simplest possible expression, then add detail if necessary. Here is an anecdote that illustrates the idea.

While on a postdoctoral fellowship in California, I took up body surfing because it looked like fun and the competition for space in the water was not vicious, as it was out with the board surfers. After I had learned how and where to catch little waves, I went out into some larger waves coming in one day at Little Corona Beach. I caught several rides over 50 meters and soon forgot about anything else except watching for the next wave. Then, having misjudged placement on a wave, I found myself being carried up to the top of a real "dumper" rather than sliding down the front of a wave with better form.

Lacking the ability to tuck out of the wave, I was carried up and then began falling head first almost straight down over the breaking crest of the wave, which by now was vertical in front. Rather than allowing myself to be driven head first into the sand, I rolled forward into a tight ball, taking a deep breath on the way down. The wave broke over me, tumbling me around at the bottom of a mass of churning water and sand. The water was too deep for me to stand up, but I assumed that buoyancy would carry me to the surface, so I waited.

Let's call this Model I. The response variable is time to get to the sea surface. The explanatory variable is buoyancy. The estimate based on this idea was that the time to float to the surface would be a few seconds.

After more than a few seconds it became evident that buoyancy was not going to carry me upward through the churning water before I ran out of breath. Model I was not adequate. I had to try something else, quickly.

I knew I had the leg strength to push through the churning water, so Model II was to add an additional force to buoyancy by pushing off the bottom. Model II was still not sufficient because, as I was tumbled around at the bottom, I could feel a lot of sand churning around and so opening my eyes was the short route to more pain than gain of information on which way was up. A blind push might not work. As I tumbled around down there, I stretched my hands out slightly, feeling for the bottom. Then at the point in time when both hands and both feet touched the bottom, I imagined a flat surface against which I pushed as hard as possible.

This is Model III, that force be applied perpendicularly to four determinations of the sea floor rather than perpendicularly toward the sunlit sea surface. This model worked. At about the time I had run out of air I burst through to the surface, which was still churning and frothing in the wake of the wave. Model III was more complex than Models I and II but no more complex than it had to be to solve the problem.

The same sequence of adding detail as needed applies to writing equations, although usually without the degree of urgency in arriving at an adequate solution to the problem of regaining the sea surface.

13.7.3 Sequential Caricatures

The equation used in Box 13.3 was presented without citation to some published work. In fact it has none; the equation was developed by reasoning about quantities. It was the result of sequential development of symbolic expressions from biological ideas. It will serve as an example of such an approach.

The first step is to separate the response from explanatory variables. The response variable is of interest for any of a variety of reasons, whether in the context of conservation, resource management, or better understanding of the natural world. In Box 13.3 the response variable, with units and a symbol, is food consumption, $\dot{M}$ = kg/day. The next step is to state one or more explanatory variables with units and symbols. Food intake is nearly impossible to measure directly in a marine vertebrate such as a penguin. Intake depends on an explanatory variable, metabolic rate, that can be measured with stable isotopes, as in the equation from Kooyman et al. (1982), which led off this chapter (Equation 13.1). Metabolic rate in units of energy per unit time is $\dot{E}$ = Watt ≡ 1 Joule/second. The next step is to state a simple functional relation between response variable (on the right) and explanatory variable(s) on the left side of the equality sign.

$$\dot{M} = \mathrm{f}(\dot{E}) \tag{13.12}$$

This is read as "intake $\dot{M}$ is a function of metabolic rate $\dot{E}$."

Expression of the idea in this functional form marks the halfway point from verbal to formal expression as an equation. This functional expression is not yet an equation because the quantity on the left side, food consumption, cannot be computed from the quantity on the right, metabolic rate. The left side (kg/day) does not have the same units as the right (Watts). The missing piece of biology is the energy density of prey, that is, the number of Joules per kg of prey. For most types of animal prey, except for something as watery as gelatinous zooplankton, energy density is around $7{\cdot}10^6$ Joules/kg.

The formal expression relating food consumption to metabolic rate and energy density of prey ($E_{/M}$ = Joules/kg) is:

$$\dot{M} = \dot{E} \cdot {E_{/M}}^{-1} \tag{13.13}$$

Metabolic rate has been divided by energy density expressed in Joules per kilogram so that the units cancel properly: Joules in the numerator of metabolic rate cancel Joules in the denominator of $E_{/M}{}^{-1}$.

This expression still falls short because intake $\dot{M}$ has been defined per day, whereas the time units of metabolic rate are second^{-1}. A rigid conversion factor $k_{s/day}$ is required

to rescale field metabolic rate $\dot{E}$ from Joules · second^{-1} to Joules · day^{-1}. The expression is revised by adding the conversion factor:

$$\dot{M} = k_{s/day} \cdot \dot{E} \cdot E_{/M}{}^{-1} \tag{13.14}$$

Another check of the units should show that the units on the right will cancel out to yield the units on the left. I leave this to the reader.

Both $k_{s/day}$ and $E_{/M}{}^{-1}$ have known values, which can be substituted, leading to a new form of the same equation:

$$\dot{M} = (0.0123 \text{ kg Watt}^{-1} \text{ day}^{-1}) \cdot \dot{E} \tag{13.15}$$

This is Model I, which is dimensionally consistent and hence can be used for calculating food intake from field metabolic rate. If the calculations are checked against detailed measurements, such as might be obtained from captive penguins in a zoo, the observed intake will exceed the calculated value. Much of this bias will be due to the inefficiency of digestion, which wastes something like 20% of the calories ingested. The energy actually burned is less than the energy ingested. The accuracy of the equation can be improved by adding more biology in the form of an *assimilation rate*, or ratio of the mass of the food ingested to the mass of the food assimilated. This typically has a value of $M_{assism} = 80\%$, or 0.8 kg assimilated per kg ingested.

The next step is to write the revised model (Model II):

$$\dot{M} = (0.0123 \text{ kg Watt}^{-1} \text{ day}^{-1}) \cdot \dot{E} \cdot M_{assim} \tag{13.16}$$

This is analogous to Model II in the surfing problem in the previous section: "Push off the bottom." As with Model II in the surfing problem, the improvement is no good if applied in the wrong direction. Assimilation is a dimensionless ratio, so there are no units to guide its application in the right direction. The solution is to reason about the quantities involved. Consumption M must exceed metabolic requirements (on the right side), so the right side must be multiplied by a factor greater than 1. This is accomplished by dividing the right side by the assimilation rate. Dividing by 80% boosts the metabolic requirement to the larger value on the left side. So, the equation has to be:

$$\dot{M} = (0.0123 \text{ kg Watt}^{-1} \text{ day}^{-1}) \cdot \dot{E} \cdot M_{assim}{}^{-1} \tag{13.17}$$

This is analogous to Model III in the surfing problem: "Apply force in the correct direction, estimated as perpendicular to the bottom."

The surfing example in the previous section and the food consumption discussion in this section both illustrate the concept of constructing a series of caricatures to address a problem. They were tackled in similar ways, adding detail only as necessary.

ANOTHER LOOK AT SECTION 13.7.3

Metabolic rate $\dot{E}$ is typically measured under standard conditions: at rest, in thermoneutrality, and not actively absorbing food. Write an expression for daily intake by a free-living organism not living at rest in thermoneutrality and not absorbing food.

13.7.4 General Procedure

Several components of good practice have now been covered: good notation, parsimony, and sequential development, including checks on dimensional homogeneity and unit conversions. Table 13.4 presents a generic recipe based on these components. As with any generic recipe, the number of variations is large. To illustrate the recipe, a model for field metabolic rate will be developed.

Many ecologically important quantities are expensive or even impossible to measure directly. Equations based on biological concepts permit calculation of these quantities. Field metabolic rate $\dot{E}$, for example, is expensive to measure. Happily, the quantity $\dot{E}$ is closely related to a much more easily measured quantity, body size (M = kg). This completes Steps 1–4 in Table 13.4. For Step 5, the idea that metabolic rate depends on body size is expressed in functional form:

$$\dot{E} = f(M) \tag{13.18}$$

As previously, the response variable (which is to be calculated) has been placed on the left side of the equation. The explanatory variables, which are to be used to make the calculation, have been placed on the right side. The units on the left do not match those on the right; more biology needs to be added to describe exactly how field metabolic rate ($\dot{E}$ = kcal/day) is related to body mass (M = kg). The simplest idea is that field metabolic rate is directly related to body mass:

$$\frac{\dot{E}_{elephant}}{\dot{E}_{mouse}} = \left(\frac{M_{elephant}}{M_{mouse}}\right)^{1} \tag{13.19a}$$

As shown in Chapter 2, this is equivalent to:

$$\dot{E} = k \cdot M \tag{13.19b}$$

A few measurements on animals of different sizes will quickly reveal problems with this 1:1 scaling. The scaling of energy to mass, $k = \dot{E}/M$, is not constant. This scaling ratio increases with decreasing mass because small organisms live more intensely than larger organisms and have higher metabolic rates per unit mass.

Table 13.4 Generic Recipe for Writing an Equation

1. State the response variable.
2. Define this quantity in words, and assign units to a symbol.
3. State the explanatory variable or variables.
4. Define each in words, and assign symbols and units.
5. Write the response variable as a function of the explanatory quantities:
   Response = f(Explanatory, Explanatory, ...)
6. Write an equation by reasoning about the quantities. Empirical description of the form of the relation, such as from exploratory data analysis, is also useful.
7. Check the equation to make sure that units cancel correctly.
8. Make a calculation and check against an independent measurement, if possible.
9. Revise the equation to include more processes as necessary.
10. Check units after each revision, before making calculations.

An improvement on Model 13.18 is that metabolic rate varies with surface area, whereas mass varies with volume. According to this notion, energy dissipation is directly proportional to area, which in turn scales as volume$^{2/3}$ and hence to mass$^{2/3}$ in organisms with a fixed density (mass-to-volume ratio). Step 6 (Table 13.4) is completed by writing the equation obtained by reasoning about the relation of metabolic rate to body mass:

$$\dot{E} = k \cdot A^1 = k \cdot V^{2/3} = k \cdot M^{2/3} \tag{13.20}$$

This turns out to be nearly correct, but not exactly. A still more accurate description is that the metabolic rates scale according to some other exponent β that on average exceeds 2/3. The scaling function that relates metabolic rate to body mass is:

$$\dot{E} = k_{Watt/kg} \cdot M^{\beta} \tag{13.21a}$$

$$\text{Watt} = (\text{Watt} \cdot \text{kg}^{-\beta}) \cdot \text{kg}^{\beta} \tag{13.21b}$$

Units are checked (Step 7, Table 13.4) by writing them beneath each symbol. For units to cancel, the symbol $K_{watt/kg}$ must have the complex units shown in Equation 13.21b.

One of the first estimates of β and $K_{Watt/kg}$ relative to field metabolic rate $\dot{E}$ was obtained by King (1974), who measured the amount of time that individual animals spent in assorted activities, then combined this with the energetic cost of each activity to obtain a daily energy budget in bird and rodent species of different body sizes.

King's (1974) parameter estimates were $\hat{\beta} = 0.6687$ and $\hat{k} = 179.8$ kcal $\cdot$ kg$^{-0.6687} \cdot$ day^{-1}. Applying the rules for rigid rescaling (Table 5.3) yields a new scaling factor $\hat{k}_{Watt/kg} = 8.74 \cdot \text{Watt} \cdot \text{kg}^{-0.669}$ for rodents. Because these estimates of $\hat{k}_{Watt/kg}$ and are completely empirical, they apply only to rodents in the 0.009 to 0.61 kg range. A calculation was then made (Step 5, Table 13.4). Box 13.4 shows calculations from King's equation.

Box 13.4 Metabolic Rate Calculated from Body Mass Following Steps in Table 13.2

1. Write the equation:
   $\dot{E} = k_{kcal/(kg-day)} \cdot M^{\beta}$ (King, 1974)
2. $\dot{E}$ = kcal·day
   $k = 179.8\,\text{kcal} \cdot \text{kg}^{-.6687} \cdot \text{day}^{-1}$
   M = kg
   $\beta = 0.6687$
3. The idea in words: Field metabolic rate scales as body mass in rodents and hence can be calculated by allometric rescaling according to body mass.
4. Check units:
   $(\text{kcal} \cdot \text{kg}^{-.6687} \cdot \text{day}^{-1})(\text{kg}^{0.6687}) = \text{kcal}^1 \cdot \text{kg}^0 \cdot \text{day}^{-1}$
5. Substitute and calculate:
   $\dot{E} = (179.8\,\text{kcal} \cdot \text{kg}^{-.6687} \cdot \text{day}^{-1})(0.5\,\text{kg})^{0.6687}$
   $\Rightarrow 113\,\text{kcal day}^{-1}$

At the time of King's work, there were no independent measurements against which to check these calculations (Step 5, Table 13.4). King's scaling factor of 8.74 Watt $kg^{-0.6687}$, which was obtained by quantitative reasoning to combine several quantities, was later confirmed by direct measurement nearly a decade later. The field metabolic rate of rodents, as measured with doubly labeled water (Garland, 1983), is $\hat{k}_{\text{Watt/kg}} = 9.28$ Watt $kg^{-0.66}$, almost exactly the same as the calculated scaling based on quantitative reasoning (King, 1974). Revision (Step 9, Table 13.4) was not needed.

ANOTHER LOOK AT SECTION 13.7.4

Which is more informative?

- A test of the hypothesis that $k_{Watt/kg} = 0$ (a null hypothesis)
- A test of the hypothesis that $k_{Watt/kg} = 8.74$ Watt $kg^{-0.6687}$

Why?

13.7.5 Combining Equations

Using quantitative reasoning, one can combine several equations to make calculations that can be tested against data (Peters, 1983; Calder, 1984; Pennycuick, 1992). For example, Calder (1984, p. 305) used reasoning about quantities to obtain a relation between foraging bouts (T = days) and body size (M = kg) based on food intake as a function of metabolic rate. The relation that Calder obtained is that:

$$T = 3.04\, M^{-0.26} \tag{13.22}$$

This relation is based on several empirical relations combined with reasoning about how intake must vary with body size. The relation was not obtained by simply regressing measurements of intake frequency against body size. The two parameters (an exponent of 0.26 and a rigid scaling factor of 3.04 days $\cdot$ $kg^{-0.26}$) represent highly specific expectations about that relation, not nominal scale or yes/no expectations about whether or not the relation is present.

Box 13.5 shows a similar example in which a calculation about average food intake by a half kg rodent is obtained by reasoning about quantities. The idea that food consumption is directly proportional to respiration rate (Box 13.3) and the idea that respiration scales allometrically according to body mass (Box 13.4) can be combined into one equation. Substitution of parameter values produces an equation that can be used to calculate food consumption by rodents, as shown in Box 13.5. The calculations follow the recipe in Table 13.2.

The calculations show that at time scale of weeks to lifetimes, a 0.5 kg rodent is expected to consume 0.085 kg food per day, or 17% day^{-1} relative to its own body mass. This calculation is based on a set of ideas about the relation of intake to metabolic rate and body mass. This calculation may be adequate in some cases but not others. As with any model (Section 13.7.2), more detail can be added if the model is not adequate to the purpose at hand.

Box 13.5 Food Intake Calculated from Body Size

Combine equations in Boxes 13.3 and 13.4 to calculate food intake from body mass.

1. The equation:

$$\dot{M} = k_{s/day} \cdot E_{/M}{}^{-1} \cdot M_{assim}{}^{-1} \cdot k_{Watt/kg} \cdot M^{\beta}$$

2. Each symbol, with units:

$\dot{M}$	kg day^{-1}
$k_{s/day}$	s day^{-1}
$E_{/M}$	Joules kg^{-1}
M_{assim}	80%
$k_{Watt/kg}$	Watt kg$^{-.6687}$
M^{β}	kg$^{.6687}$

3. In words: At the time scale of weeks to lifetimes, food intake depends on body size (M^{β}), the scaling of metabolic rate to body size ($k_{Watt/kg}$), assimilation efficiency (M_{assim}), and the energy density of prey ($E_{/M}$).
4. (Joules kg^{-1})$^{-1}$(Joule Watt^{-1} day^{-1})(Watt kg$^{-.6687}$)(kg)$^{.6687}$
   = Joule^{0} kg^{1} Watt^{0} day^{-1} kg^{0}
   = kg day^{-1}
5. Substitute parameter estimates. King's (1974) estimate is:

$$k_{Watt/kg} = 8.74 \text{ Watt kg}^{-.6687} \quad \text{(King, 1974)}$$
$$\beta = 0.6687 \quad \text{(King, 1974)}$$
$$k_{s/day} \cdot E_{/M}{}^{-1} \cdot M_{assim}{}^{-1} = 0.0154 \text{ kg day}^{-1}\text{Watt}^{-1} \quad \text{(Box 13.3)}$$

   The resulting equation is:

$$\dot{M} = 0.0154 \text{ kg day}^{-1} \text{ Watt}^{-1} \cdot 8.74 \text{ Watt kg}^{-.6687} \cdot M^{0.6687}$$

   Calculate $\dot{M}$ for a rodent of M = 0.5 kg:

$$\dot{M} = 0.0154 \text{ kg day}^{-1} \text{ Watt}^{-1} \cdot 8.74 \text{ Watt kg}^{-.6687} (0.5 \cdot \text{kg})^{0.6687}$$
$$\Rightarrow 0.085 \text{ kg day}^{-1}$$

ANOTHER LOOK AT SECTION 13.7.5

If the number of bird species N_{sp} on an island scales with the number of birds $N_{sp} = k_N \cdot N_{bird}{}^{0.3}$ and N_{bird} scales directly with island area $N = k_A \cdot A^1$, how does N_{sp} scale with the island area?

13.8 Empirical and Theoretical Scaling Functions

Dimensional analysis is a useful tool for checking equations (Section 13.5). It is also an effective tool for developing scaling functions. Box 2.4 showed the first ecological use of a measurement relation to develop an *empirical scaling function* relating species to area: $S = cA^z$. This function is empirical because it was developed by fitting data to a line (on a log-log plot). Beginning with a measurement relation that equates one dimensionless

ratio to another, Arrhenius (1921) estimated the exponent z of the measurement relation from data, then estimated the scaling parameter c of the scaling function. Box 6.1 describes this analytic sequence in general form, from measurement relation to scaling function. The machinery of dimensional analysis is a first step toward development of more theoretically based scaling functions. Boxes 6.3–6.5 demonstrated the machinery of obtaining dimensionless ratios. Box 6.6 demonstrated the derivation of scaling functions from dimensionless ratios. Box 6.8 combined biological reasoning with the logic of dimensions to develop Π ratios for the fish catch from lakes. A plot of data (Figure 6.2) showed that the Π ratio for lake shape was independent of the Π ratio for water clarity. Two more plots showed that the fish catch (per lake) was a function of lake shape but not water clarity (Figures 6.3 and 6.4). This suggests that fish catch, as an empirical function, is best described by lake shape, ignoring water clarity. Box 13.6 shows the development of a scaling function guided by dimensional analysis in Box 6.8.

Box 13.6 Dimensionless Ratios to Develop an Empirical Scaling Function

The functional expression for the measured variables:

$\dot{M} = \mathrm{f}(A, TDS, z)$
$\dot{M}$ = fish catch (kg yr^{-1})
A = lake area (ha)
TDS = total dissolved solids (ppm)
z = lake depth (meters)

Two independent Π ratios from three variables and one dimension (Box 6.7). Reduction by inspection (Box 6.8):

$\Pi_{\text{shape}} = A^{1/2}/z \qquad \Pi_{\text{organics}} = TDS \qquad \dot{M} = \mathrm{f}(\Pi_{\text{shape}}, \Pi_{\text{TDS}})$

Plots of data (Figures 6.2 and 6.3) show:

$\Pi_{\text{shape}} = (\Pi_{\text{organics}})^0$ No relation between two Π ratios
$\dot{M} = \mathrm{f}(\Pi_{\text{shape}})$ Fish catch related to Π_{shape} but not Π_{organics}

Empirical scaling relation for a power law relation:

$\ln(\dot{M}) = \kappa + \zeta \cdot \mathrm{In}(\Pi_{\text{shape}})$
$\hat{k} = 6.754 \qquad \hat{\zeta} = 2.0154$ Estimates from regression.
$\dot{M} = \exp(\hat{\kappa}) \cdot (\Pi_{\text{shape}})^{\hat{\zeta}} = 858 \cdot \Pi_{\text{shape}}^{2.0154}$

Dimensional analysis (Box 6.8) suggested that lake shape is the primary influence on fish production and catch. From this insight we can take the next step toward a theoretically based scaling function. If lake shape is the key factor, a simple model is that fish production depends on the flux of some limiting material (such as a nutrient) through the water body (Schneider and Haedrich, 1989). Box 13.7 demonstrates the use of Rayleigh's method of dimensional analysis (Box 6.4) to work out the theoretical scaling function if fish catch depends on a fixed flux of material through the volume, taking into account conversion efficiency of that material to fish. The equation developed in Box 13.7 is a *theoretical scaling function* because it has at least one theoretically derived parameter ($\delta = 2/3$).

The regression estimate of the scaling exponent δ was close to the theoretical value (Box 13.7). The scaling function in Box 13.7 was developed from analysis of Ryder's

data for Canadian Shield lakes (Ryder, 1965), so it was checked against a new set of data, fish catch from lakes and reservoirs further south, in Missouri and Iowa (Jones and Hoyer, 1982). The regression estimates of the exponent were 0.60 (Great Lakes data) and 0.77 (Mississippi valley data). These estimates were statistically indistinguishable from the theoretical value of 2/3 (Schneider and Haedrich, 1989). Box 13.7 shows the use of a scaling relation to estimate a scaling parameter with the dimensionally correct units of $(km^3)^{2/3} = km^2$. Figure 13.3 shows both sets of data relative to their theoretically derived scaling functions.

Box 13.7 Dimensionless Ratios to Develop a Semitheoretical Scaling Function

The functional expression for the measured variables:

$\dot{M} = f(A, TDS, z)$

$\dot{M}$ = fish catch (kg yr^{-1})

A = lake area (ha)

TDS = total dissolved solids (ppm)

z = lake depth (meters)

Theoretically based scaling relation $\dot{M}^{\alpha} = c^{\beta}[\dot{Q}]^{\gamma} V^{\delta}$

c = conversion efficiency (kg harvested/kg limiting material Q)

$[\dot{Q}]$ = flux of limiting material Q through lake (kg km^{-2} yr^{-1})

V = lake volume (km^3)

Exponents derived from Rayleigh's method:

	$\dot{M}^{\alpha}$	=	c^{β}		$[\dot{Q}]^{\gamma}$		V^{δ}	
[M]	α	=	0	+	1γ	+	0	$\alpha = \gamma$
[L]	0	=	0	−	2γ	+	3δ	$\delta = 2/3\gamma$
[T]	α	=	0	+	γ	+	0	

Setting the exponent of the response variable M to $\alpha = 1$, we have:

$\dot{M} = c^{\beta}[\dot{Q}]\, V^{2/3}$

$\dot{M} = 6.9 \cdot 10^4 V^{0.60}$ by regression, data from Ryder (1965)

$\dot{M} = 4 \cdot 10^5 V^{0.77}$ by regression, data from Jones and Hoyer (1982)

The estimates of $c^{\beta}[\dot{Q}]$ are $6.9 \cdot 10^4$ kg $yr^{-1}(km^3)^{0.60}$ and $4.5 \cdot 10^5$ kg $yr^{-1}(km^3)^{0.77}$. To obtain an estimate with the correct units $(km^3)^{2/3}$, we use a Euclidean scaling relation (Equation 2.2a) where the basis for scaling is the geometric mean for fish catch and for volume.

$$\frac{\dot{M}}{\text{geom mean}(\dot{M})} = \left(\frac{V}{\text{geom mean}(V)}\right)^{2/3}$$

$c^{\beta}[\dot{Q}] = (\text{geom mean}(\dot{M}))(\text{geom mean}(V))^{-2/3}$

$c^{\beta}[\dot{Q}] = 5.45 \cdot 10^3$ kg $yr^{-1}(3.83\ km^3)^{-2/3} = 2.22 \cdot 10^3$ kg yr^{-1} km^{-2}

$c^{\beta}[\dot{Q}] = 5.05 \cdot 10^5$ kg $yr^{-1}(28158\ km^3)^{-2/3} = 5.45 \cdot 10^2$ kg yr^{-1} km^{-2}

$\dot{M} = 545\ V^{2/3}$ Canadian shield lakes, Ryder (1965)

$\dot{M} = 2224\ V^{2/3}$ Mississippi valley, Jones and Hoyer (1982)

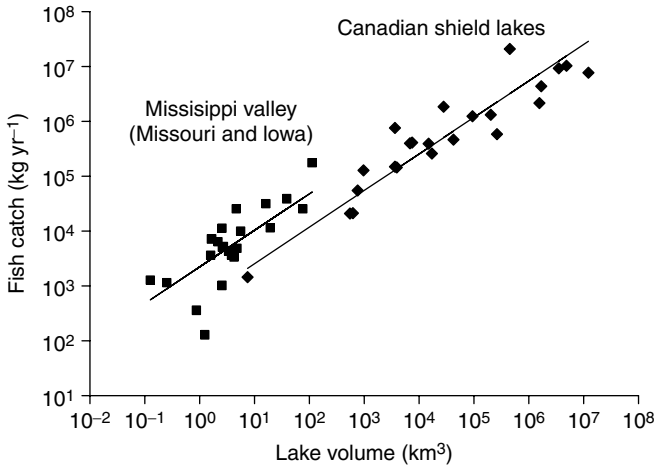


FIGURE 13.3 Fish Catch Relative to Lake Volume. Triangles show data from Canadian shield lakes (Ryder 1965). Circles show data from lakes further south (Jones and Hoyer 1982). Lines drawn through the geometric mean of each data set show theoretical scaling, with exponent of 2/3. Redrawn from Schneider and Haedrich (1989).

One bonus of the use of dimensionally correct parameter estimates is that we can compare them directly. The two estimates of $c^{\beta}[\dot{Q}]$, via regression, have different units (kg $yr^{-1}(km^3)^{0.60}$ versus kg $yr^{-1}(km^3)^{0.77}$). Consequently, we cannot compare them by taking a difference. If we compare the regression estimates as a ratio, we obtain a value with units, and hence that depends on lake volume:

$$R_{Miss/Shield} \equiv c^{\beta}{}_{Miss}[\dot{Q}_{Miss}] / c^{\beta}{}_{Shield}[\dot{Q}_{Shield}] \tag{13.23}$$

$$\begin{aligned} R_{Miss/Shield} &= (4.5 \cdot 10^5 \text{ kg yr}^{-1}(\text{km}^3)^{0.77})/(6.9 \cdot 10^4 \text{ kg yr}^{-1}(\text{km}^3)^{0.60}) \\ &= 6.5 \text{ (km}^3)^{0.17} \end{aligned} \tag{13.24}$$

If we compare the theoretically derived estimates, we obtain a dimensionless ratio, with no dependence on units:

$$R_{Miss/Shield} = (2224 \text{ kg yr}^{-1}\text{km}^{-2})/(545 \text{ kg yr}^{-1}\text{km}^{-2}) = 4 \tag{13.25}$$

So, we find that the Mississippi valley lakes are four times as efficient as the Canadian Shield lakes at yielding fish.

Theoretically based scaling functions obtained by classical dimensional analysis (Box 13.7, Figure 13.3) represent a substantial improvement over completely empirical scalings produced by regression. A well-verified theoretical scaling can be used with more confidence than an empirical scaling. The caveat is that classical dimensional analysis, which yields exponents composed of integer and integer ratios, will fail to describe the measurement or scaling relations of variable quantities subject to episodic dynamics in physical systems (Wilson, 1971; Barenblatt, 1996; Stanley et al., 1996) and biological systems (Stanley, 1996; Schneider, 2001b). In these cases (see Chapter 16), an extension of classical dimensional methods will be needed (Barenblatt, 1996) to arrive at the scaling exponent.

13.9 Dimensional Analysis of Conservation Laws

Conservation laws in conjunction with dimensional methods are an effective way of reasoning about a quantity. For any quantity of interest, one can usually write an equation called a *conservation law*, which lists source and sink terms. For example, in population biology, the quantity of interest is the number of organisms N. The rate of change, as a percentage, is:

$$\dot{N}(t)/N \equiv \frac{1}{N}\frac{dN}{dt}$$

For the entire population, the conservation law (Box 8.5) is:

$$\frac{\dot{N}(t)}{N} = \frac{\dot{B}(t)}{N} - \frac{\dot{D}(t)}{N} \tag{13.26a}$$

This equation says that the rate of change, as a percentage, is due to the proportional increase via recruitment $N^{-1}\dot{B}(t)$ and death $N^{-1}\dot{D}(t)$. The net change is conventionally represented by the symbol r, the per capita rate of change:

$$\frac{\dot{N}(t)}{N} = r \tag{13.26b}$$

At spatial scales smaller than that of the entire population, movements become important sources of change in number. If we divide the area occupied by a population into spatial units, the conservation law for each unit includes lateral fluxes running east-west$y$ $[N]\dot{y}$ or north-south $[N]\dot{x}$ (Section 8.4). We are interested in accumulation or loss due to flux, so we take the gradient in flux, which measures the rate of increase (or decrease) due to differences in flux rates on opposite sides of the unit. The flux gradients, or divergences, are $\nabla\cdot[N]\dot{x}$ and $\nabla\cdot[N]\dot{y}$ (Section 8.6). The divergence for the entire unit is $\nabla\cdot[N]\dot{\boldsymbol{x}}$ (note the boldface symbol to denote both horizontal directions). On a percentage basis the divergence is $N^{-1}\nabla\cdot[N]\dot{\boldsymbol{x}}$ (Section 8.6). If the divergence is positive, there is net loss (emigration). If the divergence is negative, there is net gain (immigration). Thus the conservation equation for any spatial unit is:

$$\frac{\dot{N}(t)}{N} = \frac{\dot{B}(t)}{N} - \frac{\dot{D}(t)}{N} + \frac{\nabla\cdot[N]\dot{\boldsymbol{x}}}{N} \tag{13.27a}$$

The conservation equation can be written as a rate of change in numbers.

$$\dot{N}(t) = \dot{B}(t) - \dot{D}(t) + \nabla\cdot[N]\dot{\boldsymbol{x}} \tag{13.27b}$$

Equation 13.27a is often used because recruitment and mortality are typically treated on a per capita basis. If they are, the divergence needs to be considered on a per capita basis as well.

Equation 13.27 displays the typical components of a conservation equation for a quantity: an *in situ* source term $\dot{B}(t)$, an *in situ* loss or sink term $\dot{D}(t)$, and a divergence

$\nabla \cdot [N]\dot{\boldsymbol{x}}$. One can write a similar conservation for other quantities. For example, the conservation equation for water on the landscape is typically written as:

$$\frac{dS}{dt} = PPT + ET + Q \tag{13.28}$$

In this equation, dS/dt is the rate of change in soil moisture, PPT is precipitation (which can be viewed as *in situ* production of water from vapor), ET is evapotranspiration (which can be viewed as *in situ* loss of water back to vapor), and Q is runoff (a lateral flux gradient or divergence).

Once we have a conservation law, a natural step is to compare its components by scaling them to each other as dimensionless ratios. For example, we expect fluxes to prevail at small spatial scales while demographics prevail at large scales.

$$\begin{aligned} \frac{\dot{B}(t) - \dot{D}(t)}{\nabla \cdot [N]\dot{\boldsymbol{x}}} &> 1 \quad \text{at} \quad \text{large} \quad \text{scales} \\ \frac{\dot{B}(t) - \dot{D}(t)}{\nabla \cdot [N]\dot{\boldsymbol{x}}} &< 1 \quad \text{at} \quad \text{small} \quad \text{scales} \end{aligned} \tag{13.29}$$

We can form these ratios any way we like, guided by our knowledge of the system. However, it helps to keep in mind that the number of nonredundant ratios for a conservation law is $n!/2!$, where n is the number of terms on the right side of the conservation equation. Given three components on the right side of Equation 13.28, the number of ratios that can be formed (number of pairs) will be $3!/2! = 3$ ratios. Having examined demographics relative to fluxes (Equation 13.29), we might then examine recruitment relative to divergence and mortality relative to divergence at a variety of spatial scales.

One common technique is to begin with a conservation equation that summarizes a full list of dynamics, then simplify the list using dimensionless ratios to identify those components that prevail at a given scale. From this the equation is rewritten as an approximation with fewer terms. A classic example in ecology is the role of advection in the dynamics of primary production in the sea (O'Brien and Wroblewski, 1973). These authors began with a conservation law for phytoplankton dynamics that included sources (production), concentration-dependent sinks (e.g., predation), concentration-independent sinks (e.g., loss during anoxic episodes), advection, and diffusion. Using the method of governing equations (Table 6.10, Box 13.8) these authors identified a dimensionless ratio Π_s, which is the horizontal flow scaled to the quantity $L \cdot ([\dot{B}]/B)$, where L is the length scale of interest and $[\dot{B}]/[B]$ is the phytoplankton production, $[\dot{B}]$ per unit of biomass concentration $[B]$. If $\Pi_s > 1$, advection needs to be included in more detailed models of phytoplankton production. If $\Pi_s < 1$, advection can be dropped as an extraneous detail. For the Sargasso Sea, $\Pi_s = 4 \cdot 10^{-2}$ at the length scale of energetically dominant eddies. At this scale advection can be omitted from a model of phytoplankton dynamics. For the Gulf Stream, $\Pi_s = 1$. Advection must be included in a dynamic model because it plays a role as important as production in determining the dynamics of phytoplankton patchiness.

One can undertake a dimensional analysis of the parameters in a conservation law via sequential elimination (Box 6.3), Rayleigh's method (Box 6.4), or Buckingham's Π

method (Box 6.5). However, if we want to evaluate the relative importance of terms in the equation, we need to analyze the equation as shown in Box 13.8. Analysis of the equation begins by selecting a characteristic value for each variable. These values are then used to reduce the equation to dimensionless form. The example in Box 13.8 was chosen for simplicity. Normally one would not set about simplifying an equation with only one source and one sink term. A realistic application of the governing equation method would typically have several source and sink terms.

Box 13.8 Dimensionless Ratios Obtained by the Method of Governing Equations

A conservation equation for red-tide organisms that bloom in isolated patches of low-salinity water and die if mixed outside the patch (Kierstead and Slobodkin, 1953):

$$\frac{dB}{dt} = rB + K_H \frac{d^2B}{dx}$$

$\dot{B} = dB/dt$	biomass production	[M]		$[T]^{-1}$
$r = \dot{B}/B$	maximum growth rate			$[T]^{-1}$
x	horizontal distance		[L]	
d^2B/dx^2	diffusive movement	[M]	$[L]^{-2}$	
K_H	horizontal eddy diffusivity		$[L]^2$	$[T]^{-1}$

Reduce variables to dimensionless ratios (Buckingham's method, Box 6.5):

$B_* = B/B'$ where B' is reference biomass
$x_* = x/x'$ where x' is patch scale
$t_* = t/r^{-1}$ where r^{-1} is time scale (e.g. doubling time)

Conservation equation rewritten in terms of dimensionless variables:

$$rB'\frac{dB_*}{dt_*} = rB'B_* + \frac{K_H B'}{x'^2}\frac{d^2B_*}{dx_*^2}$$

Divide by rB':

$$\frac{dB_*}{dt_*} = B_* + \frac{K_H}{rx'^2}\frac{d^2B_*}{dx_*^2}.$$

The first term on the right has a coefficient of 1
The second term on the right has a coefficient of $\Pi = K_H\, r^{-1}\, x'^{-2}$
Growth balances diffusive loss when the two coefficients are equal.
This gives the critical scale: $d_{crit} = (K_H/r)^{1/2}$ Compare Box 6.6.
The magnitude of each coefficient is used to decide whether to retain a term.
If a term has a coefficient that is small relative to other coefficients, that term is dropped. In this example, if Π were smaller than 1 (growth coefficient exceeds diffusivity), the eddy diffusion term would be dropped.

The governing equation method requires considerable prior knowledge compared to the method of similitude and the method of Π ratios (Table 6.9). The method requires an ecologically reasonable conservation equation, which requires familiarity with the system and its components. It requires that one choose a characteristic scale for each variable based on knowledge of the system. These requirements introduce ecological knowledge into the analysis.

The governing equation method, like other forms of dimensional analysis, produces an approximation rather than an exact solution. As with any dimensional analysis, the method is useful when an approximate solution will suffice or where an analytic solution is not available. Stephens and Dunbar (1993) used dimensional analysis to obtain an approximate solution to the marginal value theorem of Charnov (1976), which has no analytic solution, yet has proved highly useful in predicting behavior from gain constrained by cost.

It is evident from Box 13.8 that the governing equation method is not the easiest way to undertake a dimensional analysis. Sequential simplification (Boxes 6.3 and 6.5) requires less effort. Most people will find Rayleigh's method (Box 6.4) and the method of Π ratios (Box 6.5) easier than the method of governing equations. The principle reason for using the governing equation method is model simplification by permitting comparison of coefficients to determine whether a term is negligible and can be dropped (Box 13.8).

ANOTHER LOOK AT SECTION 13.8

Make a brief list of conservation equations (mass, energy, numbers) that you have encountered in biology.

13.10 Quantitative Reasoning

Equations in ecology are often developed from exploratory analysis of data rather than by reasoning about quantities. Most of the equations compiled by Peters (1983) were obtained by plotting one quantity against another and then fitting a regression line. Such equations are empirical, based largely on verbal reasoning. This analytic style relies heavily on statistics to discover pattern, and it employs yes/no hypothesis testing, in which the goal is to reject the hypothesis that the observed relation is no more than a matter of chance.

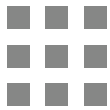
This exploratory style is highly effective in situations where little is known about what should be measured or what processes are important. But once patterns have been described and some of the important processes have been identified, it is then possible to adopt a style based on reasoning about quantities. Once knowledge has accumulated, it makes sense to use it.

The scaling relations developed in Chapter 6 for dinosaur running speed, rate of otter spread, and phytoplankton growth are all examples of quantitative reasoning. Reasoning about quantities, rather than statistical analysis, was used to decide which quantities should be considered and how to combine them. Revision of the dimensional matrix in light of biological knowledge led to scaling relations that could tested against new data.

The amount of reasoning that goes into an equation varies. Some relations between quantities are based on empirical relations derived completely from data. Some are based on the application of statistics to quantities in situations where some relation is expected on other grounds. An example would be a functional relation used to calculate primary production from wind strength in a coastal zone where wind-driven upwelling brings nutrients upward into the light. Little relation would be expected between wind strength and production by photosynthetic microbes *per se*. A stronger relation would be expected between production and upwelling-favorable wind stress, which takes into account both wind direction and amount of energy imparted to the water by the wind. Thus an equation to calculate production from wind stress is based on more specific reasoning than an equation based on wind speed alone, in much the same way that the solution to the surfing problem required reasoning about direction in which to apply force rather than relying on buoyancy alone.

Defined Concepts and Terms for Review and Future Reference

____ conservation law
____ dimensional homogeneity of an equation
____ parametric vs. variable quantity
____ parsimony
____ quantitative grammar
____ terms of an equation
____ theoretical vs. empirical scaling function



14 Coordinating Equations: Derivatives

It is only by bitter experience that we learn never to trust a published mathematical statement or equation … Misprints are common. Copying errors are common. Blunders are common. Editors rarely have the time or the training to check mathematical derivations. The author may be ignorant of mathematical laws, or he may use ambiguous notation. His basic premises may be fallacious even though he uses impressive mathematical expressions to formulate his conclusions. The present book—in text and in exercise—points again and again at published errors. But there are bound to be similar errors in this very book. Caveat lector! Let the reader beware!

—D. S. Riggs, *The Mathematical Approach to Physiological Problems*, 1963

14.1 Synopsis

Derivatives coordinate one function with another. Time derivatives coordinate descriptive functions with dynamics. Spatial derivatives coordinate descriptive functions with gradients. The rules for working with measured quantities apply to functions of measured quantities, including the derivatives and integrals of such functions. The chain rule turns out to be particularly useful in working with functions of scaled quantities.

Time and space derivatives also coordinate descriptive functions with scaling functions. The derivative with respect to temporal or spatial scale is called a zoom operator, to distinguish it from taking the derivative with respect to time or distance (panning).

Ecologists work with measurements that have finite ranges and resolutions. Differential equations produce values with infinite ranges and infinitely fine resolution. As a result, judgment is required in applying derivatives and integral functions to measured quantities.

14.2 Derivatives

Within *quantitative ecology*, defined as the use of scaled quantities in understanding ecological patterns and processes (Chapter 1), the role of calculus is to reexpress a concept written as an equation. Derivatives coordinate one equation with another. Thus they are an important means of relating measurements to underlying dynamics. For example, if aphid numbers increase exponentially during the spring because

doi: 10.1016/B978-0-12-627865-1.00014-3

recruitment b exceeds death d, the descriptive model reports a number $N(t)$ for any time t that we want:

$$N(t) := N_o e^{(b-d)t} \tag{14.1}$$

The dynamics are described by the derivative function $N(t)$:

$$\frac{dN}{dt} \equiv \dot{N}(t) := (b - d) \cdot N \tag{14.2}$$

This example demonstrates one of the principle uses of derivatives in ecology, which is to work back and forth from the dynamics, or time rate of change in a quantity, to the value at any point in time. The rules for taking derivatives and integrals permit rapid and accurate passage from the descriptive function $N(t)$ to the dynamic function $\dot{N}(t)$.

Many readers will have taken a calculus course that demonstrates the rules for derivatives and integrals (antiderivatives). Few ecologists regularly apply the tool, which rusts from disuse. The underlying idea is, however, a common mode of reasoning in ecology: examining the way the value of one quantity changes in relation to another. Expressed verbally, the idea of a *derivative* is the change in the value of one quantity with respect to another. The idea of a derivative is thus the same as that of scaling one quantity to another. Expressed graphically, the derivative is of course the slope of the curve of the function relating one quantity to another. The formal expression, as applied to a measured a quantity, is that the derivative is the ratio of the difference of one quantity over the difference of the second quantity, taken at a mathematically convenient (although hard to imagine) limit of infinitely fine differences. To rescue calculus for our purposes, which is to work with equations that express ideas about measured quantities, we have to introduce a "kluge": that a derivative has units (refer back to Box 13.1). This allows us to use the rules of calculus, as in Table 14.1. In this table, abstract notation is

Table 14.1 Commonly Used Rules for Derivatives of Functions with Units

L, Q, and M are symbols for quantities with different units. $\dot{L}, \dot{Q}$, and $\dot{M}$ are symbols for the derivative of these quantities with respect to time:

$\dot{L} \equiv dL/dt \qquad \dot{Q} \equiv dQ/dt \qquad \dot{M} \equiv dM/dt$

The derivative relative to the quantity is:

$$\dot{L}/L \equiv \frac{d\ln(L)}{dt} \qquad \dot{Q}/Q \equiv \frac{d\ln(Q)}{dt} \qquad \dot{M}/M \equiv \frac{d\ln(M)}{dt}$$

α and β are numbers, having no units.

K is a quantity that does not change with time: $dK/dt \equiv 0$

Here are four rules that relate functions to derivative functions. The derivative has been taken with respect to time.

Function	Derivative Function
1. $Q = k_{Q/k} \cdot K$	$\dot{Q} = 0$
2. $Q = k \cdot t^{\beta}$	$\dot{Q} = k \cdot \beta \cdot t^{\beta-1}$
3. $Q = k_{Q/L} \cdot L + k_{Q/M} \cdot M$	$\dot{Q} = k_{Q/L} \cdot \dot{L} + k_{Q/M} \cdot \dot{M}$
4. $Q = k_{Q/LM} \cdot L^{\alpha} \cdot M^{\beta}$	$\dot{Q}/Q = \alpha \cdot \dot{L}/L + \beta \cdot \dot{M}/M$

The chain rule is easy to remember because it works like unit cancellation.

$$\frac{dQ}{dt} = \frac{\cancel{dL}}{dt} \cdot \frac{dQ}{\cancel{dL}}$$

used for the sake of generality. Time derivatives *d/dt* are used because these are familiar. The rules apply to any derivative, not just *d/dt*.

Derivatives are ratios that can be added and subtracted, following the same rules as any other quantity (Table 4.4). For our purposes derivatives cannot be added unless they have the same units. Derivatives with different units can be added if there is a conversion factor that transforms one of the derivatives to the same units as the other derivative. As was the case with the algebraic operations described in Chapter 4, care in the use of symbols contributes to effective use of derivatives in quantitative ecology.

ANOTHER LOOK AT SECTION 14.2

Work out the derivative of the species area curve with respect to island area A. The species area curve is $N_{sp} = kA^z$. Provide a verbal interpretation of the derivative function.

Box 14.1 Seed Density Gradients Calculated from Rules in Table 14.1

$[N] \equiv$ seeds per unit area, with dimensions of $\#\,L^{-2}$

$\nabla[N] = \dfrac{d[N]}{dr}$ The one-dimensional radial gradient

1. The density is a constant value: $[N] := K = 100/m^2$
   $\nabla[N] = 0$ (the gradient is zero at all points) (Rule 1)
2. The density is inversely proportional to the square of radial distance r from point of release.
   $$[N] := k \cdot r^{-2} \qquad k = 250$$
   $$\nabla[N] = -2k \cdot r^{-3} \quad \text{(Rule 2)}$$
   The density gradient at $r = 5\,m$
   $$\nabla[N] \Rightarrow -2(250)(5\,m)^{-3} = -(4\,m^{-2})\,m^{-1}$$
3. The density of viable seeds depends on distance from point of release and on a survival gradient k_{live} along a radius.
   $$[N] := k \cdot r^{-2} + k_{live} \cdot r \qquad k = 250 \qquad k_{live} = (20\,m^{-2})m^{-1}$$
   $$\nabla[N] = -2k \cdot r^{-3} + k_{live} \quad \text{(Rule 3)}$$
   The density gradient at $r = 5\,m$
   $$\nabla[N] \Rightarrow -(4\,m^{-2})m^{-1} + (20\,m^{-2})m^{-1} = (16\,m^{-2})m^{-1}$$
4. The density is defined as the ratio of numbers to area:
   $$[N] \equiv \frac{N}{A} \qquad \nabla[N] = \nabla(N^1 A^{-1})$$
   $$[N]^{-1} \nabla [N] = 1 \cdot \nabla N/N + -1 \cdot \nabla A/A \quad \text{(Rule 4)}$$
   $$= \nabla\, N/N - 2/r$$
   The percent gradient along a radius. $\nabla A/A = 2r/r^2 = 2/r$
   Let the percent gradient be $\nabla N/N := -10\%/m$
   At $r = 5\,m$ the density gradient is
   $$\frac{1}{[N]}\frac{d[N]}{dr} \Rightarrow \frac{-10\%}{m} - \frac{2}{5\,m} = -50\%/m$$

14.3 Spatial Gradients from the Density Function

The *spatial derivative d/dx* yields the spatial gradients from functions that describe density relative to location. To demonstrate the computations, each of the four rule pairs in Table 14.1 is applied to a function describing the density of seeds (Box 14.1). Units of density, *[N]* = numbers/area, are used in all four examples. Gradients in one dimension (such as along a transect) are used to illustrate the spatial derivative *d/dx*. Each calculation in Box 14.1 begins with a function describing density based on an idea. The one-dimensional or along-transect gradient in density is then obtained by applying rules for derivatives from Table 14.1. A gradient has been calculated for a distance of $x = 5\,\text{m}$ in each case.

In the first example in Box 14.1, the seed density is constant, and hence the density gradient comes out to be zero at all radial distances, including $r = 5\,\text{m}$. In the second example, the seed density decreases as the square of the distance from release point. The second rule pair in Table 14.1 is used to obtain the gradient as function of distance from point of release. The gradient was calculated as $-(4\,\text{m}^{-2})\text{m}^{-1}$ at 5 m from the point of release. The gradient has units of density per m, which is to say, seeds m^{-2} m^{-1}. As expected, the gradient is negative along a transect running out from the point of release. In the third example in Box 14.1, seed density again decreases with the square of the distance from release point. Survival is taken as positive with increasing distance from point of release, at a constant value of k_{live}. In order for the expression to be dimensionally homogeneous, the parameter k_{live} must have units of (seed m^{-2})m^{-1}. The third rule pair in Table 14.1 is then used to derive the gradient in seeds along a radius. The density gradient is computed at a distance of 5 m from point of release. The gradient is positive at 5 m because the survival percentage at this distance exceeds the gradient due to decreasing density with distance.

In the fourth example in Box 14.1, the density is defined as the ratio of seed numbers to area. The fourth rule pair in Table 14.1 is used to calculate the gradient for seeds that disperse along a radius from point of release, with gradient in numbers of -10% m^{-1} along the radius. The negative gradient in numbers along the transect ($\nabla N/N < 0$) amplifies the rate of decrease in density with radial distance relative to an even distribution of seeds along a radius ($\nabla N/N = 0$).

Gradients in energy gain or loss shape the behavior of mobile organisms. An example is the costs of defending nesting territories, which vary considerably in size among individuals within a bird species. Territories provide food for adults and chicks during a period when chicks are immobile. So energy gain and cost depend on area as a first approximation. What are the costs? If energy expenditure ($\dot{E}$ = Watt) for defense depends on the territory diameter ($2r$, in meters), a simple function that expresses this idea is:

$$\dot{E}(r) = 2r \cdot K_{watt/m} \tag{14.3}$$

The factor $K_{Watt/m}$ is present to scale energy expenditure to radial distance. The functional expression is not dimensionally homogeneous unless K has units of Watt m^{-1}. This is a rigid scaling factor that describes how much energy the bird expends per unit of radius of area defended (radius = $d/2$). A rough estimate of K for a 30 gram bird with an existence energy of $\dot{E} = 1.2$ Watt (104 kJ day^{-1}) defending a 75 m radius (1.75 ha) territory is $K_{Watt/m} = 0.016$ Watt m^{-1}. Existence energy $\dot{E}$ for a 30 gram bird was

calculated from Walsberg's (1983) expression for nonpasserine birds. Territorial area of a 30 g bird was calculated from Schoener's (1968) equation. Radius was calculated from area, assuming a roughly circular area $A = \pi r^2$.

Once the relation between energy cost and radius has been stated as a first approximation or caricature, the change in cost with change in radius can be derived:

$$\frac{d\dot{E}(r)}{dr} = 2 \cdot K \tag{14.4}$$

This is the gradient in energy cost. A convenient symbol for this is $\nabla\dot{E}$, read as "the gradient in energy expenditure." The symbol $\nabla\dot{E}$ represents a quantity derived from $\dot{E}$, just as the symbol for energy expenditure $\dot{E}$ represents a quantity derived from energy E. A bird trying to expand its territory faces a constant spatial gradient in energy expenditure $\nabla\dot{E}$ much like walking up a hill with constant slope:

$$\nabla\dot{E} = 2 \cdot \mathrm{K} \tag{14.5}$$

If this caricature is accurate (it will not be exact), the cost of expanding a 30 m radius territory is the same as expanding to a 40 m radius territory. This idea can be expressed as a proportion: the ratio of the energy cost of a large territory $\dot{E}_{\text{big}}$ to a small territory $\dot{E}_{\text{small}}$. The statement of proportion is:

$$\dot{E}_{\text{big}} \text{ is to } \dot{E}_{\text{small}} \text{ as radius}_{\text{big}} \text{ is to radius}_{\text{small}}$$

This is readily translated into a scaling relationship:

$$\frac{\dot{E}_{big}}{\dot{E}_{small}} = \frac{r_{big}}{r_{small}} \tag{14.6}$$

This scaling relationship says that if the ratio of the bigger to smaller territory radius is, for example, 1.5:1, the ratio of energy costs is also 1.5:1. The scaling relation (Equation 14.6) is readily transformed into a scaling relation that expresses the gradient in energy expenditure:

$$\frac{\dot{E}_{big}}{r_{big}} = \frac{\dot{E}_{small}}{r_{small}} \tag{14.7}$$

Because this is a model, it needs to be tested against measurements.

ANOTHER LOOK AT SECTION 14.3

Name several important physical and biological gradients going up a mountain. For each, state a testable density function $Q = f$ (height), then derive the gradient $\nabla_{vertical}(Q)$. Use your own symbol, not Q.

14.4 Quantitative Reasoning with the Chain Rule

The chain rule for working with derivative functions, shown in Table 14.1, has the same format as unit cancellation. In quantitative ecology, the *chain rule* is a simple device for drawing conclusions about the relation of one quantity to another. The chain rule is a convenient way to reason about quantities, as in the following example.

What is the relation between population increase and mean crowding? To phrase this more precisely, how is the rate of change in numbers N related to the rate of change in mean crowding $\dot{M}^*(N)$? Here is the same question written in symbolic form:

$$\frac{dM^*}{dt} \stackrel{?}{=} \frac{dN}{dt} \tag{14.8a}$$

This puzzle can be solved analytically by taking the derivative of M^* with respect to time. The chain rule for derivatives gives a reliable answer with just algebra. The chain rule works the same way as cancelling units and dimensions (Table 14.1), so we use it to "fill in the blanks." In other words, what quantity relates crowding to population density? Here is the question written in symbolic form, with question marks for missing quantities:

$$\frac{d\,M^*(N)}{dt} = \frac{d\,N}{d\,t} \cdot \frac{?}{?} \tag{14.8b}$$

The question marks are replaced so that the "units" (in this case, dN, dt, $dM^*(N)$, etc.) cancel correctly:

$$\frac{d\,M^*(N)}{dt} = \frac{d\,N}{d\,t} \cdot \frac{d\,M^*(N)}{d\,N} \tag{14.8c}$$

The relation between time rate of change in numbers $\dot{N}$ and the time rate of change in mean crowding $\dot{M}^*(N)$ depends, logically enough, on a quantity that describes how crowding changes with density. A lot can be learned by visualizing this new quantity, which turns out to be a measure of patchiness (Lloyd, 1967). Verbal models of ways in which crowding changes with overall change in population size can be developed from knowledge of the biology of a species. Many animals, for example, tend to crowd into already established areas rather than spreading to new areas as population increases. Far more than if new recruits move into new areas, this will increase crowding as population size increases, and so $d\,M^*(N)/dN$ exceeds zero. In contrast, completely passive dispersal stages of plants have no mechanism for preferential settlement into populated sites, so we expect $d\,M^*(N)/dN$ to be zero. But there are mechanisms that raise $d\,M^*(N)/dN$ above zero. For example, passive dispersal stages of benthic marine invertebrates sample the substrate before committing themselves to a site, thus avoid settling into poor sites. Thus they can crowd into populated sites even though their dispersal movements are passive.

ANOTHER LOOK AT SECTION 14.4

What is the relation of the gradient in mean crowding dM^*/dx to the gradient in numbers dN/dx? Provide an interpretation.

14.5 Variance Functions with the Delta Method

The production and loss of variability (Chapter 10) represents a distinctly biological approach to highly variable phenomena. Because of the central role of variability in ecological thought, it is of interest to be able to calculate a variance either from another variance or from a functional expression. For example, what is the variance in respiration rate, given a variance in body size? What is the variance in swimming speeds in units of $cm^2\ s^{-2}$, given a variance in units of $m^2\ s^{-2}$? What is the variance in territory area, given a variance in distance between nests?

An expression for one variance in terms of another is called a *variance function*. These can be worked out using the Delta method (e.g., Seber, 1980), which relies on the rules for taking derivatives, as shown in Table 14.2. The first example is for the variance of any quantity Q plus a constant k, shown in Box 14.2. The analytic sequence in Box 14.2 has been recast in Table 14.2 as a generic recipe for obtaining one variance from another, using the rules for derivatives.

Box 14.3 shows another example, this time for the variance of the product of a quantity and constant. The analytic sequence is presented without narrative using the steps in Table 14.2. This result establishes the units of a variance. For the case of Q in units of grams m^{-2}, each value of Q is written as a number q and its unit: q grams m^{-2}. The result in Box 14.3 is then applied:

$$\text{var}(Q) = \text{var}(q \text{ grams } m^{-2}) = \text{var}(q) \cdot \text{ grams}^2 m^{-4} \tag{14.9}$$

The variance of a quantity has squared units, which are inconvenient and often uninterpretable. One solution is to reduce the quantity to a dimensionless ratio and then take the variance. An example is Box 11.1, where the variance in per capita change in organism number was used. Another solution is to take the square root of the variance, as in Box 10.1. Yet another solution is to scale the variance to the mean, resulting in an ensemble quantity cd(Q) with the same units as the mean (Table 10.1).

Another useful application is the variance of the sum of two quantities. Box 14.4 shows the derivation. One application of this result is visualizing the results of taking a spatial or temporal variance at a larger scale by grouping contiguous measurements.

Table 14.2 Generic Recipe for Obtaining One Variance from Another Using the Delta Method

1. State, in words, the variance that is to be found.
2. Translate this into a formal expression:
   var*(such and such)* = ?
   "What is the variance in such and such?"
3. Set the terms inside the parentheses equal to a dummy or stand-in variable u: u = such and such.
4. Take the derivative du: $du = d$(such and such)
5. Solve for du, using the rules for derivatives.
6. Square both sides of this expression, and rearrange terms as necessary.
7. Substitute var(u) for $(du)^2$, var(Q) for $(dQ)^2$, etc.
8. Write out the result in convenient form:
   var*(such and such)* = ….

Box 14.2 Use of the Delta Method to Obtain the Variance of a Quantity Q Plus a Constant k

The variance to be found, in symbolic form:

$$\text{var}(Q + k)$$

Set the terms inside the parentheses equal to a dummy variable u.

$$u = Q + k$$

Take the derivative du:

$$du = d(Q + k)$$
$$du = dQ$$

The symbol du is short for du/dx, where x in this case is any variable, not just distance eastward from a point. Because x can be any variable, the operator d/dx has been shortened to just the letter d.

Take the square of both sides of this expression:

$$(du)^2 = (dQ)^2$$

This expression does not require rearrangement, but more complex expressions will at this point.

Replace $(du)^2$ with $\text{var}(u)$, $(dQ)^2$ with $\text{var}(Q)$:

$$\text{var}(u) = \text{var}(Q)$$

The solution:

$$\text{var}(Q + k) = \text{var}(Q)$$

The variance of Q plus a constant is equal to the variance of Q.

Box 14.3 Use of the Delta Method to Obtain the Variance of the Product of a Quantity and a Constant

1. What is the variance of the product of a quantity Q and a constant k ?
2. $\text{var}(kQ) = ?$
3. $u = kQ$
4. $du = d(kQ)$
5. $du = k dQ$
6. $(du)^2 = k^2 (dQ)^2$
7. $\text{var}(u) = k^2 \text{var}(Q)$
8. $\text{var}(kQ) = k^2 \text{var}(Q)$

If we group a sequence of measurements by adjacent pairs, take the sum of each pair, then take a new variance for the pairs, the resulting variance $\text{var}(Y + Z)$ will be equal to the variance of the original measurements, $\text{var}(Y) + \text{var}(Z)$ plus an additional component of variance that depends on whether neighboring values are correlated. This

Box 14.4 Use of the Delta Method to Obtain the Variance of the Sum of Two Quantities

1. What is the variance of the sum of two quantities, Y and Z ?
2. $\text{var}(Y + Z) = ?$
3. $u = Y + Z$
4. $du = d(Y + Z)$
5. $du = dY + dZ$
6. $(du)^2 = (dY)^2 + (dZ)^2 + 2\, dY\, dZ$
7. $\text{var}(u) = \text{var}(Y) + \text{var}(Z) + 2\text{cov}(Y,Z)$
8. $\text{var}(Y + Z) = \text{var}(Y) + \text{var}(Z) + 2\text{cov}(Y,Z)$

additional component is the $\text{cov}(Y,Z)$ at lag 1, a quantity that in turn can be converted into an autocorrelation coefficient. Thus binning by adjacent pairs will measure autocorrelation. The autocorrelation is positive if $\text{var}(Y + Z)$ exceeds $\text{var}(Y) + \text{var}(Z)$, negative if $\text{var}(Y + Z)$ falls short of the sum of variance of the two components.

Another useful variance is that for the product of two quantities. The prospect of working through this algebraically is daunting. With the delta method, it is relatively straightforward. The covariance between the two quantities again figures in the result:

$$\text{var}(\text{Y} \cdot \text{Z}) = \text{Z}^2 \cdot \text{var}(\text{Y}) + \text{Y}^2 \cdot \text{var}(\text{Z}) + 2 \cdot \text{Y} \cdot \text{Z} \cdot \text{cov}(\text{Y,Z}) \qquad (14.10)$$

The derivation of Equation 14.10 is left to the reader.

If we take the view that one of the distinguishing features of biological thought is that the variance is an interesting quantity subject to loss and gain, a means of computing variances becomes important. The delta method, based on the rules for derivatives, makes such computations relatively easy.

ANOTHER LOOK AT SECTION 14.5

The expression $\dot{N} = rN$ describes exponential population growth in a patch. What is the variance in $\dot{N}$ among patches?

14.6 Scaling Operators

A *scaling operator* is a derivative with respect to change in scale. These operators provide the mathematical apparatus for working with scaling functions and scaling relations. Scaling operators can be taken with respect to noniterative scaling (Equations 2.1 and 2.2). An example is a comparison of species numbers across plots of different sizes (Figure 1.2) or across islands of different sizes. Another example is a comparison of running speed in

Table 14.3 Operators for Temporal Analysis (Panning) and Temporal Scaling (Rating, Coarse Graining, Accumulating, or Lagging)

Temporal analysis (panning across contiguous units of same duration):		
$\boldsymbol{h}$	unit vector in time t	
$t\boldsymbol{h}$	time since $t = 0$	d/dt
Temporal scaling (rating across units of different duration):		
T_{ref}	duration of reference unit.	
i	label linking a quantity to size of its unit	
$T = T_i/T_{ref}$	ratio of unit size to reference unit size	$^*d/dT$
Temporal scaling (zooming through units of different duration):		
n	units per bin at resolution of analysis	d/dn
$t = n\boldsymbol{h}$	bin size (inner scale) at resolution of analysis	$^*d/dt$
N	number of steps at range of analysis	$^*d/dN$
$T = N\boldsymbol{h}$	range size or outer scale of analysis	d/dT
$f = l/T = n/N$	frequency of measurement	d/df
$2/f$	wavelength	
Temporal scaling (increasing unit size by accumulation):		
$T = N\boldsymbol{h}$	range size or outer scale of analysis	d/dT
Temporal scaling (skipping through units of same duration):		
k	lag or separation between two points in time	d/dk

*Possible operators, not usually encountered.

relation to leg length across a series of animals that differ in body size (Chapter 6). This is the classic form of scaling found in text treatments of dimensional analysis. Scaling operators can also be taken with respect to iterative scaling (Equations 2.3 and 2.4). An example is the change in murre and capelin patchiness with change in spatial resolution along a transect (Figure 10.3e). Another example is the change in autocorrelation with change in lag (spatial separation) of murres and capelin (Figure 10.3b).

Tables 14.3 and 14.4 list commonly used temporal and spatial scaling operators. The tables demonstrate four different *scaling maneuvers* (rating, coarse graining, accumulating, and lagging, as shown in Figure 7.1). These maneuvers differ from *panning*, which passes from one equal size unit to the next. Panning underpins the familiar procedures of dynamic analysis by taking the time derivative d/dt or descriptive spatial analysis by taking the gradient d/dx. Scaling operators are far more diverse than panning. The first scaling maneuver, *rating*, compares values across separate units that differ in size, as in Figure 1.2. The second scaling maneuver, that of *coarse-graining*, changes the unit size by successive aggregation (or subdivision) of contiguous and equal sized units, keeping the range fixed, as in Figure 10.3e. The third scaling maneuver, *accumulating*, changes scale by adding new units to increase the range (Figure 10.3d). Accumulating and coarse graining are analogous to zooming with a camera. The fourth scaling maneuver, *lagging*, changes scale by comparing equal-sized units at increasingly large separations (lags), as shown in Figures 10.3b or 10.3c.

The formalism of scaling operators will allow us to work from one type of scaling to another. For example, the operator d/dL expresses the change in a quantity as measurements

Table 14.4 Operators for Spatial Analysis (Panning) and Spatial Scaling (Rating, Coarse Graining, Accumulating, or Lagging)

Spatial analysis (panning across contiguous units of same size):		
$\boldsymbol{i}$	unit vector in direction x	
$x\boldsymbol{i}$	distance from point $x = 0$	d/dx
Spatial scaling (rating across units of different size):		
L_{ref}, A_{ref}, V_{ref}	Length, Area, or Volume of reference unit.	
i	label linking a quantity to size of its unit	
$L = L_i/L_{ref}$	ratio of unit size to reference unit size	d/dL
Spatial scaling (coarse graining to units of different size):		
n	units per bin at resolution of analysis	d/dn
$l = n\boldsymbol{i}$	bin size (inner scale) at resolution of analysis	d/dl
N	number of steps at range of analysis	
$L = N\boldsymbol{i}$	range size or outer scale of analysis	d/dL
$f = l/L = n/N$	frequency of measurement	d/df
$2/f$	wavelength	
Spatial scaling (increasing unit size by accumulation):		
N	number of steps at range of analysis	d/dN
Spatial scaling (skipping through units of same size):		
k	lag or separation between two points in space	d/dk

accumulate along a transect. The operator d/df expresses change in a quantity as the frequency of measurement goes from small (a few large bins) to large (many small bins). How are the two zoom operators related? The chain rule provides the answer (Box 14.5).

The second example in Box 14.5 shows the relation between temporal panning to temporal zooming. Interesting simplifications of Pollett's equation arise when either the

Box 14.5 Use of Chain Rule to Work Out the Relationship Between Scaling Operators

1. What is the relation of the zoom operator d/dL to the operator d/df?
   Begin with the chain rule:

$$\frac{d}{d\ln L} = \frac{d}{d\ln f}\,\frac{d\ln f}{d\ln L}$$

$$f = L^{-1} \quad \text{hence } \ln f = -\ln L$$

$$\text{and } d\ln f = -d\ln L$$

   Consequently:

$$\frac{d}{d\ln L} = (-1)\frac{d}{d\ln f}$$

2. What is the relation of the temporal zoom operator $d/d\boldsymbol{h}$ to the temporal panning operator $d/d\boldsymbol{t}$?
   Position in time $\boldsymbol{t}$ is the product of step size $\boldsymbol{h}$ and step number t.
   Apply the chain rule*.

$$\frac{dQ}{d\boldsymbol{h}} = \frac{dQ}{dt\boldsymbol{h}} - \frac{dt\boldsymbol{h}}{d\boldsymbol{h}}$$

$$\frac{dQ}{d\boldsymbol{h}} = \frac{t + \boldsymbol{h}\ dt/d\boldsymbol{h}}{\boldsymbol{h} + t\ d\boldsymbol{h}/dt}\ \frac{dQ}{dt}$$

*Thanks to Troy Pollett for this derivation.

resolution or the range is fixed, which is the case for the common scaling maneuvers (Figure 7.1).

Scaling operators are useful in working with measurement relations, scaling relations, scaling functions, and the results of dimensional analysis. Box 14.6 shows the application of scaling operators to scaling functions derived from scaling relationships. Distinguishing one scaling operator from another forces us to consider the measurement operation used to obtain the scaling function. It tells us whether scaling functions are comparable (based on the same scaling operator) or not.

Box 14.6 begins with a general formula for applying a scaling operator to a noniterative scaling relation. The first example in Box 14.6 demonstrates the close relation between dimensional analysis (Chapter 6) and derivatives interpreted as scaling operators. The noniterative scaling operator is that of rating, because each organism is considered a separate system and a measure is made only once. This scaling operator summarizes, as an exponent, the information obtained from dimensional analysis. This operator yields the scaling of running speed to leg length that applies across systems (animals). The second example in Box 14.6 illustrates an iterative scaling relation for phytoplankton patch size. Dimensional analysis showed that as a first approximation, patch size scales as the ratio of diffusivity K_H to growth rate r. If we consider contiguous measurements along a transect across patches with diameter L, the derivative function of interest is either d/dL or, equivalently, d/dN, where $N = L/\boldsymbol{i}$, the number of steps of length $\boldsymbol{i}$ across the patch (see Table 14.4). The logarithm will be taken so the dimensionless ratio N needs to be used. The derivative function for patch diameter with respect to N is more complex than the scaling function. However, this derivative function can be simplified by considering its components. We expect the percentage growth rate to vary with several factors, but not with distance L or step number N across the patch. Consequently, $dr/dN = 0$. In contrast, we do expect measured diffusivity to change as measurements are accumulated along the diameter. Taking the derivative functions leads us to the conclusion that measured diffusivity will depend on length scale of measurement, relative to the diameter.

Box 14.6 Scaling Relationships, Scaling Functions, and Scaling Operators (Rating and Accumulation)

General formula:

The scaling relationship:

$$Q_* = \frac{Q}{Q_{ref}} = \left(\frac{X}{X_{ref}}\right)^{\beta} = X_*$$

The scaling function:

$$Q = (Q_{ref} \cdot X_{ref}^{\beta})\, X^{\beta} = k \cdot X^{\beta}$$

The derivative function:

$$\frac{d \ln Q_*}{d \ln X_*} = \frac{X_*}{Q_*}\frac{dQ_*}{dX_*} = \frac{X_*}{Q_*}\, \beta\, k\, X_*^{\beta-1} = \beta\frac{kX_*^{\beta}}{Q_*} = \beta$$

Example from Box 6.4 (rating units of different size):

The scaling function:

$$v_* = \prod \cdot legL_*^{1/2} \cdot g_*^{1/2}$$

The scaling relation:

$$\frac{v}{v_{ref}} = \left(\frac{legL}{legL_{ref}}\right)^{1/2}$$

The derivative function:

$$\frac{d \ln v_*}{d \ln legL_*} = 1/2$$

Example from Box 6.5 (scaling by accumulation):

The scaling function:

$$diam = \prod \cdot (K_H/r)^{1/2}$$

The derivative function:

$$\frac{d\ln(diam_*)}{d\ln(N)} = \frac{N \cdot \Pi}{diam}\left(r^{-1/2}\frac{d(K_H^{1/2})}{dN} + K_H^{1/2}\frac{d(r^{-1/2})}{dN}\right)$$

$dr/dN = 0$

$$\frac{d\ln(diam_*)}{d\ln(N)} = \frac{N}{2}\left(K_H^{-1}\frac{dK_H}{dN}\right)$$

Scaling operators allow computation from local to larger-scale expectations. To illustrate the concept, we examine the variance in the density of a quantity. The value of this quantity at one resolution scale cannot be computed directly from the variance at another scale, a result confirmed again and again, from Mercer and Hall (1911) onward. As we saw in the last section, this is because variance at one scale depends on the covariance among units at a finer scale. Applying the result in the last section is cumbersome, however. In practice we require a function that expresses the rate of change in spatial variability with change in resolution scale l, for which $l = n\boldsymbol{i}$, $\boldsymbol{i}$ is the unit vector in the x direction, and n is the number of such units at the resolution of analysis. The zoom operator with respect to spatial resolution is the derivative d/dl. This operator coordinates the expression for the variance at any given resolution with the expression for the rate of change in variance relative to change in resolution scale. A convenient model is that variance is a fixed quantity with no additional components of variance at larger scales:

$$\text{var}(Q) = k \tag{14.11a}$$

$$\frac{d\,\text{var}(Q)}{dl} = 0 \tag{14.11b}$$

The model is convenient because additional components of variance do not appear as the resolution scale (i.e., grain size) increases. If the model is supported by evidence, an estimate at a small-scale var(Q) could be used at a larger scale.

In general, variances are not fixed in ecology. A better model is that variance increases in direct proportion to resolution scale $l = n\boldsymbol{i}$.

$$\text{var}(Q) = k\,l \tag{14.12}$$

$$\frac{d\,\text{var}(Q)}{dl} = k \tag{14.13}$$

This model says that doubling the bin size will double the variance, trebling the bin size will treble the variance, and so on.

There is much room for better understanding of patterns of change in variance with change in scale (as in Figures 10.1–10.4). There is even more room for better knowledge of patterns of change in association of ecological variables as a function of change in scale (as in Figure 10.5). These are theoretically interesting topics (e.g., Levin, 1992) that are driven by the practical problems of making ecosystem scale calculations from data that are almost inevitably confined to limited spatial and temporal scales (Holling, 1992; Levin, 1992; Schneider, 2001). Scaling analysis has undergone a proliferation of disparate techniques (Legendre and Fortin, 1989; Turner and Gardner, 1991; Legendre and Legendre, 1998), with far more attention to statistical technique than to underlying measurement and scaling maneuver (rating, coarse graining, lagging). Progress will depend in part on the capacity to compare and contrast findings from apparently disparate scaling techniques. Scaling operators provide the mathematical basis for comparing the results of different scaling techniques, from coarse graining to lagging and rating.

ANOTHER LOOK AT SECTION 14.6

Of the scaling operators listed in Tables 14.3 and 14.4 how many have you encountered?

Defined Concepts and Terms for Review and Future Reference

____ chain rule
____ derivative
____ panning
____ scaling maneuvers (compare to Chapter 7)
____ scaling operators
____ spatial derivative
____ variance function

15 Equations and Uncertainty: The General Linear Model

Fisher's famous paper of 1922, which quantified information almost half a century ago, may be taken as the fountainhead from which developed a flow of statistical papers, soon to become a flood. This flood, as most floods, contains flotsam much of which, unfortunately, has come to rest in many text books. Everyone will have his own pet assortment of flotsam; mine include most of the theory of significance testing, including multiple comparison tests, and non parametric statistics.

—J. Nelder, *Mathematical Models in Ecology*,
British Ecological Society Symposium, 1971

15.1 Synopsis

The general linear model provides a comprehensive framework for analysis of scaling functions. The model based approach emphasizes biological relevance and the interpretation of parameters. The focus is on the degree of uncertainly associated with each parameter estimate and on the adequacy of the error model. Most statistical packages now include a general linear model (GLM) routine, often with a graphics interface to assist in writing the model. The GLM relates a response variable to one or more explanatory variables. These can be categorical (ANOVA), ratio scale (regression), or both (ANCOVA). Each observation is compared to the model with a data equation. The deviations due to each term in the model are obtained by estimating model parameters, usually means or slopes. These parameters are biologically interpretable quantities with units and dimensions.

In the statistical analysis of scaling functions. parameter estimates and their confidence limits are often useful. The conventional machinery of hypothesis testing is most useful in comparing one scaling parameter to another. Hypothesis testing for the GLM uses the F-ratio, which scales the improvement in fit to the residual variation. A p-value is then computed from a statistical distribution (F or t) to decide whether the improvement in fit is too large to be due to chance. The accuracy of this p-value rests on four assumptions: (1) residuals sum to zero; (2) residuals are independent; (3) residuals are homogeneous; and (4) residuals are normal. Graphical evaluations of these assumptions are preferable to the use of hypothesis testing, which detects violations when they do not matter (at large sample sizes) while missing violations when they do matter (at small sample sizes). If serious violations are evident, p-values can be computed by randomization methods.

doi: 10.1016/B978-0-12-627865-1.00015-5

A generic recipe serves as a guide to model-based analysis of scaling functions. The first phase entails writing the model, estimating parameters, computing the residuals, and evaluating the model, including a check on whether the form of the model is appropriate and a check on assumptions concerning the distribution of errors. The next phase is hypothesis testing, although this step can often be skipped in favor of a report of parameter estimates with some measure of uncertainty.

Power law scaling and measurement relations are evaluated within the framework of the GLM by taking the logarithm of both response and explanatory variable. Power laws are compared by analysis of covariance, a special case of the GLM. Compound scaling relations contain more than two quantities, usually a quantity that is the ratio of two quantities. The assumption that both parts of the ratio have the same exponent (with opposite sign) is readily evaluated with multiple regression.

The standard method of parameter estimation for the general linear model assumes no error in the explanatory variable. If the variable is measured with error, reduced major axis (RMA) regression is often recommended because it assigns uncertainty to both the explanatory and the response variables. Reducing uncertainty in the explanatory variable is more effective than adjustment via RMA regression. One technique to reduce uncertainty is to use means instead of single values of the explanatory variable. Another technique is to estimate parameters via the scope of measurement relations.

15.2 Model-Based Statistical Analysis

In ecology, statistics are routinely presented as a collection of recipes. The basic ingredients are null and alternative hypotheses, a statistic (*F, t,* or chi-square), a *p*-value, and the declaration of a decision. The recipes focus on the logic of the null hypothesis rather than on the biological relevance of the model, the interpretation of parameters, the adequacy of the error model, or the degree of uncertainty associated with each parameter estimate. The recipe collection is huge. Widely used texts typically cover the following tests: one-sample hypotheses, two-sample hypotheses, paired-sample hypotheses, one-way ANOVA, multiple comparisons, two-way ANOVA, hierarchical ANOVA, multiway ANOVA, regression, multiple regression, analysis of covariance (ANCOVA), polynomial regression, logistic regression, goodness of fit tests, and contingency tests. The menus of widely used statistical packages (Minitab, SPSS, SAS, Systat) contain even longer lists of tests. Choosing from such a long list is daunting and, as it turns out, often unnecessary.

For problems in scaling, statistical analysis will usually entail some form of regression: How does some quantity Q vary as a function of another set of quantities $X1, X2...$ etc.? We can employ *model-based statistics*, which focus on a response variable in relation to one or more explanatory variables. The generalized linear model (Nelder and Wedderburn, 1972; McCullagh and Nelder, 1989), one of the major developments in statistics in the last quarter of the 20th century, allows analysis based on any of several error distributions. For practical reasons the presentation here will emphasize a special case, the *general linear model* (Figure 15.1). The GLM can be undertaken in nearly all statistical packages, almost always with a graphics interface and menu to assist in learning to use model-based analysis. The generalized linear model will be used as an organizational device in presenting model-based statistics, with emphasis on the special case of the general linear model. This approach is readily grasped and executed by biology students at

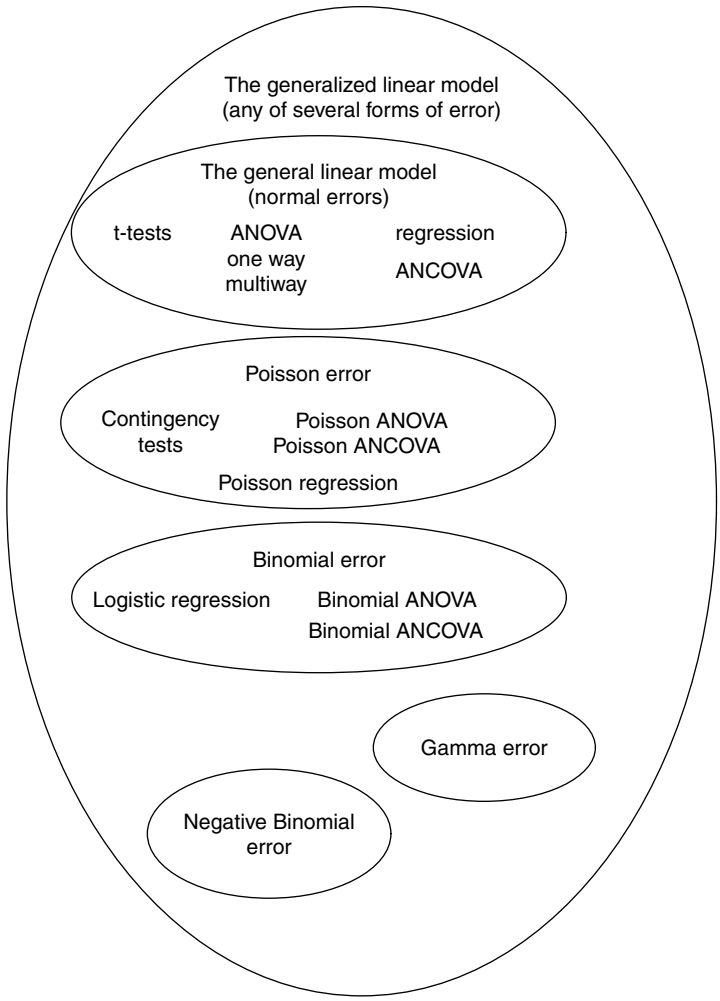


FIGURE 15.1 Named Statistical Tests that are Special Cases of the General Linear Model (GLM) and the Generalized Linear Model (GzLM).

the fourth-year undergraduate and first-year graduate levels. The presentation here begins with the concept of data equations. It then moves on to computing the fit of the model to the data within the familiar framework of least squares and the ANOVA table. After a discussion of assumptions for computing *p*-values in an ANOVA table, it moves to a generic recipe illustrated by several examples. Finally there is a brief tour of the differences between the general and the generalized linear model.

ANOTHER LOOK AT SECTION 15.2

Biologists agree that the list of current bird species is finite and rapidly approaching completion. Do you think that a list of statistical tests is finite or could ever be complete? Why or why not?

15.3 Data Equations for the General Linear Model

The *general linear model* matches data to expected values. It has three components: a *response variable* Y; a *structural model* consisting of one or more explanatory variables $X1$, $X2$, *etc.*; and an *error term*. Table 15.1 shows equivalent expressions. Each term in the model (response variable, explanatory variables, error) represents a vertical string (vector) of numbers. Consequently the symbolic expressions in Table 15.1 represent a series of *data equations* (see Box 15.1). A data equation is written for each value of the response variable. Data equations connect observations to a model and residual term.

In Table 15.2 the symbol $\underline{\circ}$ means "equal by observation." The model in Table 15.2 is the simplest possible: M is equal to a fixed value called the *expected value* β_o, which is the average value for the population, obtained by making an infinite number of measurements under the same conditions. Of course, we almost always do not know nor can we obtain the true value of β_o. Our best estimate of the expected value is the mean $\hat{\beta}_0 = 59\text{ g}$, computed from the data we have. Table 15.2 shows the data equations for this simple model, called the *null model* H_o.

The null model is not a particularly interesting one. We are more interested in a model with some biological content. A more interesting model might be the scaling of cod mass to its length, a scaling that measures nutritional state. Box 15.1 shows another example of the contrast of the null model with a more interesting model. The more interesting biological model is motivated by the question: Does the amount of algal growth on the walls of experimental tanks depend on the shape of the tank? Chen et al. (1997) carried out experiments in tanks that differed in volume and shape (length relative to radius) to work out how production dynamics and algal accumulation scale with tank volume, area, and height. Surprisingly, algal biomass appeared not to scale with wall

Table 15.1 Equivalent Expressions of the General Linear Model

Data	=	Model	+	Residual
Observed	=	Expected	+	Residual
Response	=	f(explanatory variables)	+	Residual
Y	=	f($X1 + X2 + \ldots$)	+	Error
Y	=	$\Sigma\, \beta_i X_i$	+	ε

The explanatory variables can be on a nominal type of scale (ANOVA), on a ratio type of scale (regression), or both (ANCOVA). The residuals are distributed normally.

Table 15.2 Data Equations for Measurement of the Mass of Three Juvenile Cod *Gadus morhua*

The null model is H_o : $M = 59$ g.

Data	=	Fitted values	+	Residual
M	=	mean(M)	+	ε
55 g	$\underline{\circ}$	59 g	+	−4 g
60 g	$\underline{\circ}$	59 g	+	+1 g
62 g	$\underline{\circ}$	59 g	+	+3 g

area in a simple 1:1 fashion. The ratio of biomass to wall area increased, going from small to large tanks. Biomass appeared to scale instead with tank volume. This generated the research hypothesis that biomass was a function of tank volume. The research hypothesis, in statistical jargon, is called the *alternative model* H_A.

Box 15.1 Analysis of Algal Biomass in Relation to Tank Volume. Data from Table 1 and Figure 5 in Chen et al. (1997).

Data

B = biomass on tank walls, mg chl*a*.
V = volume of tank, m^3

V	***B***
10.0	520
10.0	200
1.0	40
0.1	11
0.1	7

Data equations for H_o

An estimate of the mean: $\text{mean}(B) = n^{-1}\sum B_i = 155.5$

Data	**=**	***mean(H)***	**+**	***Residual***		
Data	=	$\hat{\beta}_o$	+	Residual	Residual^2	
520	=	155.5	+	364.5	132831	
200	=	155.5	+	44.5	1977	
40	=	155.5	+	−115.5	13349	
11	=	155.5	+	−144.8	20979	
7	=	155.5	+	−148.5	22064	
		H_o			191200	$= SS_{tot}$

Parameter estimates

An estimate of the change in B with volume:

$$\hat{\beta}_V = \frac{520 - 7}{10 - 0.1} = 51.28 \text{ mg chl}a \text{ m}^{-3}$$

General linear model format:

$$B - 155.54 = 51.28\,(V - 4.24)$$

Slope-intercept form:

$$B = -64.17 + 51.28\,V$$

Data equation for H_A

Nsp	=	β_o	+	$\beta_V \cdot V$	+	ε		
Data	=	mean(B)	+	$\hat{\beta}_v(V - \bar{V})$	+	Residual	Residual^2	
520	=	155.5	+	298.5	+	66.0	4354	
200	=	155.5	+	298.5	+	−254.0	64522	
40	=	155.5	+	−167.9	+	52.4	2741	
11	=	155.5	+	−214.5	+	69.7	4856	
7	=	155.5	+	−214.5	+	66.0	4354	
		H_o		H_A			80828	$= SS_{res}$

SS_{tot}	191200	Fit to the null model H_o
SS_{res}	80828	Fit to the alternative model H_A
ΔSS	110372	Improvement in fit due to straight line model

The first set of data equations in Box 15.1 is for the null model: that species number is adequately characterized by a single value, the mean, with only random fluctuation around that value. The null model is the "just chance" model that the variation around the mean value of the response variable is unrelated to the explanatory variable. The GLM here consists of three terms: the response variable, the mean (null model) term, and the error. The second set of data equations in Box 15.1 is for the alternative model: that species number increases in a straight line fashion with increasing area. In general the alternative model will be that the response variable is related in some way to the explanatory variable.

We are interested in the improvement in fit due to the explanatory variable, which in this case is tank volume. So we write the alternative model as two terms: the null model term plus the volume term. This GLM now consists of four terms: the response variable B, the null model term β_o, the area term $\beta_V \cdot V$, and the error ε. In Box 15.1 the GLM has been written in symbolic form with an equivalent computational form immediately below it. The latter form replaces the symbol β_V with V, the estimate of the rate of change in biomass with tank volume. The symbol V has been replaced by the differenced value $(V - \bar{V})$, which is used in computing the data equations. The differenced value forces the regression line through the mean value of the explanatory variable, which we can always estimate. The intercept is the value of the response variable at $V = 0$, but we avoid making an estimate at $V = 0$, because we often have no information on the response variable at this point. Instead, we combine our estimate of the slope with the estimates of the mean values of B and V to calculate the y-intercept (value of B at $V = 0$), as in Box 15.2.

Box 15.2 Parameter Estimates and Calculation of the Slope-Intercept Form of a Linear Model from the General Linear Model Form. Data from Box 15.1

The general linear model: $B - \beta_o = \beta_V(V - \bar{V}) + \text{Residual}$

The structural model is: $B = \beta_o - \beta_V\bar{V} + \beta_V V$

The slope-intercept form is: $B = \alpha + \beta_V V$

$$\hat{\beta}_o = \frac{1}{n}\sum B_i \qquad \hat{\beta}_o = 155.54 \qquad \bar{V} = \frac{1}{n}\sum V_i \qquad \bar{V} = 4.24\ \text{m}^3$$

Estimate from largest and smallest tank:

$$\hat{\beta}_V = \frac{520 - 7}{10 - 0.1} = 51.82\ \text{mgChl}a\ \text{m}^{-3} \qquad \hat{\alpha} = 155.54 - 51.82 \cdot (4.24)$$
$$\hat{\alpha} = -64.17\ \text{mgChl}a$$

Estimate from ratio of average squared deviations:

$$\hat{\beta}_V = \frac{\sqrt{\sum(B_i - \bar{B})^2}}{\sqrt{\sum(V_i - \bar{V})^2}} = \frac{\sqrt{191200}}{\sqrt{111.13}} = 41.48\ \text{mgChl}a\ \text{m}^{-3} \quad \hat{\alpha} = 155.54 - 41.48 \cdot 4.24$$
$$\hat{\alpha} = -20.33\ \text{mgChl}a$$

Estimate from ordinary least squares:

$$\hat{\beta}_V = \frac{\sum B_i V_i - n^{-1}\sum B_i \sum V_i}{\sum V_i^2 - n^{-1}\left(\sum V_i\right)^2} = \frac{7241.77 - (21.20)(777.70)/5}{201.02 - (21.20)^2/5} = 35.49\ \text{mgChl}a\ \text{m}^{-3}$$
$$\hat{\alpha} = 155.54 - 35.49 \cdot 4.24$$
$$\hat{\alpha} = 5.053\ \text{mgChl}a$$

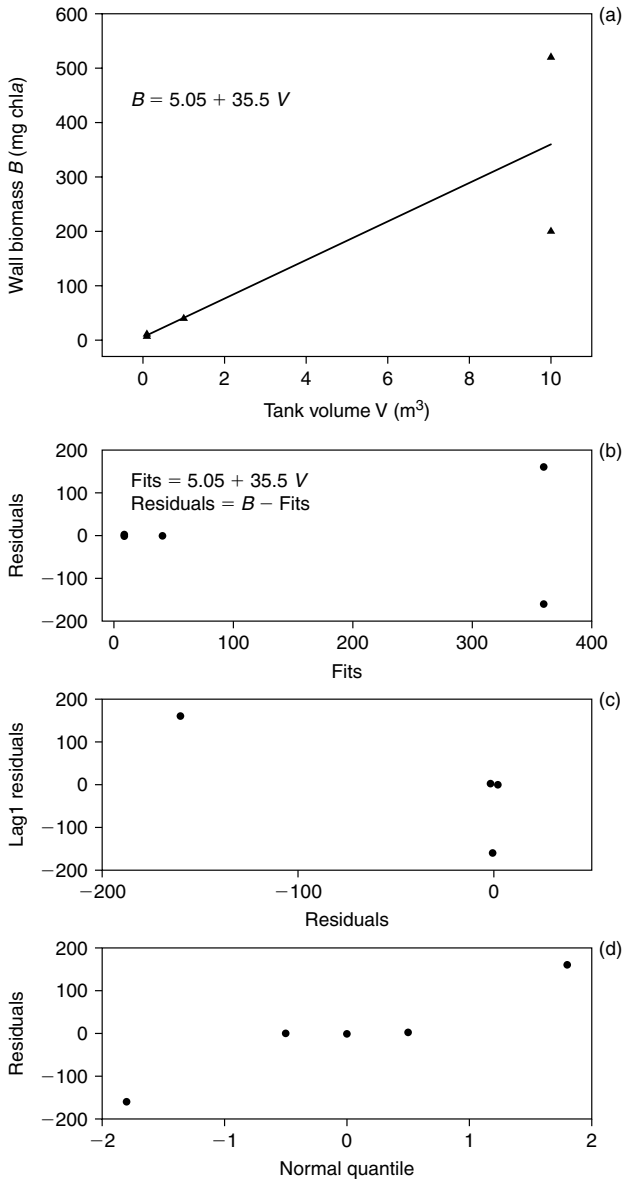


FIGURE 15.2 Graphical Evaluation of the Assumptions for Regression Analysis. Data from Chen et al. (1997); (a) Arithmetic axes display biomass as a linear function of tank volume; (b) Residuals, plotted against fitted values via least squares estimate of the exponent; (c) Residuals plotted against adjacent values; (d) Normal probability plot of residuals.

Once we have parameter estimates, we can plot the regression line (Figure 15.2a) and compute fitted and residual values to evaluate the model. The plot of the residuals versus the fitted values (Figure 15.2b) allows us to evaluate the structural part of the GLM ($Y = \Sigma\ \beta_i\ X_i$). If the structural model is correct, we expect to see the residuals form a band from left to right, with no evidence of bowl or arch patterns. If we do see a trend (a bowl, an arch, or both), the structural model is incorrect; the data are not consistent with a straight line. The model would need to be revised to some other functional form. In Figure 15.2b there is no obvious bowl or arch pattern, so a straight line model is acceptable.

ANOTHER LOOK AT SECTION 15.3

If you use regression, do you make it a practice to examine the residuals? Do you think your practice is typical? What are the pitfalls of using a straight line regression model, even though the fit is good, if the residuals show a curvilinear relation to the fitted values?

15.4 Interpreting Models and Their Parameters

Units and dimensions are typically not considered in the statistical analysis of ecological data. They should be. The parameters (means and slopes) that result from statistical analyses are parametric quantities with units and dimensions that depend on the units and dimensions of the measured variables being analyzed. They are not simply numbers, which is how they are often reported. A glance at the set of the three data equations for cod weights (Table 15.2) will reveal that the mean has the same units and dimensions as the response variable, which appears on the left side of the $\hat{=}$ sign. In a regression equation ($Y = \alpha + \beta_x X$ + residual), the intercept α must have the same units and dimensions as the response variable Y. The residual term must also have the same units and dimensions as the response variable Y. The regression coefficient β_x will have the same units and dimensions as the ratio Y/X in order for the equation to be dimensionally consistent. In the tank biomass example (Box 15.1), the slope β_V quantifies the increase in biomass (mg chl*a*) for each unit increase in tank volume. The rate of increase is 35 mg per cubic meter, or about 1 gram for 30 cubic meters of tank volume.

There are several reasons that GLM parameters should be recognized as scaled quantities rather than treated as simply numbers. First, the rules for operations on scaled quantities, which differ from those for numbers, apply to parameters estimated from data. Two means can be added only if they have the same units. The rules for rigid and elastic rescaling apply to parameters, a fact that is not evident if parameters are treated as mere numbers. Erroneous calculations can result if a parameter is treated as a number. For example, a regression coefficient that is an estimate of increase in biomass per cubic meter cannot be used to calculate the rate of increase at another scale (e.g., cubic centimeters) unless that coefficient is rescaled according to its units and dimensions, as in the previous example. Though this fact might seem obvious, it is too easily overlooked if the slope is reported as a number.

Parameters in a regression equation can have complicated units. An example is the regression of energy expenditure (on a log scale) against body mass (on a log scale), which results in a power law with fractional exponents. King (1974) used the following equation to relate daily energy expenditure ($\dot{E}$ = kcal/day) of free-living birds to body mass (M = kg):

$$\dot{E} = k \cdot M^{\beta} \tag{15.1}$$

King's regression estimates of the parameters were $\hat{k} = 179.8\ \text{kcal}\cdot\text{kg}^{-0.6687}\cdot\text{day}^{-1}$ and $\hat{\beta} = 0.6687$. These complicated units are necessary if the equation is to be of any use in making calculations about scaled quantities. The parameter k is a rigid rescaling factor (refer back to Chapter 5) that scales energy expenditure (with units of kcal/day) to a quantity with units of $\text{kg}^{-0.6687}$.

ANOTHER LOOK AT SECTION 15.4

Fee (1979) developed the idea that primary production in lakes (grams of carbon fixed per year) will scale with the area (m^2) of lake bottom shallow enough to be illuminated.

$$\frac{P}{P_{ref}} = \left(\frac{A}{A_{ref}}\right)^{\beta}$$

The exponent as estimated from data in Fee (1979) is $\hat{\beta} = 0.98$, which is indistinguishable from a value of 1. Rewrite the scaling relation as a scaling function (Equation 2.5; also, see Boxes 2.3 and 6.1). Work out the units of the two parameters that relate production to area and state an interpretation of each.

15.5 Goodness of Fit and Hypothesis Testing

Once we have a model consistent with the data, we can move on to questions of goodness of fit and hypothesis testing. The two most widely used *measures of goodness of fit* are both deviations; the smaller the deviation, the better the fit. These two measures are the information statistic G and the sum of the squared deviations (Boxes 10.2 and 15.1). For historical reasons, the G-statistic tends to be used for count data, whereas sums of squares are used for other analyses. However, statistical practice in ecology is moving toward increasing use of the information statistic (Burnham and Anderson, 1998), which can be used for any data, not just counts (McCullagh and Nelder, 1989). Box 15.3 shows an analysis of goodness of fit based on sums of squared deviations. For the parameter estimates based on ordinary least squares, the measure of deviation is the *total sum of squares* ($SS_{tot} = 191200$), which is the deviation of the data from the null model. The fit is better for the alternative model, where the additional term reduces the deviation measure, that is, reduces the *residual sum of squares* ($SS_{res} = 51207$). The observed *improvement in fit* is $SS_{model} = 139993$. This improvement in fit will be greater than with any other estimate of the slope because the criterion for this estimate was that it minimized SS_{res}. The ordinary least squares estimate is "optimal" in the sense that it gives the best possible improvement in fit. In statistical jargon this estimate maximizes the likelihood of having observed the data.

Though the least squares estimate is "optimal," we still must ask whether the improvement in fit is more than just chance. How much weight should we put on the conclusion that the fit is improved and hence that the research model is better than the null model? This is addressed by computing the probability of obtaining the observed improvement by chance alone. The improvement could be measured as the difference in the sums of squares, but the calculation of a p-value is more convenient if we measure the improvement as a *variance ratio*, the ratio of the improvement in fit to the residual variance. The *analysis of variance (ANOVA) table* displays the flow of computations leading to this variance ratio (Box 15.4). We begin with the sums of squares, which are partitioned according to the general linear model as shown in Box 15.4. We then move to a similar partitioning of the *degrees of freedom*, which take into account the number of parameters estimated from the data, relative to the number of observations. We begin with as many degrees of freedom as observations. We lose a degree of freedom for each

parameter estimated from the data. In the analysis of the tank biomass data, a degree of freedom was lost for the estimate of the mean value of B; another degree of freedom was lost for the estimate of the slope of the regression line.

Box 15.3 Improvement in Fit Due to Regression for Three Estimates of the Slope Parameter. Data from Boxes 15.1 and 15.2

The general linear model: $B - \beta_o = \beta_V(V - \bar{V}) + \text{Residual}$

The slope-intercept form is: $B = \alpha + \beta_V V$

Estimate from largest and smallest tank		SS_{tot}	191200
$\hat{\beta}_v = 51.8$		SS_{res}	−80828
$\hat{\alpha} = -64.17$	Improvement in fit	ΔSS	110372
Estimate from ratio of average squared deviations		SS_{tot}	191200
$\hat{\beta}_v = 41.48$		SS_{res}	−55190
$\hat{\alpha} = -20.33$	Improvement in fit	ΔSS	136010
Estimate from ordinary least squares		SS_{tot}	191200
$\hat{\beta}_v = 35.49$		SS_{res}	−51207
$\hat{\alpha} = 5.053$	Improvement in fit	ΔSS	139993

Once the degrees of freedom and *SS* are partitioned according to the model, they are gathered into the ANOVA table. Each term in the model is listed in the source column. Degrees of freedom and *SS* corresponding to each term in the model are listed in their respective columns. The computational flow continues to the next column to the right, where we compute the *mean squared deviations*, $MS = SS/df$. The *F-ratio* measures the improvement in fit as the ratio of variance due to the model term relative to the error term. This variance ratio measures the improvement in fit relative to the error. The last step is to compute the *p-value*, the probability of obtaining this variance ratio (improvement in fit) by chance alone. The *p*-value of a variance ratio is computed from the ratio of two chi-square distributions under the assumption that the numerator and denominator, being variances, are distributed as chi-square. As there are two chi-square distributions in any *F*-distribution, the degrees of freedom of the numerator and denominator must both be used to compute the *p*-value. Box 15.4 shows the Minitab calculation, which returns a value of 0.936. This means that 93.6% of the frequency distribution is less than the variance ratio of 8.20, and only 6.4% is greater. The chance of obtaining the observed *F*-ratio, if there is no relation of wall biomass to volume, is the *p*-value of 6.4%.

Software packages automatically generate *p*-values along with the ANOVA table. Nevertheless, direct calculation of *p*-values from the *F*-distribution, as shown in Box 15.4, is useful to know. For complex model designs, it can be hard to coax the package into computing the appropriate *F*-ratio and *p*-value. In these cases the *F*-ratio can be computed from the appropriate mean squares, followed by a computation of the *p*-value, as in Box 15.4.

Box 15.4 Model-Based Statistical Analysis of Biomass *B* in Relation to Tank Volume *V*. Data from Boxes 15.1 and 15.2

$\hat{\beta}_v = 35.49$ mg Chl*a* m^{-3}

Partition the sum of squares according to the model:

$\beta - \beta_o$	=	$\beta_V \cdot (V - \bar{V})$	+	Residual
SS_{tot}	=	SS_{model}	+	$SS_{residual}$
191200	=	139993	+	51207

Partition the degrees of freedom according to the model:

$\beta - \beta_o$	=	$\beta_V \cdot (V - \bar{V})$	+	Residual
df_{tot}	=	df_{model}	+	$df_{residual}$
$5 - 1$	=	1	+	3

ANOVA table:

Source	*df*	*SS*	*MS*	*F*	*p*
V	1	139993	139993	8.20	0.064
Residual	3	51207	17069		
Total	4	191200	*		

*MS_{tot} = var(*Nsp*)
var(*B*) = 191200/4
var(*B*) = 47800

Use *F*-distribution to compute *p* in Minitab:

```
MTB  > cdf 8.20;
SUBC > F 1 3.
     8.20    0.936
MTB > let k1 = 1 - 0.936
MTB > print k1
          0.064
```

Assumptions for *p*-value:

1. Σ residuals = 0 ? Yes Σ residuals $< 10^{-12}$
2. Residuals independent? ? Figure 15.2c
3. Residuals homogeneous? No. Figure 15.2b
4. Residuals normal? Yes? Figure 15.2d

The *p*-value is close to the criterion of significance of $\alpha = 5\%$ and the assumptions for the *p*-value are not met. A decision based on this *p*-value may well be incorrect.

The *p*-values computed from the *F*-distribution rest on four assumptions (Cochran and Cox, 1957; Scheffe, 1959; Seber, 1980). These *assumptions for* p-*values* arise from the pedigree of the *F*-distribution, which is two chi-square distributions. The first assumption is that the residuals sum to zero. With the least-square estimate (Box 15.2), the residuals automatically sum to zero (within rounding error).

The second assumption is that the residuals are independent. A simple diagnostic device is to plot each residual against a neighboring value. If the residuals are independent, the graph will show no pattern. Figure 15.2c shows the residuals plotted against a neighboring value. The plot suggest a negative relation due to the ordering of the data, where the two largest residuals are adjacent because they are both from tanks of the same volume. The autocorrelation of the residuals at several lags (see Box 10.2) can be computed, but this is no substitute for the sparseness of the information, which limits any conclusion about whether this assumption has been met.

The third assumption is that the residuals are homogeneous. It has long been known that violation of this assumption can strongly distort *P*-values (Cochran, 1947; Eisenhart, 1947; Green, 1979). The best diagnostic for this assumption is the plot of the residuals against fitted values (Figure 15.2b). Here we are looking at whether the residuals fall within a horizontal band of uniform vertical extent (residuals are homogeneous) or deviate from a band (residuals are heterogeneous). The most common form of heterogeneity is a cone expanding to the right. Other forms (rare) are cones expanding to the left or a spindle with greater heterogeneity near the middle of the plot. The residuals in Figure 15.2b are judged heterogeneous because there was strong pattern (cone) in the plot.

The fourth assumption for computing a *p*-value from an *F*-distribution (Box 15.3) is that the errors are normally distributed. Violations of this assumption usually have a less distorting effect than violations of the homogeneity assumption, so it is curious that in the ecological literature, the normality assumption draws more attention. The general linear model is known to perform well under a variety of deviations from normal errors if the sample size is moderate to large (Green, 1979; Draper and Smith, 1981). Violation of this assumption has little effect provided that the error distribution is not too skewed and the tails are well behaved (Hotelling, 1960). A simple diagnostic plot for this assumption is the histogram of the residuals. However, this plot is less useful at small sample sizes, when violations matter. A more useful diagnostic is a normal plot in which each residual is plotted against its expected value from a normal distribution. Normally distributed residuals fall along a straight line rising to the right in such plots. Residuals that depart from the normality assumption deviate from the line. The residuals for the scaling of biomass to tank volume (Figure 15.2d) suggest departure from normality, but as with the lagged residual plot, the information is too sparse to draw much of a conclusion. A statistic measuring fit to the normal distribution could be computed, but this again is no remedy for the sparseness of the information in the graph.

Some software packages encourage the egregious practice of checking assumptions before undertaking the analysis. This is dangerously wrong because the assumptions apply to the errors for the research model (as in the second set of data equations in Box 15.2), not to the distribution of the data before analysis. The practice of checking assumptions before analysis wastes an unconscionable amount of effort because it is not uncommon for residuals to be normally distributed, even though a plot of the response variable itself is not normally distributed around its mean value. Checking the data (response variable) for normality is voodoo statistics.

Another poor practice, encouraged by some statistical packages, is to use the machinery of hypothesis testing to check assumptions. This seems like a good idea because it provides

a criterion for violation of an assumption. But it is a dangerous practice because statistical tests of assumptions become more and more effective in detecting violations as the sample size increases and violations matter less and less. Conversely, statistical tests become less and less effective as sample size decreases and violations matter more and more (as in Box 15.4). Hypothesis testing identifies violations when they don't matter (large sample sizes), whereas it misses violations when they matter (small samples sizes). Hypothesis testing to check assumptions reliably leads in the wrong direction because violations become more detectable the less they matter. The general linear model is robust to violations of normality, and so statistical testing of this assumption is, in the words of Roger Green of the University of Western Ontario, like "going out in a rowboat to see if it is too windy to launch the *Queen Mary*."

The data from Chen et al. (1997) were chosen for presentation because they illustrate the conundrum presented by a *p*-value close to the 5% criterion, where violations matter because the sample size is small, where the small number of observations make it difficult to judge whether violations are substantial, but where there is evidence that the most important assumption (homogeneity) has been violated. If an assumption is violated and sample size is small enough that violation can distort the *p*-value, a number of remedies are available. One highly effective remedy is to retain the *F*-ratio but compute the *p*-value from a distribution that is free of the assumptions. This is called a *randomization test* because the *p*-value is computed from a distribution generated by randomizing the response variable with respect to the explanatory variable. Box 15.5 shows an example. We begin by generating an *F*-ratio from randomized data, generating many F-ratios by repeated randomization, and constructing a frequency distribution. From this we compute the *randomized* p-*value* for the *F*-ratio obtained from the unrandomized data.

Box 15.5 Analysis Algal Biomass in Relation to Tank Volume. Data from Table 1 and Figure 5 in Chen et al. (1997)

10.00	40
10.00	11
1.00	7
0.10	7
0.10	520
V	*B*

B = biomass on tank walls, mg chl*a*. Volume $V = m^3$

Sample the biomass values at random, as in the example at right.

Random sampling makes the null hypothesis true. Note that sampling with replacement will result in some values appearing more than once, others not appearing at all.

Compute a new F-ratio for the randomized data. ANOVA table:

Source	*df*	*SS*	*MS*	*F*
V	1	139993	139993	0.88
Residual	3	479583	159861	
Total	4	619576		

Repeat to obtain a large number of *F*-ratios. Compute the proportion of *F*-ratios that exceed $F = 8.20$. Out of 4000 randomization, 118 resulted in *F*-ratios exceeding $F = 8.20$. Hence p = 118/4000 = 2.95%.

This *p*-value is smaller than our 5% criterion for significance. $2.95\% = p < \alpha = 5\%$

This *p*-value is free of assumptions.

We conclude that the relation of biomass to tank volume is not due to chance.

To declare our decision, we compare this assumption-free p-value to our criterion of statistical significance, which is traditionally set at 5%. We conclude that biomass on the walls of the experimental tanks scaled in a linear fashion with tank volume. This completes the formal machinery of statistical hypothesis testing with the GLM.

ANOTHER LOOK AT SECTION 15.5

Introductory texts in statistics often present analysis of variance as a set of procedures. Consistent with this practice, the statistical results are often presented as an ANOVA table rather than as a model, as in Box 15.4. List the advantages and disadvantages of the following presentation of results, compared to the conventional ANOVA table.

	B	=	$\beta_V \cdot \mathrm{V}$	+	Residual
SS	191200		139993		51207
df	4		1		3
F-ratio			8.20		

15.6 Generic Recipe for Applying the General Linear Model

The general linear model is not part of the traditional undergraduate curriculum for biology students. However, it can be taught at this level, and in doing so there are many advantages for the student. First, students learn unifying concepts rather than a sequence of apparently unrelated procedures. Students can see the relationship of one test to another rather than having to memorize a set of special procedures. For example, ANCOVA can be presented as minor variants of the same model rather than as two separate procedures, one for comparing slopes and one for statistical control of a regression variable. Remedies for recurring problems (e.g., heterogeneous variances) are presented once rather than several times in different guises. The model-based approach means that students can learn general remedies instead of specific remedies peculiar to individual tests. The mechanics of analysis are presented once rather than as a different procedure for each test. Students are able to accomplish more with the general linear model than by learning statistics as a set of named procedures. For example, with this approach students can set up and execute the analysis of a response variable in relation to two categorical variables and a single regression variable. There is no name for this analysis, and hence it is outside any list of tests. This greater flexibility leads to better quantitative work in biology. The GLM is a way of thinking in quantitative terms using formal models that relate one quantity to another. The GLM approach takes slightly longer to present than any single technique such as regression, one-way ANOVA, two-way ANOVA, or the like. The GLM is, after all, more abstract and general than any

particular procedure. Students in biology at the fourth-year undergraduate or first-year graduate level readily grasp the approach and can successfully employ the following generic recipe (Table 15.3), even if they have limited backgrounds in mathematics or statistics.

The analysis of algal biomass in relation to tank volume (Boxes 15.1 and 15.3) illustrated most of the steps in the generic recipe (Table 15.3). The next two sections work through the generic recipe step by step. The first application evaluates a power law scaling function using the general linear model (Section 15.7.1). The second compares two scaling functions (Section 15.7.2).

Table 15.3 Generic Recipe for Statistical Analysis with the General Linear Model

1. Construct model: Begin with verbal and graphical model.
   Distinguish response from explanatory variables
   Assign symbols, state units and type of measurement scale for each.
   Write out statistical model.
2. Execute model:
   Place data in model format, code model statement.
   Compute fitted values from parameter estimates.
   Compute residuals and plot against fitted values.
3. Evaluate the model using residuals:
   If straight line inappropriate, revise the model (back to Step 1).
   If errors not homogeneous, consider using generalized linear model back to (Step 1).
   If *n* small, evaluate assumptions for using chi-square, *t,* or *F* distribution.
      Residuals homogeneous? (residual versus fit plot)
      Residuals independent? (plot residuals versus residuals at lag 1)
      Residuals normal? (histogram of residuals, quantile or normal score plot)
      If not met, empirical distribution (by randomization) may be necessary.
4. State population and whether the sample is representative.
5. Decide on mode of inference. Is hypothesis testing appropriate?
   If yes, go to Step 6; otherwise, skip to Step 10.
6. State H_o/H_A pair (some analyses may require several pairs).
   State test statistic, its distribution (*t* or *F*), and α, the tolerance of Type I error.
7. ANOVA: Partition *df* and *SS* according to model.
      Table source, *SS*, *df*, *MS*, *F*-ratio.
      Type I error (*p*-value) from distribution (*F* or *t*).
8. Recompute *p*-value if necessary.
   If assumptions are not met compute better *p*-value by randomization if:
   Sample small ($n < 30$) and if p near α.
9. Declare decision about model terms:
   If $p < \alpha$, then reject H_o and accept H_A
   If $p \geq \alpha$, then accept H_o and reject H_A
   Report conclusion with evidence: Either the ANOVA table or
   F-ratio (*df1*,*df2*) or *t*-statistics (*df*) and *p*-value (not α) for terms of interest.
10. Report and interpret parameters of biological interest (means, slopes) along with one measure of uncertainty (standard error, standard deviation, or confidence intervals).
    Use appropriate distribution (Step 8) to compute confidence limits.

15.6.1 Linearized Power Laws

To evaluate power law scaling functions within the framework of the general linear model, the standard approach is to linearize the power law by taking the logarithm of the response and explanatory variables:

$$\ln(Q) = \beta_o + \beta_x \ln(X) + \varepsilon \tag{15.2}$$

When we exponentiate both sides of this equation, we obtain a power law:

$$Q = e^{\beta o}\, X^{\beta x}\, e^{\varepsilon} \tag{15.3}$$

Application of the general linear model to a linearized power law is demonstrated by reanalysis of the biomass data from Chen et al. (1997).

The first step (Box 15.6) is to construct the model. With practice and experience, writing the model can be done directly. To reach this level of skill, we break the activity into preparatory steps: distinguishing response from explanatory variables, stating the research model about these variables in words, sketching the expected relation of variables as a graph, then writing the formal model. We start by distinguishing the response from explanatory variables. We are interested in the scaling of algal biomass to tank volume, which we can visualize as a plot of algal biomass (Y or vertical axis) against tank volume (X or horizontal axis). Biomass is the response variable, which we seek to scale to an explanatory variable, tank volume. When plotted on a logarithmic scale, the data fall nicely along a straight line (Figure 15.3a). Having stated our model verbally (biomass scales as tank volume), then graphically (Figure 15.3a), we proceed to writing out our concept as a linear model with response variable on the left and a series of terms on the right: a reference value (β_o the mean value of B) from which we will calculate the Y-intercept, an explanatory term (product of the explanatory variable $\ln V$ and the slope β_V), and the error term ε:

$$\ln B = \beta_o + \beta_V \cdot \ln V + \varepsilon \tag{15.4}$$

These components (response variable, reference value, one or more explanatory terms, and an error term) will appear in every general linear model.

The model format (Equation 15.4) will be our guide in executing the analysis in any statistical package (Box 15.6). Most statistical packages now have a general linear model routine that allows us to execute the analysis with a single command that estimates parameters and their standard errors, computes residual and fitted values, and produces diagnostic plots. The most useful plot will be the residual versus fitted values, which we will use to evaluate both the structural part of the model and the error distribution. Additional plots, notably a histogram of the residuals, a normal plot of the residuals, and a plot of residuals against their neighbors, are useful in evaluating the error distribution when the response variable is ordered in some way (high to low values) or when an explanatory variable is ordered in some systematic fashion.

Moving to Step 3, we use the diagnostic plots to decide whether the structural model is appropriate. The structural model is the response variable and all the terms except the error term. In this example the structural model is a linearized power law, which has

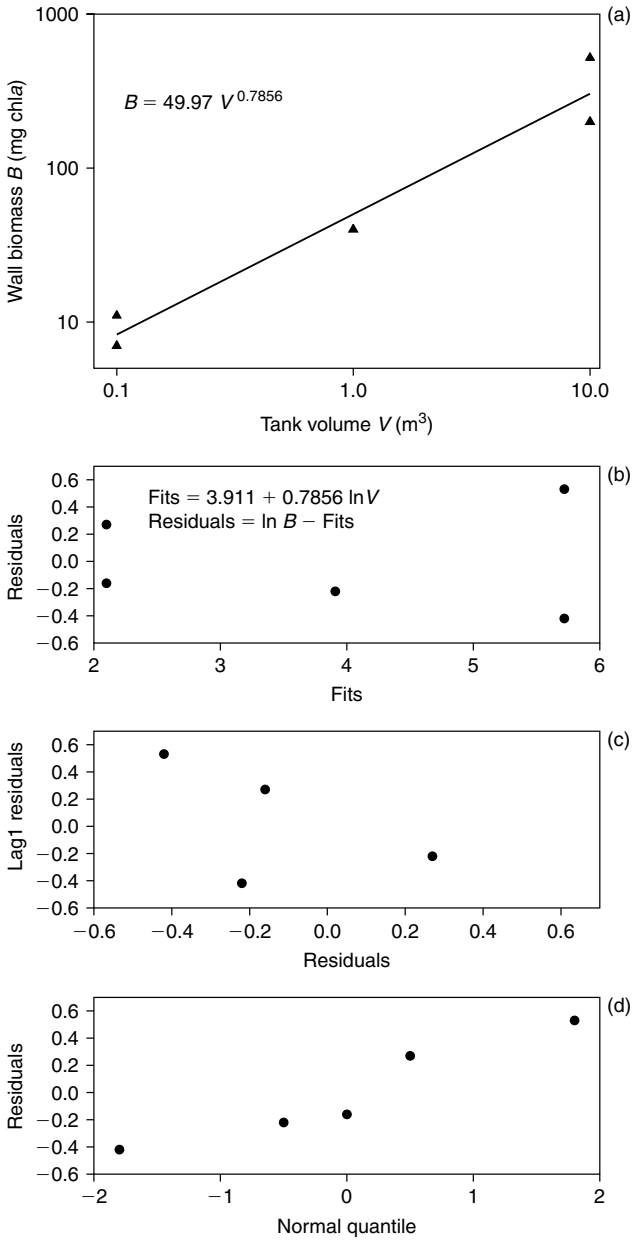


FIGURE 15.3 Scaling of Algal Biomass on Tank Walls to Tank Volume. Data from Chen et al. (1997); (a) Logarithmic axes display biomass as a power law function of tank volume; (b) Residuals, plotted against fitted values from linearized power law using least squares estimate of the exponent; (c) Residuals plotted against adjacent values; (d) Normal probability plot of residuals.

three terms (response variable ln B , reference value β_o, and explanatory term $\beta_V \cdot \ln V$). The residual versus fitted value plot (Figure 15.3b) shows no evidence of a bowl or arch pattern, indicating that the structural model (linearized power law) is an appropriate model. If a bowl or arch is present, revising the model should be considered (return to Step 1).

Box 15.6 Set-Up (Steps 1–3 in Table 15.3) of Analysis of Power Law Relation of Algal Biomass to Tank Volume. Data from Chen et al. (1997), in Box 15.1

1. Construct model:
   Response variable is logarithm of algal biomass $\ln B = \ln(B/1\,\text{mg Chl}a)$.
   Explanatory variable is logarithm of tank volume $\ln V = \ln(V/1\text{m}^3)$.
   Verbal model; algal biomass scales with tank volume according to a power law.
   Graphical model; plot of log of algal biomass against log of tank volume, Figure 15.3a.
   Write the linear model:
   Response = f(Explanatory) + residual
   $\ln B = \beta_o + \beta_V \cdot \ln V + \varepsilon$
2. Execution can be accomplished easily with the general linear model routine available in most statistical packages. The linear model (above) guides the use of these routines:

Model	$\ln B = \beta_o + \beta_V \cdot \ln V + \varepsilon$
SAS	Proc GLM; Model $\ln B = \ln V$;
Minitab	MTB > glm '$\ln B$' = '$\ln V$'; covariate '$\ln V$'.
SPlus(R)	lm $\ln B \sim \ln V$

   The routines estimate and report $\hat{\beta}_v = 0.78564$ $sterr = 0.09779$
   parameters with standard errors $\hat{\alpha}_v = 3.911$ $sterr = 0.2002$
   They use parameter estimates to obtain fitted values and residuals, then produce diagnostic plots, notably the residuals versus fitted values (Figure 15.3b).

3a. Evaluate the structural model with residual versus fit plot:
   No bowl or arch pattern evident (Figure 15.3b).
   Hence the structural model ($\ln B = \beta_o + \beta_V \cdot \ln V$) is acceptable.

3b. Evaluate assumptions concerning the error distribution:

a. Errors sum to zero?	Yes	$\Sigma\text{res} < 10^{-14}$
b. Errors independent?	?	See Figure 15.3c
c. Errors homogeneous?	Yes	See Figure 15.3b
d. Errors normal?	Yes	See Figure 15.3d

Once we have an appropriate structural model (this can require several returns to Step 1), we move on to evaluating the error distribution. The most important criterion is that the residuals be homogeneous with respect to the fitted values (uniform band from left to right of the plot). The heterogeneous residuals evident with the linear model of algal biomass versus tank volume (Figure 15.2b) are no longer evident for the linearized power law model (Figure 15.3b). Taking the logarithm of the response and explanatory variables substantially reduced the heterogeneity in the residuals, as is often the case. We continue the evaluation of the errors by asking whether the residuals are independent and normally distributed. There

are too few observations to be able to say much about independence of errors (Figure 15.3c). The normal plot (Figure 15.3d) is an improvement, because the residuals fall closer to a straight line than previously (Figure 15.2d).

Once we have adopted a structural model and evaluated the error assumptions, we move to the statistical analysis of uncertainty (Steps 4–10, Table 15.3). We begin (Step 4) by considering the population to which we will be inferring from our sample. In this example the mean values of the response and explanatory variable cannot be used to infer the true values of either of these two means because we did not take a random sample of tanks. However, our interest is in the relation of algal biomass to tank volume, which we visualize as a line that we can estimate regardless of whether the tanks we use are randomly chosen so as to represent mean values of biomass over hundreds of experiments. We are prepared to infer from the data points we have to the infinite number of points along the line that relates biomass to volume, from the smallest to largest tank. Because the degree of experimental control was rigorous and well reported (Chen et al., 1997), we are prepared to infer from these results to an infinite number of points described by a power law that applies to any tank within the size range the experiment.

Next (Step 5, Table 15.3, Box 15.7), we consider the mode of inference to be used. We have several choices. We can decide to use the machinery of hypothesis testing. This mode, with its emphasis on *p*-values and on a decision between the alternative and null hypothesis, is widely used in biology. It was developed for experimental work where a clear decision on a nominal scale (yes or no) is required. Hypothesis testing is not always appropriate and its use has come into question by prominent statisticians, as is evident from the quote at the beginning of this chapter. Alternatives to hypothesis testing are Bayesian inference (a topic that is beyond the scope of this book) and likelihood inference (reporting parameter estimates with confidence limits). Confidence limits are informative because we can examine hypotheses other than the null model prescribed by hypothesis testing. For example, in the case of linearized power laws we are often at least as interested in whether the response variable scales in a 1:1 fashion with the explanatory variable (exponent is equal to 1) as we are in whether there is a relation (exponent not equal to zero). Confidence limits are usually more useful than *p*-values in evaluating scaling and measurement relations. For the example at hand (algal biomass in relation to tank volume), it is appropriate to ask whether or not biomass depends on the volume of experimental tanks, so we will use hypothesis testing (Steps 6–9, Table 15.3) to illustrate its use.

Hypothesis testing rests on the concepts of Type I and Type II error in relation to null and alternative hypotheses (Table 15.4). The null hypothesis is that any pattern we see (such as the association of one variable with another) is due to chance. The alternative hypothesis is that the pattern we see is not a matter of chance. All is well when we accept a null or alternative hypothesis that is true (Table 15.4). But if we reject the null (just chance) hypothesis and it is true, we have made a *Type I error*. If we accept the null hypothesis and it is false, we have made a *Type II error*. Hypothesis testing focuses on Type I errors, which we can calculate from a distribution (chi-square, *t,* or *F* distribution, or distribution via randomization). If the *p*-value (Type I error) is small (conventionally, small means 5%), we reject the null hypothesis, knowing that we will be wrong 5% of the time. Students who have learned a collection of statistical tests are often surprised to learn that in fact all we have done is to reject the "just-chance hypothesis" at a fixed level of error.

Box 15.7 Analysis of Uncertainty (Steps 4–9 in Table 15.3) for the Analysis of Algal Biomass in Relation to Tank Size. Data from Box 15.6

4. The population is taken to be all possible measurements on similarly constructed tanks within the size range of the sample.
5. The mode of inference will be hypothesis testing.
6. State research hypothesis. Algal biomass scales with tank volume, hence $\beta_V \neq 0$.
   State null hypothesis. Biomass does not scale with volume, hence $\beta_V = 0$.
   State statistic. The *F*-statistic from analysis of variance in $\ln B$.
   State tolerance for Type I error of accepting a relation when in fact there is none.
   α set at 5%, the traditional level in biology.
7. Calculate variance, partition df and SS according to model, place in ANOVA table.

$\ln B$	$= \alpha +$	$\beta_V \cdot \ln V$	$+$	ε	Model
4	$=$	1	$+$	3	df
13.70	$=$	13.09	$+$	0.61	*SS*

Compute $MS = SS/df$
$F = MS_{lnV}/MS_{residual}$

Source	*df*	*SS*	.*MS*	*F*
$\ln V$	1	13.09	13.09	64.54
residual	3	0.61	0.203	
Total	4	13.70		

Calculate Type I error from *F*-distribution. $p < 0.004$

```
MTB > cdf 64.54;
SUBC> f 1 3.
   64.54  0.996
```

8. Assumptions met, no need to recompute *p*-value.
9. Declare decision. Reject the null hypothesis, accept alternative hypothesis of relation of algal biomass to tank volume. $F_{1,3} = 64.54$, $p = 0.004$
10. Report parameters with measure of uncertainty. $\hat{\beta}_v = 0.78564$

$$st.err = \sqrt{\frac{MS_{err}}{\sum(\ln V - \overline{\ln V})^2}} \qquad st.err = \sqrt{\frac{0.203}{21.21}} \qquad st.err = 0.09779$$

Hypothesis testing centers on the null hypothesis, but because we are usually more interested in the alternative hypothesis, it is easier to start with this (Step 6, Table 15.3, Box 15.7). The research hypothesis is then inverted to the null hypothesis. To complete the setup, we decide on the test statistic and the distribution to be used to compute the

Table 15.4 Type I and II Statistical Errors in Relation to the Null H_o and Alternative H_A Models

	Decision	
	Accept H_o (Reject H_A)	**Reject H_o (Accept H_A)**
H_o true	Ok	Type II error
H_A true	Type I error	Ok

p-value. The criterion for statistical significance is set before undertaking the analysis, to avoid any adjusting of the criterion to the outcome (e.g., taking 'significant' to mean 6% if the *p*-value is 6%).

Once the conditions for the test are established, the next step is to compute the test statistic and *p*-value. For the general linear model, this means computing the measure of improvement in fit due to a term in the model (Step 7, Box 15.7). The analysis of variance table, widely used in biology, is a general linear model displayed vertically. Each term is listed on its own line with its sum of squares and degrees of freedom. Computation then proceeds from left to right in the table. The mean square for a term is computed as $MS = SS/df$ for explanatory and error terms. The *F*-ratio is the ratio of appropriate mean squares, usually the mean square for an explanatory term relative to the mean square for the error term. The *p*-value is computed from the *F*-ratio, given the degrees of freedom assigned to the numerator and denominator mean squares that make up the *F*-ratio.

Before declaring a decision based on the *p*-value delivered in the ANOVA table, we review the information we have and decide whether the *p*-value needs to be recomputed by randomization (Step 8, Table 15.3, Box 15.7). This step calls for judgment, just like Step 3 (evaluate model assumptions). In making such judgments we take into account several factors. Violations are a matter of degree. Their distorting effect on the *p*-value diminishes as sample size rises. Recomputation is not going to change the *p*-value much unless the sample size is small and violations are serious, especially the violation of the homogeneity assumption. Recomputation is not going to change the decision if the *p*-value is far from 5%, even if there are violations. As a rule of thumb, violations have little effect on *p*-values if sample sizes exceed 100 and rarely have much effect unless samples sizes fall below 30. As another rule of thumb, randomization rarely changes a *p*-value by more than a factor of five. That is, randomized *p*-values are usually within a factor of 5 of the *p*-value from the statistical distribution (*t, F,* or chi-square). Consequently, if the *p*-value is less than 0.01 or greater than 0.25, the decision at 5% will usually go unchanged despite the better *p*-value. Violations need to be dramatic to have an effect on the *p*-value computation. The most frequent source of distortion is substantial heterogeneity, particularly that due to a few outliers when sample size is small. Judgment in deciding whether to recompute a *p*-value comes with experience in diagnosing violations and then recomputing a *p*-value for serious violations. The analysis of algal biomass in relation to tank volume was chosen because it was a good candidate for a randomization test. The sample size was small ($n = 5$), the *p*-value was close to 5%, and the violation of the heterogeneity assumption was evident (Figure 15.2b).

ANOTHER LOOK AT SECTION 15.6.1

Have you ever regressed the logarithm of one variable against the logarithm of another? If you have, write the function that relates the log of the response variable to the log of the explanatory variable. Then rewrite the expression as a power law. If you have not, write a function relating the log of catch/effort to the log of effort. Then rewrite this expression as a power law.

15.6.2 Comparing Regression Lines (ANCOVA)

The next application of the general linear model compares two scaling functions expressed as straight lines. The data are from Fee (1979). Does annual primary production in lakes scale with benthic epilimnion area A, the area of bottom shallow enough to be illuminated? Does the scaling of production to epilimnion area A in lakes treated with nutrients differ from untreated lakes? The analysis demonstrates the use of a categorical variable (fertilizer treatment yes or no) within the framework of the general linear model. The analysis also demonstrates the logic of interaction terms, which are examined before the main effects (which in this case are treatment and epilimnion area). This mixture of regression and categorical explanatory variables is called *analysis of covariance* (ANCOVA). Its primary use in ecological scaling is comparing scaling coefficients.

Figure 15.4a shows the data. Box 15.8 sets up the analysis (Steps 1–3, Table 15.3). In this analysis there are two explanatory variables. One is categorical (Treatment = yes or no); the other is on a ratio type of scale (epilimnion area). The ratio scale variable appears in the model as a term that is the product of the variable times a parameter. Because lake production and area are on a log scale, the parameter β_A is the exponent that scales production to area. The categorical variable also appears as a term that is the product of the variable times a set of parameters. The parameter β_{Tr} represents the difference in production between the mean value of ln(production) of treated lakes and the mean value of ln(production) of lakes in the reference category. The reference category can be the mean for all lakes.

$$\begin{aligned} \beta_o &= \text{mean}(\ln(\dot{M})) \\ \beta_o + \beta_{Tr} &= \text{mean}(\ln(\dot{M}_{Tr=Yes})) \end{aligned} \tag{15.5}$$

Alternatively, the reference category can be just one category, such as the mean for untreated lakes:

$$\begin{aligned} \beta_o &= \text{mean}(\ln(\dot{M}_{Tr=No})) \\ \beta_o + \beta_{Tr} &= \text{mean}(\ln(\dot{M}_{Tr=Yes})) \end{aligned} \tag{15.6}$$

The difference in notation follows from differences in the way that categories (in this case, two) are coded by the package used to implement the analysis. If the package codes Tr = No as 0 and Tr = Yes as 1, the result is Equation 15.6. If the package codes Tr = No as −1 and Tr = Y as +1, the codes add to zero and the result is Equation 15.5.

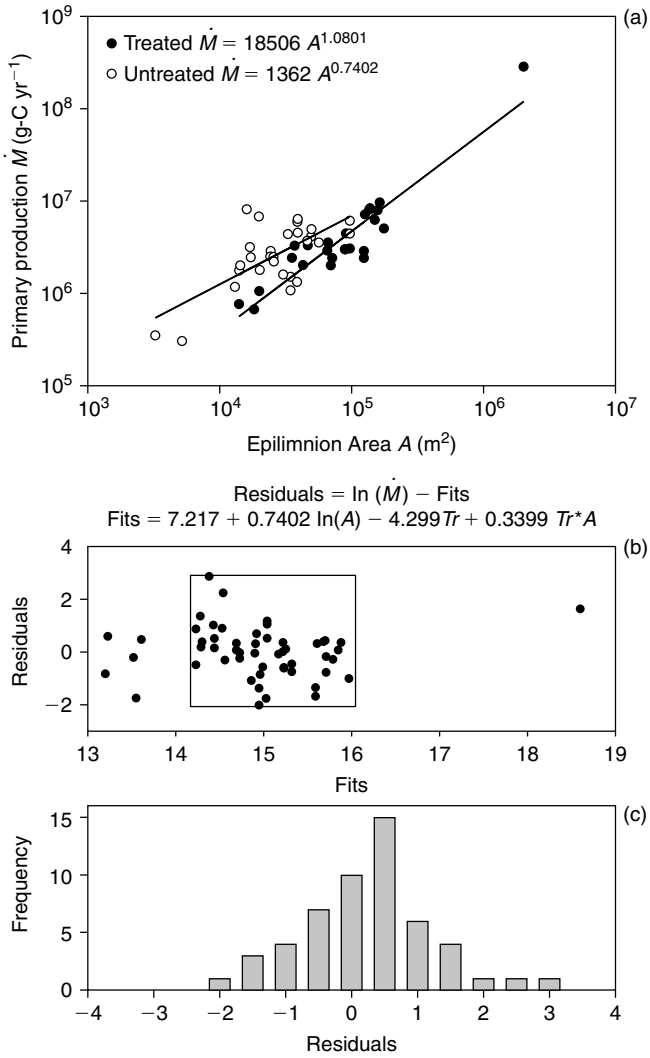


FIGURE 15.4 Scaling of Primary Production to Area of Epilimnion in Fertilized and Unfertilized Lakes. Data from Fee (1979); (a) Logarithmic axes display production as a power law function of area with line estimated by least squares regression in 19 lakes; (b) Residuals plotted against fitted values from analysis of 19 lakes; (c) Histogram of residuals.

In addition to a term for each variable, the model includes an *interaction term*. This term expresses the interactive effects of the explanatory variables on the response variable. In this analysis, the interaction term is the product of a categorical and a ratio scale (regression) variable. It measures the heterogeneity of slopes (Figure 15.4a). The more divergent the slopes across categories, the larger the interaction term. The term is written as the product of its two component variables and a parameter $\beta_{A \cdot Tr}$, which represents deviations in slope from the slope in the reference category β_A. If the reference category is all lakes (Equation 15.5), we have:

$$\beta_A = \text{regression of } \ln(\dot{M}) \text{ on } \ln A$$

$$\beta_A + \beta_{Tr^*A} = \beta_{Tr=Yes} \tag{15.7}$$

If reference category is one of the categories (Equation 15.6), we have:

$$\begin{aligned} \beta_A &= \beta_{Tr=Yes} \\ \beta_A + \beta_{Tr*A} &= \beta_{Tr=No} \end{aligned} \tag{15.8}$$

Box 15.8 Model Identification (Steps 1–3 in Table 15.3) for Analysis of Primary Production $\dot{M}$ Scaled to Epilimnion Area A in Treated and Untreated Lakes. Data are from the Experimental Lakes Area in Ontario (Fee, 1979).

1. Construct model:
   Response variable is logarithm of primary production
   $\ln \dot{M} = \ln(\dot{M}/1\,\text{g-C yr}^{-1})$.
   Explanatory variable is logarithm of epilimnion area $\ln A = \ln(A/1\text{m}^2)$.
   Explanatory variable is Treatment at two levels (Tr = Yes or no)
   Write the linear model.

$$\begin{aligned} \text{Response} &= \text{f(Explanatory)} + \text{residual} \\ \ln \dot{M} &= \beta_o + \beta_A \cdot \ln A + \beta_{Tr} \cdot Tr + \beta_{A \cdot Tr} \cdot \ln A \cdot Tr + \varepsilon \end{aligned}$$

2. Execute analysis. The linear model (above) guides the use of GLM routines.

SAS	Proc GLM; Class Tr; Model lnM = lnA Tr lnA*Tr;
Minitab	MTB > glm 'lnM' = 'lnA' 'Tr' 'lnA'*'Tr'; covariate 'lnA'.
SPlus	lm lnM ~ lnA + Tr + lnA:Tr

   The routines estimate parameters, compute fitted values and residuals, then produce diagnostic plots.
3. Evaluate the structural model with residual versus fit plot. The outlier evident in this plot tilts the regression line such that most of the lakes do not follow the regression (data points inside the box in Figure 15.4b). The lake responsible is much larger than the other lakes (Lake 228/next largest: surface area = 1677 ha/56.1 ha = 30). There is no information from a control (untreated) lake of comparable size, so lake 228 was dropped from the comparison of treated and untreated lakes (Figure 15.5a). The resulting scaling functions will be more reliable (less influenced by a single lake) but restricted to a smaller range in sizes (largest/smallest = 56.1 ha/1.7 ha = 33).
2. Return to step 2.Execute analysis without Lake 228.
3. Evaluate the structural model. No bowl or arch is evident (Figure 15.5b). Evaluate error term. Residuals are homogeneous (Figure 15.5b) and normal (Figure 15.5c).

Once the model is written (Step 1, Box 15.8) the calculations are readily executed (Step 2) in any statistical package with a GLM command. The command generates parameter estimates, standard errors, and diagnostic plots to evaluate the model (Step 3). The residual versus fitted value graph (Figure 15.4b) shows a pattern where most of the residuals (inside the box drawn round them) trend downward from left to

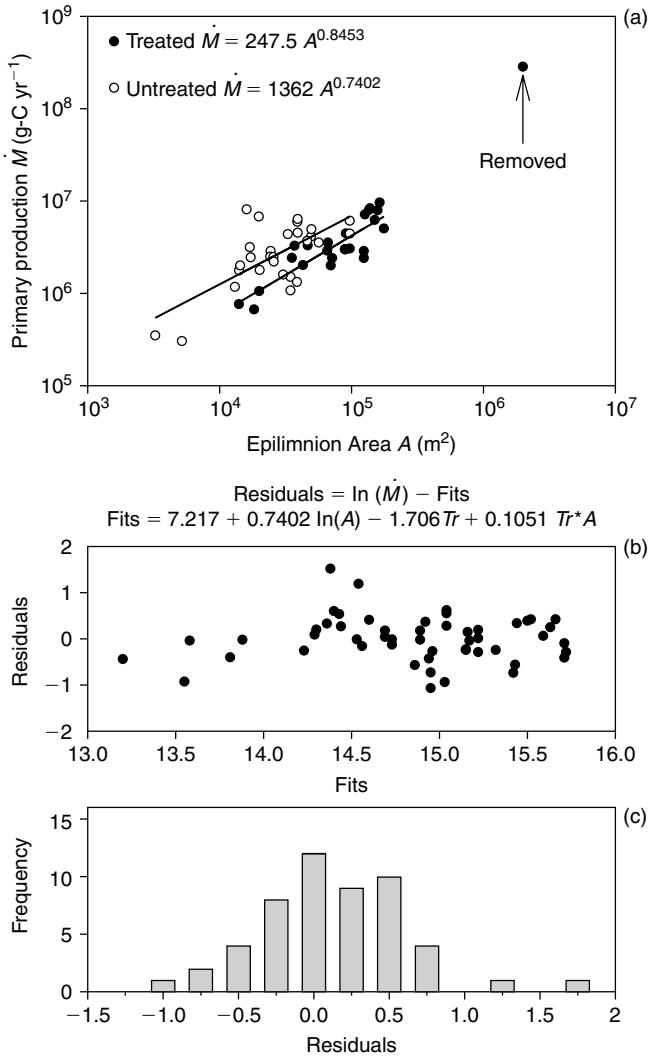


FIGURE 15.5 Scaling of Primary Production to Area of Epilimnion in Fertilized and Unfertilized Lakes. Data from Fee (1979); (a) Logarithmic axes display production as a power law function of area via least squares regression in 18 lakes; (b) Residuals plotted against fitted values from analysis of 18 lakes; (c) Histogram of residuals.

right. This is the result of the undue influence of a single large lake with an epilimnion area 30 times that of the next largest lake. This single value, because it is far from the mean value of production and area, tilts the regression line upward on the right side in Figure 15.4a. Because of this tilting of the line, the data points (inside the box) do not follow the line for treated lakes. Because lake 228 differs substantially from the others in epilimnion area and because there is no untreated lake of comparable size, the lake was dropped from the analysis. It is interesting to note that with lake 228 in the analysis, the difference in exponents (β_{Tr} = 1.0801–0.74021 = 0.3399) is almost significant ($F_{1,49}$ = 3.71, p = 0.0599), where the assumptions of homogeneous (Figure 15.4b) and normal (Figure 15.4c) errors were met. However, the conclusion is suspect because of the undue influence of a single lake. Dropping this lake strengthens the analysis but at the same time restricts the conclusions to a small range of lake sizes.

The analysis is executed again (return to Step 2). This gives revised estimates of the scaling parameters (Figure 15.5a), with new ANOVA table and diagnostic plots (Figures 15.5b,c). The data (with one lake dropped) are consistent with the linearized power law, as indicated by absence of bowls or arches in the residual versus fit plot (Figure 15.5b). The residuals are homogeneous (Figure 15.5b) and acceptably normal (Figure 15.5c).

Box 15.9 Setting Up the Analysis of Uncertainty Via Hypothesis Testing (Steps 4-6 in Table 15.3) for Analysis of Primary Production Scaled to Epilimnion Area in Treated and Untreated Lakes. Data from Fee (1979)

4. Sample consists of 52 measurement of annual production in 16 lakes (10 treated, 5 untreated, 1 treated in 1 year and untreated in 2 years, lake 228 omitted). Statistical inference cannot be made to the population of all lakes in the Experimental Lakes Area in Ontario. Inference is to a hypothetical population, all possible measurements with this measurement protocol. The conclusion about this hypothetical population can be used to form expectations about other lakes with surface area from 1.7 to 56.1 ha under the assumption that the lakes measured are representative of other lakes.
5. Mode of inference will be hypothesis testing because we seek a decision on whether the scaling exponent differs been treated and untreated lakes.
6. State research hypotheses:
   Production depends on area of epilimnion, hence $\beta_A \neq 0$.
   Production depends on treatment, hence mean($\dot{M}_{Tr=Yes}$) $\neq$ mean($\dot{M}_{Tr=No}$).
   Equivalently, $\beta_{Tr} \neq 0$
   Relation of production to area depends on treatment,
   hence $\beta_{A,\ Tr=Yes} \neq \beta_{A,\ Tr=No}$.
   That is, the regression coefficients differ for treated and untreated lakes.

   State null hypotheses:
   Production does not depend on area, hence $\beta_A = 0$.
   Production does not depends on treatment,
   hence mean($\dot{M}_{Tr=Yes}$) = mean($\dot{M}_{Tr=No}$).
   Equivalently, $\beta_{Tr} = 0$
   Relation of production to area does not depend on treatment,
   hence $\beta_{A,\ Tr=Yes} = \beta_{A,\ Tr=No}$.
   Tolerance for Type I error set at $\alpha = 5\%$, the traditional level in biology.

The target of inference (Step 4, Table 15.3, Box 15.9) deserves reconsideration in light of the removal of the largest lake (Box 15.8). Inference will now be restricted to lakes within a narrower size range. Dropping the outlier has a cost in that this restricts the target of inference and hence what we can say from the analysis. Moving to Step 5, we will use hypothesis testing, given the goal of declaring whether treatment alters the scaling of production to epilimnion area. The model is our guide to writing the

alternative/null hypothesis pair (Step 6). There is one pair for each of the three explanatory terms. If the interaction term proves significant, hypotheses concerning the scaling exponent or the average difference between treated and untreated lakes β_{Tr} will not be tested. It would be inconsistent to test for these overall effects if interaction is present and hence the effect of one explanatory variable depends on the other explanatory variable.

Box 15.10 Analysis of Uncertainty (Steps 7–10 in Table 15.3) for Analysis of Primary Production Scaled to Epilimnion Area in Treated and Untreated Lakes. Data from Fee (1979), Continued from Boxes 15.8 and 15.9

7. Calculate variance, partition *df* and *SS* according to model.

$\ln \dot{M} - \beta_o$	=	$\beta_A \cdot \ln A$	+	$\beta_{Tr} \cdot Tr$	+	$\beta_{A \cdot Tr} \cdot \ln A \cdot Tr$	+	ε
51	=	1	+	1	+	1	+	48 *df*
29.982	=	14.651	+	2.878	+	0.07445	+	12.379 Sequential *SS*

List sources of variance with *df* and *SS*.
Compute *MS* = *SS*/*df* for each term in the model.
Compute variance ratios relative to residual variance.

Source	***df***	**Adj *SS***	**Adj *MS***	***F***	***p***
ln*A*	1	8.1785	8.1785	31.71	0.0000009
Tr	1	0.1689	0.1689	0.65	0.422
ln*A*·*Tr*	1	0.07445	0.07445	0.29	0.594
Residual	48	12.379	0.2579		
Total	51	*	<----Adjusted *SS* do not sum to 29.98		

8. Recomputing *p*-value not judged necessary: assumptions met.
9. Declare decision. This is done sequentially, beginning with the interaction term. The interaction term is not significant ($F_{1,48} = 0.289$, p = 0.59). The scaling exponents for treated and untreated lakes are statistically indistinguishable. Consequently, continue upwards in table, to hypotheses concerning main effects. Primary production depends on epilimnion area ($F_{1,48} = 31.71$, $p < 0.001$) adjusted for ln*A* and ln*A*·*Tr*. Production does not depend on treatment ($F_{1,48} = 0.65$, p = 0.42) adjusted for ln*A* and ln*A*·*Tr*. Production does depend on treatment ($F_{1,49} = 2.878/(12.379/49) = 11.39$, $p = 0.0015$) adjusted for ln*A*, without considering ln *A*·*Tr*. This *F*-ratio is adjusted for ln*A* only because it was calculated from sequential *SS* above.
10. Estimate of scaling exponent:
$\hat{\beta}_A = 0.7875$ *st. err.* = 0.0965 CI = 0.5935 to 0.9815 adjusted for treatment difference
$\hat{\beta}_A = 0.6002$ *st. err.* = 0.0867 CI = 0.426 to 0.774 not adjusted
The 95% confidence limits exclude a 1:1 scaling ($\hat{\beta}_A = 1$) of primary production to epilimnion area.

GLM commands automatically partition the degrees of freedom according to the number of terms in the model, following the rules in Table 15.5. Degrees of freedom are assigned only to terms in the model. For example, if the interaction term B x R were omitted from the model in Table 15.5, the residual term would have $df = 25$ rather than $df = 23$.

GLM commands also partition the total sums of squares according to the model (Step 7, Box 15.10). Partitioning can occur in several ways. One common method is to partition the total *SS* in the order in which the terms occur in the model. With this sequential method the variability assigned to each term depends on where it occurs in the list if the explanatory variables are correlated. Packages will also produce an adjusted *SS*, which is the variability for each term if it were to occur last in the list. This adjusts for any correlation among explanatory variables. Box 15.10 shows the adjusted *SS* and *MS*.

Once the degrees of freedom and *SS* are partitioned according to the model they are tabled and the mean squares are calculated from the degrees of freedom in the usual way. The *F*-ratios are taken relative to the error term. GLM commands use the *F*-distribution to compute *p*-values for each *F*-ratio. The *p*-value is the chance of obtaining the observed improvement in fit (measured as an *F*-ratio) if the null hypothesis is true. In the analysis in Box 15.10, the interaction term was far from significant, leading to the conclusion that the same scaling exponent applies to treated and untreated lakes. In the absence of interactive effects, hypotheses concerning the average effects of treatment ($\hat{\beta}_{Tr}$) and the overall scaling with area ($\hat{\beta}_A$) are then examined. The treatment effect was significant when the effects of area are controlled statistically (ln$A$ included in the model). The scaling of primary production with area did depend on epilimnion area ($\hat{\beta}_A \neq 0$). Of more interest is that the 95% confidence limits for the exponent (Box 15.10) exclude $\beta_A = 1$. Consequently, we can exclude a 1:1 scaling of production to epilimnion area. It is worth noting that if the residuals had not been examined and the statistical output with lake 228 present were accepted on face value, a different conclusion (treated and untreated lakes differ in scaling of production to epilimnion area) would be the result.

Table 15.5 Computing Degrees of Freedom of Terms in a General Linear Model

n = number of observations; example based on $n = 3 \times 4 \times 6$ observations.

Source	*Df*	Example	Result
Response variable	$n - 1$	(72 − 1)	71
Explanatory variables			
Regression variable *R*	1	1	
Categorical variable *A*, with *na* levels	$na - 1$	(4 − 1)	
Categorical variable *B*, with *nb* levels	$nb - 1$	(3 − 1)	
Interaction term, $A \times B$	$(na - 1)(nb - 1)$	(3)(2)	
Interaction term, $A \times R$	$(na - 1)(1)$	(3)(1)	
Interaction term, $B \times R$	$(nb - 1)(1)$	(2)(1)	
Interaction term, $A \times B \times R$	$(na - 1)(nb - 1)(1)$	(3)(2)(1)	
Σdf_{model}			23
Residual	$n - 1 - \Sigma df_{model}$	72 − 1−23	48

ANOTHER LOOK AT SECTION 15.6.2

For a variable of interest to you, sketch a graph of the line relating this variable to an explanatory variable that you think influences your variable. Next, sketch the relation in each of several categories (before/after, male/female, and so on). Define symbols for each of the three variables, then write a general linear model comparing regression lines.

15.6.3 Compound Scaling with Multiple Regression

A widespread practice in the literature is to scale one variable to another by division before undertaking analysis. For example, in an experiment with fish, it is common practice to divide the mass of the fish by a measure of size (length or length^3 are typically used) to correct for size differences before undertaking further analysis. Another example is from Ryder (1965), who adjusted annual fish catch $\dot{M}$ (kg yr^{-1}) for lake area A (ha) before undertaking analysis in relation to his morphoedaphic index (MEI), which is the ratio of total dissolved solids TDS (ppm) to lake depth z (meters). The MEI was constructed as a composite measure of lake morphology (as indicated by depth) and trophic status (low TDS in oligotrophic lakes). Taking the logarithm of both ratios, Ryder (1965) obtained a scaling exponent of 0.4461. Conversion of Ryder's data to meters and hectares results in slightly different parameter estimates for the linearized power law relating catch per unit area to the MEI:

$$\frac{\dot{M}}{A} = e^{0.3215}\left(\frac{TDS}{z}\right)^{0.4468} \tag{15.9a}$$

The scaling relation incorporates more than two variables and thus is aptly described as a *compound scaling relation*. These arise when a variable is adjusted for another by taking a ratio before estimating the scaling exponent.

Compound scaling relations assume that the two components of a ratio have the same exponent with opposite sign. In Equation 15.9a, the assumption is that $\dot{M}$ and A have the same exponent, as follows:

$$\dot{M}^1 = e^{0.3215} A^1 \left(\frac{TDS}{z}\right)^{0.4468} \tag{15.9b}$$

If we check this assumption with Ryder's data, we find that the evidence is otherwise. The exponent that scales annual fish catch to lake area is estimated from the data as $\hat{\beta}_A = 0.84$, with a standard error of 0.0430 on 21 degrees of freedom and hence confidence limits of 0.751 to 0.929. These limits exclude the null hypothesis ($\beta = 0$), but of more importance, they exclude the 1:1 scaling of catch with area ($\beta = 1$) assumed in Equation 15.9a. A 1:1 scaling is assumed for the components of a ratio on the right side of a compound scaling relation, as in Equation 15.9a. That is, the exponent is of the same magnitude and opposite sign:

$$\dot{M}^1 = e^{0.3215} A^1 \, TDS^{0.4468} \, z^{-0.4468} \tag{15.9c}$$

Analysis of Ryder's data shows that the exponent that scales dissolved solids to depth is −0.0551 with a standard error of 0.1277, hence confidence intervals of −0.321 to 0.210. These limits include the null hypothesis ($\beta = 0$), but of far more relevance, they exclude $\beta = -0.4468$. Consequently, we should not be surprised if the scaling of annual catch to MEI differs from the scaling of catch to each of the components of the MEI. Catch may scale with one component of a ratio, not with the other, and as a result catch may scale with MEI when in fact only one component of the explanatory variable is responsible.

The scaling of catch with area assumed in Equation 15.9a is incorrect, and the scaling of catch with MEI may be due to only one component of the MEI ratio. To correct the scaling with area and to obtain more information about the scaling of catch with the components of the morphoedaphic index, we rewrite the compound scaling relation by separating the components:

$$\dot{M} = e^{\alpha} A^{\beta_A} TDS^{\beta_{TDS}} z^{\beta_Z} \tag{15.10}$$

Compound scaling relations are evaluated with *multiple regression* (more than one independent variable). Before analysis, we examine the degree of correlation of the explanatory variables. If the components are correlated, the scaling exponents as estimated by multiple regression (the partial regression coefficients) will differ from those estimated by regression of catch on each component separately (simple regression coefficients). Of particular concern are built-in correlations, which can arise in any of several ways, are avoidable, and inevitably produce estimates of exponents that are inconsistent with each other and with what we may already know about the scaling of one variable to another. One way that built-in correlations arise is construction of a new variable from two or more measured variables. If, for example, we use the measured value of lake surface area A to compute the area in the epilimnion A_e, there will be a built-in correlation between surface area and epilminion area. This in turn will influence our estimates of how some other quantity, such as catch, scales with the two correlated variables. Another source of built-in correlation is discoverable by dimensional analysis. We expect area and depth of a lake to be correlated in dimensional grounds; the surface area of an object such as a lake will scale with depth, even though the scaling of area with depth may not be as the square of depth, as in Euclidean lakes rather than real lakes.

The problem of built-in dimensional correlation is removed, as described in Chapter 6, by constructing dimensionless ratios. The compound scaling function (Equation 15.9a) is then rewritten in terms of dimensionless ratios ($\Pi_1 = A^{1/2}/z$, $\Pi_2 = TDS$) from Box 6.8:

$$\dot{M} = e^{\alpha} {\Pi_1}^{\beta 1} {\Pi_2}^{\beta 2} \tag{15.11a}$$

$$\dot{M} = e^{\alpha} (A^{1/2}/z)^{\beta 1}\ TDS^{\beta 2} \tag{15.11b}$$

At this point we undertake statistical evaluation (Box 15.11) of the results of dimensional analysis (Equation 15.11a, b), following the steps in Table 15.3.

Box 15.11 Analysis of Compound Scaling Relation of Catch to Lake Area, Depth, and Total Dissolved Solids. Data from Ryder (1965)

1. Construct model. Catch depends on lake shape (Π_1) and *TDS*.

$$\dot{M} = e^{\alpha}(\Pi_1)^{\beta 1}\ TDS^{\beta 2}$$

$$\ln \dot{M}\ = \beta_o + \beta_1 \cdot \ln(\Pi_1) + \beta_2 \cdot \ln(TDS) + \varepsilon$$

2. Execution: Take logarithms and execute multiple regression according to model above.
3. Evaluate the structural model with residual versus fit plot.
   No bowl or arch pattern evident. Hence the structural model acceptable.
   Evaluate error assumption. No cone or other pattern of heterogeneity evident in residual plot.
4. Sample consists of 1 value of annual fish catch in each of 21 Canadian shield lakes. Inference is to a hypothetical population, all possible values of fish catch in relation to the scaling function, for lakes in the size range from Lake Superior (8.2×10^6ha) to Lake Heming (259 ha).
5. Mode of inference will be hypothesis testing via construction of confidence limits:

Parameter	Estimates	St. Error	Confidence Limits	
Intercept	6.5020	3.1272		
$\ln\Pi_1$.	2.0177	0.4536	1.07	2.96
ln*TDS*.	0.0529	0.5787	−1.15	1.26

The confidence limits for log(*TDS*) include zero, so we accept the null hypothesis that catch does not scale with *TDS*. The exponent that scales catch to lake shape (Π_1) lies between 1.1 and 3.0 (with 95% certainty). We conclude that fish catch depends on lake morphometry but not on trophic status as measured by total dissolved solids.

The conclusion from the statistical evaluation of the compound scaling relation (Box 15.11) is that catch depends on lake morphometry (as measured by the dimensionless ratio of $\Pi_1 = A^{1/2}/z$) but does not depend on trophic status.

It is of interest to examine the outcome of statistical analysis of the compound scaling relation with both area and depth. When Equation 15.10 is linearized by taking the logarithm of catch, area, *TDS*, and depth, the parameter estimates for depth and *TDS* are not consistent with previous estimates. The exponent for area is now nearly 1, whereas the exponent for *TDS* rises from 0.05 (Box 15.11) to 0.3 (Table 15.6).
The estimates have changed because of the correlation among these variables. The scaling exponent for the morphoedaphic index does not accurately represent the scaling exponents for the two components of the *MEI*. The exponents for depth and TDS are not similar in magnitude, as implied by Equation 15.9a. The scaling with *TDS* appears to be

Table 15.6 Parameter Estimates and Confidence Limits for Multiple Regression Analysis of Scaling Relation of Catch to Lake Area, Depth, and Total Dissolved Solids (Equation 15.10)

Source	Parameter	St. Error (Exponent)	Confidence Limits Upper	Lower
Intercept	1.0218	0.6958	−0.43	2.47
Area	1.0275	0.0469	0.93	1.13
Depth	−0.5606	0.1162	−0.80	−0.32
TDS	0.2939	0.1201	0.043	0.55

significant (Table 15.6) because of the distorting effects of correlation among the regression variables. When these effects are removed by regression of catch against uncorrelated Π ratios, catch turns out to have no significant dependence on dissolved solids.

ANOTHER LOOK AT SECTION 15.6.3

To see whether mortality is density dependent, an ecologist computes *%mortality* as ((*Nfinal*–*Ninitial*)/*Ninitial*), then graphs *%mortality* against *Ninitial*. Write out the function that relates *Ninitial* to *%mortality*. Rewrite the function so that single variables appear to the left and the right of the equal sign. Have you ever plotted one variable against another, where either variable is a ratio? If so, what are the potential problems with statistical analysis of the relation in the graph?

15.7 Reduced Major Axis Estimates of Scaling Parameters

The general linear model command available in software packages automatically computes the least-squares estimates of the parameters. This estimate assumes no (or at least negligible) uncertainty in the explanatory variable. This assumption is always met for a measurement relation because we define the explanatory variable as an operation. The assumption is tenable (Allen, 1939) when the explanatory variable consists of categories, and hence the values of the explanatory variable consist of class marks (midpoints of each category). But this assumption is not met for scaling relations where the independent (or explanatory) variable consists of single measurements. When the assumption concerning the independent (explanatory) variable is not met, *reduced major axis regression* (Allen, 1939; Kermack and Haldane, 1950; Ricker, 1973) is often recommended (Sokal and Rohlf, 1995). The RMA estimate of the power law exponent assigns error to both variables in the regression. Consequently, the RMA estimate of the exponent for *Y1* versus *Y2* is the inverse of the exponent for *Y2* versus *Y1* (Box 15.12). The same is not true for the least-squares regression. If the association between *Y1* and *Y2* is anything less than perfect, the estimate of the exponent via regression of *Y1* and *Y2* will differ from regression of *Y2* on *Y1*. This property of the RMA estimate (Box 15.12) is of no advantage for a measurement relation, since we have no interest in an inverse

relation, such as calculating box size from measurements of lake perimeter. This property is an advantage for scaling relations, where the inverse relation can be of interest. Unfortunately, when we apply our recipe for analysis of uncertainty (Table 15.3), the analysis of residuals will uncover some undesirable properties of the RMA estimate of the exponent of scaling relations.

Box 15.12 Computing One Scaling Exponent from Another Using the Reduced Major Axis Estimate of the Exponent of a Scaling Function Based on a Scaling Relation. *Y1* and *Y2* Both Measured with Error

Theory:

$Y1 = k\ Y2^{\beta}$

$Y1^{(1/\beta)} = (k\ Y2^{\beta})^{(1/\beta)}$

$Y2 = k^{(-1/\beta)}\ Y1^{(1/\beta)}$

Example:

$Vol = k\ A^{3/2}$

$A = k^{2/3}\ Vol^{2/3}$

For the lake data in Fee (1979) the scaling of lake volume (V, m^3) to lake surface area (A, hectares) has an exponent of $\hat{\beta} = 1.418$, as estimated by least-squares regression.

The estimate by RMA regression was = 1.467. The estimates are both close to the theoretical value of 3/2 for an Euclidean object.

$V = e^{9.697}A^{1.467}$

$1/\hat{\beta} = 0.682$ and hence: $A = e^{-6.61}\ V^{0.682}$

Strictly speaking, one cannot apply this to the least-square estimate of the exponent, but in this case the estimates are so close as to make little practical difference.

Box 15.13 shows the computations for reduced major axis regression. Once the model is written (Step 1), it is executed (Step 2) using the RMA estimate. The intercept α, computed from the RMA estimate of the slope, is used to calculate the scaling coefficient as $k = \exp(\alpha)$. The RMA estimate is the line of best fit that minimizes the sum of the squared residuals perpendicular to the fitted line. These are calculated and then plotted against the fitted values. Inspection of this plot (not shown) resulted in a plot similar to Figure 15.5b, where the presence of an outlier resulted in downward trending residuals instead of a horizontal band.

The outlier was then dropped, as it was for the least-squares analysis. Execution (Step 2) is repeated, resulting in revised parameter estimates and a new residual versus fit plot. The RMA line (Figure 15.6a) appears to be consistent with the data. However, the residuals continue to show a downward trend from left to right (Figure 15.6c), unlike the residuals from the least-squares line (Figure 15.6b). Inspection of the residuals showed that in this case, the least-squares estimate was a better representation of the data than the RMA estimate.

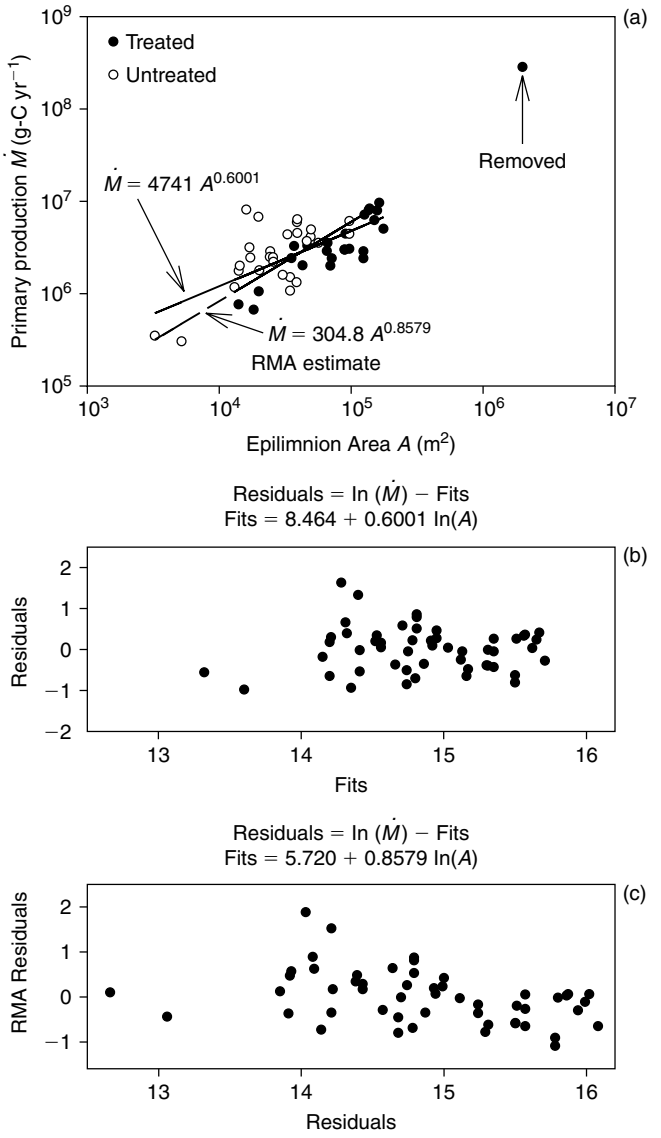


FIGURE 15.6 Scaling of Primary Production to Area of Epilimnion in Fertilized and Unfertilized Lakes. Data from Fee (1979; (a) Logarithmic axes display production as a power law function of area via reduced major axis regression in 18 lakes. Exponents estimated by regression ($A = 0.6001$) and RMA regression ($A = 0.8579$); (b) Regression residuals plotted against fitted values from analysis of 18 lakes; (c) RMA residuals potted against fitted values.

The RMA estimate is 0.858/0.6001 = 1.43 times higher than the least squares estimate for the same data. Based on the algebra of the RMA estimate the degree of inflation of β_{RMA} depends on the strength of correlation, where r is the Pearson correlation of the two variables:

$$\beta_{LS} = r \cdot \beta_{RMA} \tag{15.12}$$

Box 15.13 Analysis of Primary Production Scaled to Epilimnion Area in Treated and Untreated Lakes. Reduced Major Axis Estimates of Exponent of Linearized Power Law. Data from the Experimental Lakes Area in Ontario (Fee 1979)

1. Construct model: $\ln \dot{B} = \beta_o + \beta_A \cdot \ln A + \varepsilon$
2. Execute model (all lakes). Estimate β_A from ratio of the standard deviations:

$$\hat{\beta}_A = \frac{\sqrt{\sum (B_i - \bar{B})^2}}{\sqrt{\sum (A_i - \bar{A})^2}} = \frac{\text{stdev}(B)}{\text{stdev}(A)} = \frac{\sqrt{50.77}}{\sqrt{55.36}} = 0.958$$

$$\hat{\alpha} = 14.94 - 0.958 \cdot 10.72$$

$$\exp(\hat{\alpha}) = \exp(4.675) = 107 \text{ g-C yr}^{-1} \text{ ha}^{-0.958}$$

Compute fitted values and residuals from estimates: $\hat{\dot{B}} = 107A^{0.958}$ $\varepsilon = \dot{B} - \hat{\dot{B}}$
Compute residuals perpendicular to the fitted line: $\varepsilon_{RMA} = \varepsilon \cdot \cos(\arctan(\hat{\beta}))$
Residuals differ by a constant, so either residual versus fit plot can be used for evaluation.

3. Evaluate model: Residual plot shows a downward trend created by the outlying value (Figure 15.4b). This creates a negative correlation between the residual and the fitted values ($r = -0.33$).
2. Return to step 2. Execute model (lake 228 removed).

$$\hat{\beta}_A = \frac{\sqrt{29.93}}{\sqrt{40.67}} = 0.858$$

$$\hat{\alpha} = 14.85 - 0.858 \cdot 10.65$$

$$\exp(\hat{\alpha}) = \exp(5.720) = 305 \text{ g-C yr}^{-1} \text{ ha}^{-0.858}$$

Compute fitted values and residuals from estimates: $\hat{\dot{B}} = 305A^{0.858}$
Compute residuals perpendicular to the fitted line $\varepsilon_{RMA} = (B - 305A^{0.858}) \cdot \cos(\arctan(0.858))$ Residuals differ by a constant, so either residual versus fit plot can be used for evaluation.

3. Evaluate model. Residual plot shows downward trend despite removal of outlier (Figure 15.6c). The negative correlation between the residual and the fitted values persists ($r = -0.39$).

As the correlation of the variables drops, the RMA estimate becomes more inflated relative to the least-squares estimate. This, and the demonstrably inappropriate model resulting from the RMA estimate (Figures 15.6b), suggest that RMA estimates need to be treated with caution. RMA estimates are a statistical solution to the problem of explanatory variables measured with error. This adjustment becomes less and less

satisfactory as the error in the explanatory variable increases, thus hiding a true relation between response and explanatory variable. A better solution is to reduce or eliminate the problem. One way to eliminate the problem is to construct categories of the explanatory variable, then regress the mean value of the response against the class mark of the explanatory in each category. This eliminates the need for RMA regression (Allen, 1939). The next section describes the use of dimensional methods to eliminate the problem.

ANOTHER LOOK AT SECTION 15.7

Burger (1993) reported the number of seabirds killed (N) in relation to the size of an oil spill (M = tonnes). The exponent that scales number killed to spill size is 0.223, with confidence limits of 0.024 to 0.423, as estimated by least squares for a log/log regression (linearized power law). If the correlation of $\ln(N)$ with $\ln(M)$ is 0.374, compute the RMA estimate of the scaling exponent. Is this estimate consistent with the least squares estimate as defined by its confidence limits? Discuss how you would decide which estimate to use if you were asked to calculate the expected kill from a spill of known size.

15.8 Estimates of Scaling Parameters Via Measurement Scope

Estimating the exponent of a scaling relation is problematic if both variables are measured with error. The problem of measurement with error can be eliminated by working with measurement relations. For example, we might be interested in how species number *Nsp* per quadrat scales with habitat diversity defined as the number of distinguishable habitats *Nhab* in the quadrat. The scaling relation is:

$$\frac{Nsp}{Nsp_{ref}} = \left(\frac{Nhab}{Nhab_{ref}} \right)^{\beta} \tag{15.13}$$

The exponent β would conventionally be estimated by regressing *Nhab* on *Nsp*, which entails regression against an explanatory variable with considerable potential for measurement error. If the number of habitats in a unit were small (say, just 5), a counting error of 1 would result in an error of 1 in 5 or 20%.

This scaling relation is necessarily based on a measurement relation for each variable. Any of several measurement protocols are possible. One possibility is equally sized measurement units, such as quadrats. The scaling relation based on equally sized units will have a spatial scope of 1. Another possibility is a scaling relation based on spatial units ranging from small to large using one of the maneuvers in Figure 7.1. These might be discrete units of different sizes, resulting in scaling by rating. These might be constructed from units of equal size via sequential addition, resulting in scaling by accumulation. If equally sized units are evenly spaced (or contiguous), we can use coarse graining. Regardless of the maneuver, we can use the definition of measurement protocol

to set up measurement relations for each of the two variables of interest, provided the spatial scope exceeds 1.

$$\frac{Nhab}{Nhab_{ref}} = \left(\frac{A}{A_{ref}}\right)^{\beta Nhab} \tag{15.14a}$$

Equivalently,

$$\left(\frac{Nhab}{Nhab_{ref}}\right)^{1/\beta Nhab} = \left(\frac{A}{A_{ref}}\right) \tag{15.14b}$$

We scale species number to habitat diversity from counts of species and counts of habitat types in the same spatial units, so the measurement relation is the same for species data and habitat data:

$$\frac{Nsp}{Nsp_{ref}} = \left(\frac{A}{A_{ref}}\right)^{\beta Nsp} \tag{15.15a}$$

Equivalently,

$$\left(\frac{Nsp}{Nsp_{ref}}\right)^{1/\beta Nsp} = \left(\frac{A}{A_{ref}}\right) \tag{15.15b}$$

The measurement operation (A/A_{ref}) is that of altering quadrat area relative to some reference quadrat A_{ref}. Defining the measurement operation for each variable quantity allows reformulation of the scaling relation:

$$\left(\frac{Nsp(A)}{Nsp(A_{ref})}\right)^{1/\beta Nsp} = \left(\frac{A}{A_{ref}}\right) = \left(\frac{Nhab}{Nhab_{ref}}\right)^{1/\beta Nhab} \tag{15.16a}$$

$$\frac{Nsp}{Nsp_{ref}} = \left(\frac{Nhab}{Nhab_{ref}}\right)^{\frac{\beta Nsp}{\beta Nhab}} \tag{15.16b}$$

The ratio $\beta Nsp/\beta Nhab$ in Equation 15.16b is the estimate of β in 15.14. Thus, to obtain an estimate of β free of the problem of error in measuring *Nhab*, we use least-squares regression to estimate both βNsp and $\beta Nhab$, then form the ratio.

To obtain the scaling function (see Equation 2.5c), we rearrange the scaling relation to:

$$Nsp = \left(Nsp_{ref} \cdot (Nhab_{ref})^{-\beta Nsp/\beta Nhab}\right) Nhab^{\beta Nsp/\beta Nhab} \tag{15.17a}$$

The scaling function (see Equation 2.5c) is:

$$Nsp = k \cdot Nhab^{\beta Nsp/\beta Nhab} \tag{15.17b}$$

Table 15.7 extends this example to a generic procedure for using dimensional analysis to estimate scaling relations. Box 15.14 illustrates the procedure with another example.

Table 15.7 Generic Recipe for Estimating Scaling Relations Via Dimensional Analysis

1. Define a scaling relation between the quantities of interest.
2. Define a measurement relation for each quantity.
3. Rewrite each measurement relation so that the operation has an exponent of 1.
4. Write the relation between the two measurement operations. Usually this will be that the measurement operations are the same.
5. Using the relation between measurement operations, rewrite the scaling relation in terms of exponents from the measurement operations.
6. Use least-squares regression to estimate the exponents of each measurement relation.
7. Combine estimates from the measurement relations to obtain the exponent for the scaling relation.

In the example of *Nsp* scaled to *Nhab*, the measurement relations were identical. Box 15.14 applies the recipe to an example where the measurement relations are not identical. In this example the perimeter and the area of multiple clouds (or rain areas) were measured on a 1 km by 1 km grid. Thus the example is for a noniterative scaling relation based on noniterative measurement relations.

Box 15.14 Scaling Relation of Cloud Area to Cloud Perimeter Via Measurement Relations

1. Lovejoy (1982) examined whether the area/perimeter relation for rain and cloud areas follows the Euclidean scaling, where Perimeter $\sim$ Area$^{1/2}$. The scaling relation is:

$$\frac{Pcloud}{Pcloud_{ref}} = \left(\frac{Acloud}{Acloud_{ref}}\right)^{D/2}$$

   where D = 1 for Euclidean clouds, and D $>$ 1 for fractal clouds.

2. The measurement relations are:

$$\frac{Acloud}{Acloud_{ref}} = \left(\frac{A}{A_{ref}}\right)^{\beta area}$$

$$\frac{Pcloud}{Pcloud_{ref}} = \left(\frac{L}{L_{ref}}\right)^{\beta perim}$$

3. These are rewritten as:

$$\left(\frac{Acloud}{Acloud_{ref}}\right)^{1/\beta area} = \left(\frac{A}{A_{ref}}\right)$$

$$\left(\frac{Pcloud}{Pcloud_{ref}}\right)^{1/\beta perim} = \left(\frac{L}{L_{ref}}\right)$$

4. Based on the description in Lovejoy (1982), the measurement relations appear to be related as follows:

$$\left(\frac{L}{L_{ref}}\right) = \left(\frac{A}{A_{ref}}\right)^{1/2}$$

5. The scaling relation is then:

$$\left(\frac{Pcloud}{Pcloud_{ref}}\right)^{1/\beta perim} = \left(\frac{L}{L_{ref}}\right) = \left(\frac{A}{A_{ref}}\right)^{1/2} = \left(\left(\frac{Acloud}{Acloud_{ref}}\right)^{1/\beta area}\right)^{1/2}$$

$$\left(\frac{Pcloud}{Pcloud_{o}}\right) = \left(\frac{Acloud}{Acloud_{o}}\right)^{\beta perim/2\beta area}$$

The relation of the two measurement operations (Box 15.14, Step 4) is based on the procedural statement (Lovejoy, 1982, p. 186). Area was measured by counting the number of 1 km by 1 km boxes for each cloud, hence $A_{ref} = 1\,km^2$. Perimeter appears to have also been measured by counting boxes that contain cloud perimeter, hence $L_{ref} = 1\,km$ and so $L_{ref} = A_{ref}^{1/2}$. When the scaling relation of perimeter to area is rewritten from the measurement relations, the estimate of the scaling exponent is $D/2 = \beta perim/2\beta area$, or $D = \beta perim/\beta area$. The procedure in Table 15.7, as illustrated in Box 15.14, brings out the fact that the exponent that scales one quantity to another depends on the relation of the underlying measurement relations.

ANOTHER LOOK AT SECTION 15.8

For a quantity of interest to you, write a scaling relation to another variable as a power law. Define the measurement relation for each variable. Is the relation of the two scaling operations clear? If not, describe how you might determine the relation.

15.9 Scaling and Uncertainty via the General Linear Model

In the statistical evaluation of scaling functions, we are interested first in whether one quantity scales with another. Often we are interested in whether the scaling of one variable to another depends on category. And often we are interested in whether a scaling relation holds, controlled for an additional explanatory variable. The GLM provides a coherent model-based approach for statistical evaluation of these questions. Table 15.8 shows special cases of the general linear model in relation to the number of explanatory variables, the type of explanatory variable (regression, categorical, or both), and the

Table 15.8 Special Cases of the General Linear Model. Variables are either ratio scale or nominal scale (factors).

Analysis	Response Variable	Explanatory Variable	Interaction?	Comments
t-test	1 ratio	1 nominal	Absent	Compares two means
1-way ANOVA	1 ratio	1 nominal	Absent	Compares three or more means in one factor
2-way ANOVA	1 ratio	2 nominal	Present	Tests for interactive effects
			Absent	Compares means in two factors, if no interaction
Paired Comparison	1 ratio	2 nominal	Assumed Absent	Compares two means in one factor, controlled for second factor (blocks or units)
Randomized Blocks	1 ratio	2 nominal	Assumed Absent	Compares three or more means in one factor, controlled for second factor (blocks or sampling units)
Hierarchical ANOVA	1 ratio	≥2 nominal	Absent	Nested comparisons of means
ANCOVA	1 ratio	≥1 ratio	Present	Compares two or more slopes
		≥1 nominal	Absent	Compares means, controlled for slopes
Regression	1 ratio	1 ratio	Absent	Tests linear relation of response to explanatory
Multiple Regression	1 ratio	≥2 ratio	Often	Linear relation to two or more explanatory variables
			Absent	Assumes independent effects on response variable

presence of interaction terms. Though ANOVA designs are generally not used in scaling analysis, the ready incorporation of categorical variables into the GLM opens the door to comparison of scaling and measurement relations across multiple categories. For example, the analysis of the scaling of primary production to area of epilimnion in treated and untreated lakes (Boxes 15.8, 15.9, 15.10) could be extended to include year effects, since most lakes were measured in successive years. The full model for an analysis that includes time contains several interaction terms, some of which address key questions.

$$\ln\dot{M} = \beta_o + \beta_A \cdot \ln A + \beta_{Tr} \cdot Tr + \beta_{Yr} \cdot Yr + \beta_{A \cdot Tr} \cdot \ln A \cdot Tr \ \ldots$$

$$+\beta_{A \cdot Yr} \cdot \ln A \cdot Yr + \beta_{Yr \cdot Tr} \cdot Yr \cdot Tr + \beta_{Yr \cdot Tr \cdot \ln A} \cdot Yr \cdot Tr \cdot \ln A + \varepsilon \qquad (15.18)$$

We have already considered the dependence of the scaling on treatment $\beta_{A \cdot Tr}$. Extending the analysis to year introduces a new interaction term, with a parameter $\beta_{Yr \cdot Tr}$ that represents change in scaling exponent with year. The coefficient $\beta_{Tr \cdot lnA \cdot Yr}$ quantifies the degree to which the dependence of the scaling exponent on treatment differs across years. The GLM opens the door to a complete analysis of the dependencies in the data, including the question of whether the scaling exponent changes with both treatment and year.

This extension of the lake example illustrates how the model-based approach permits comprehensive statistical analysis of scaling functions. The model-based approach employs biological reasoning. It is not constrained to a search for the "right test" from within a list of named tests. The software is available from the graphical (menu-driven) interface in any of the widely used statistical packages. The procedure for writing the model differs somewhat from package to package but is readily implemented by using the logical sequence (Steps 1–3) in Table 15.3. Assumptions are readily diagnosed with residual plots, as in Figures 15.2 through 15.6.

Set against these advantages are limitations. The general linear model only handles a normal (homogeneous) error structure. This is a problem because one of the commonest violations of the assumptions for ecological data is heterogeneous errors, which produce inaccurate *p*-values when sample size is small. Violation of the homogeneous error assumption results in parameter estimates that are unduly influenced by those observations that are the least certain (have the highest underlying variability).

One widespread solution is to reduce data to ranks, then compute a measure of association (rank correlation) or a measure of difference (such as the *H* statistic of the Kruskal-Wallis test, Sokal and Rohlf, 1995). This solution discards information, reduces the ability to detect differences, and can only be applied to the simplest of designs. Because they reduce data on a ratio type of scale to a rank scale, these tests are of no utility in the analysis of scaling and measurement relations. The justification for rank-based tests (that they are relatively easy to calculate) disappeared in the 1980s, when computers became widely available.

A far better solution to the problem of heterogeneous error is to use randomization to generate a *p*-value. Randomization methods eliminate the problem of assumptions concerning the error distribution, but at a cost: they are laborious to compute. This computational limit disappeared once computers became widely available. Execution time is small, even for complex models. Even on a spreadsheet (an extraordinarily slow computation because the screen is updated on each randomization), the 4000 runs in Box 15.5 took only 6 minutes. But though execution time is small, setup time can be substantial. Setting up the spreadsheet calculations in Box 15.5 took several hours. Randomization methods are straightforward for simple models but not for models with explanatory terms that consist of categories. For these, it might be more appropriate to randomize the response variable within levels of one of the explanatory variables, which can be cumbersome.

The problem of heterogeneous errors often disappears when the response variable is transformed to a logarithmic scale. The log transform often (but not always) eliminates the heterogeneity visible as a cone in the residual versus fit plot. For example, transformation of the algal biomass data (Box 15.1) to a log scale eliminated the cone-shaped heterogeneity in the residuals. It must be kept in mind, however, that a logarithmic transformation alters the model. The log transformation of the response variable results in a linear model (Equation 15.19a) that becomes an exponential function when converted back (Equation 15.19b).

$$\ln(Q) = \beta o + \beta x\ X + \varepsilon \tag{15.19a}$$

$$Q = \mathrm{e}^{\beta o}\ e^{\beta x X}\ e^{\varepsilon} \tag{15.19b}$$

The errors are additive on a logarithmic scale, hence multiplicative when working with scaling of the response to the explanatory variable. Log transformation of both the response and explanatory variable results in a linear model (Equation 15.2) that becomes a power law when converted back (Equation 15.3). The errors are again multiplicative when working with the scaling of response to explanatory variable.

Another solution to the problems created when assumptions are violated is to adopt a more appropriate error distribution within the framework of the generalized linear model (Nelder and Wedderburn, 1972; McCullagh and Nelder, 1989). This approach, one of the major developments in statistics in the last quarter of the 20th

century, comes at a cost in that it is not widely available in the easier-to-use statistical packages, nor is there an undergraduate text. In practice, the approaches in this chapter, notably randomization methods for computing *p*-values, will suffice in most instances where statistical evaluation of a scaling function is appropriate.

ANOTHER LOOK AT SECTION 15.9

For a quantity of interest to you, write a scaling relation to another variable as a power law, then write a general linear model to analyze the dependence of the power law on an additional explanatory variable. Above each term in the model, sketch a graph showing the null ("just chance") hypothesis.

Defined Terms and Concepts for Review and Future Reference

____ analysis of covariance (ANCOVA)
____ analysis of variance (ANOVA)
____ four assumptions for *p*-values
____ compound scaling relation
____ data equations
____ degrees of freedom
____ expected value
____ F-ratio (variance ratio)
____ interaction term
____ measures of goodness of fit
____ model-based statistics
____ null and alternate model
____ randomization test
____ reduced major axis regression
____ response variable, structural model, and error term
____ Type I error (*p*-value) and Type II error

16 Power Laws and Scaling Theory

Bacteria are—and always have been—the dominant forms of life on earth. Our failure to grasp this most evident of biological facts arises in part from the blindness of arrogance, but also, in large measure, as an effect of scale. We are so accustomed to viewing phenomena of our scale—sizes measured in feet and ages in decades—as typical of nature.

—Stephen Jay Gould, *Life's Grandeur*, 1996

16.1 Synopsis

This chapter briefly reviews the role of theory in body-size scaling and biodiversity scaling. It then considers the prospects for theoretical development in spatial scaling.

Body-size scaling has a long history of theoretical development, beginning in the 19th century, when the classic theory of surface-limited metabolic activity replaced a more intuitive one-to-one scaling. By the middle of the 20th century it was evident that the exponent relating metabolic rate to body size was closer to ¾ than to the ⅔ exponent from classical theory. Early theories for ¾ scaling include elastic similarity, exchange with the environment through fractal surfaces, and the energy density of tissue. A more recent and comprehensive theory extends ¾ scaling to any supply system, ranging from the vascular systems of plants and animals to the drainage system of watersheds.

The scaling of biodiversity to area has a history of empirical findings that extend back to the 19th century. Biodiversity scales with area, time, collection size, and body size, but not in a one-to-one fashion. Theoretical development began in the middle of the 20th century, with the concept that divergence from a 1:1 scaling arises from the vagaries of sampling (at the scale of plots), from habitat diversity (at the scale of biogeographic regions), and from evolutionary history (at the scale of continents). Alpha, gamma, and epsilon diversity correspond to these three domains. Quantitative scaling of taxonomic diversity (species number) with area rests on two ideas: that the total number of organisms scales with area and that the number of species scales with number of organisms to the ¼ power. Scaling of species number with productivity is expected when density increases with primary productivity. The scaling of species number to area depends on degree of isolation. The exponent is steeper for islands than for blocks of

doi: 10.1016/B978-0-12-627865-1.00016-7

land on continents. The exponent is steeper for biogeographic provinces, which are isolated at evolutionary time scales, than for blocks within a province. An analytic review is needed to integrate the vast literature on biodiversity with the literature on the structure and dynamics of landscapes.

Ecological theory is founded on exponential rates and equilibrium dynamics. When production and loss act episodically, power laws emerge. This theory, called universal scaling or complexity theory, has the potential to become the theoretical basis for spatial and temporal scaling in ecology.

16.2 The Role of Theory in Body-Size Scaling

The word *allometry* usually refers to a special case of allometric rescaling: the scaling of organism form or function according to body size (Gould, 1966; Calder, 1984). Allometric rescaling to body size was developed by Darcy Thompson in his 1917 treatise *On Growth and Form* (Thompson, 1961). Thompson used the principle of geometric similitude to scale organism form and function with size. The rationale for geometric similarity is the observation that organisms are nearly incompressible and have densities (mass per unit volume) close to that of seawater. Consequently, any scaling of form or function with body volume can be rescaled with body mass according to the same exponent. Thompson's scaling of form and function to body size initiated a major line of biological research (Huxley, 1932; Brody, 1945; Gould, 1966; Pedley, 1977; Vogel, 1981; Peters, 1983; Calder, 1984; Schmidt-Nielsen, 1984; Alexander, 1989), even though many of his specific conclusions have not survived (Schmidt-Nielsen, 1984). In the latter part of the 20th century, body-size allometry was extended to population and community-level phenomena (Gold, 1977; Platt and Denman, 1978; Damuth, 1981; Platt, 1981; Platt and Silvert, 1981; Calder, 1983, 1984; Peters, 1983; Dickie et al., 1987; Rosen, 1989).

The role of theory in body-size allometry can be traced to the 19th century, when Sarrus and Rameaux (1839) proposed that respiration scales with surface area, for which the Euclidean scaling is volume$^{2/3}$ and hence mass$^{2/3}$. Rubner (1883) presented evidence that in dogs, metabolic rate does scale as mass$^{2/3}$. As studies of metabolic rate accumulated through the early 20th century it became evident that across a wide range of animals the exponent is slightly greater than $\frac{2}{3}$. Kleiber (1947) amassed a substantial body of evidence that over this wider range the scaling of metabolic activity with body size was closer to $\frac{3}{4}$ than $\frac{2}{3}$. Two decades later, McMahon (1973) developed a scaling of metabolic rate to body size based on mechanical similarity of organisms. McMahon reasoned that large animals must be stockier than smaller animals because of structural limitations on the skeleton. Using the concept of elastic loading, McMahon arrived at a scaling in which surface area increases as volume$^{3/4}$ rather than volume$^{2/3}$. The concept of elastic loading is readily extended to plants (McMahon, 1975). This allometric scaling is closer to the observed exponent than that based on the $\frac{2}{3}$ surface law. Confirmation comes from the finding that limb bone length scales as body mass$^{3/4}$ in ungulates, which is consistent with McMahon's similarity statements (Schmidt-Nielsen, 1984). Further support comes from the failure to find $\frac{3}{4}$ scaling in marine mammals (Kovacs and Lavigne, 1985), animals that do not use their limbs to support themselves. Alexander

et al. (1979) noted that ¾ scaling does not hold within several groups of terrestrial mammals. It is worth noting that scaling within groups (e.g., within rodents) often occurs with a smaller exponent than scaling across groups, as from mice to elephants (Schmidt-Nielsen, 1984).

Platt and Silvert (1981) developed scalings of respiration to body size based on the energy density of tissue rather than on the area across which energy is exchanged with the environment. Based on their similarity statements, the scaling of respiration to body mass was:

$$\dot{E} \cong M^{3/4} \quad \text{in terrestrial organisms} \tag{16.1a}$$

$$\dot{E} \cong M^{2/3} \quad \text{in aquatic organisms} \tag{16.1b}$$

A check against Peters' (1983) power law compilation shows that the aquatic scaling does not apply to some 13 studies of fish and another 29 studies of aquatic metazoans. A few aquatic groups do show exponents below 0.70.

A modification of the classic ⅔ surface-to-volume scaling is that respiration scales with the surface area of the lung. Weibel (1979, p. 156, in Section 4.3.7) reported a scaling exponent 1.17 for the surface of the lung. Citing this same passage from Weibel, Mandelbrot (1982, p. 114) reported the exponent as 2.17. Weibel reports using the dividers method, which yields an exponent of $\beta = -0.17$, obtained by regressing number of steps against step length. Weibel reported the result as a fractal dimension $D_f = 1.17$, which means the fractal dimension was taken relative to a line ($D_f = 1 - \beta$). It appears that Mandelbrot then used Weibel's estimate to report the fractal dimension as $D_f = 2.17$, that is, relative to a plane ($D_f = 2 - \beta$). A fractal surface with dimension of 2.17 increases the scaling exponent for flux across the surface of a volume from $V^{2/3}$ to $V^{2.17/3} = V^{0.72}$. The resulting scaling is:

$$\dot{E} \cong V^{0.72} \cong M^{0.72} \tag{16.2}$$

which is indistinguishable from ¾. This derivation was reported independently three times (Barenblatt and Monin, 1982; Pennycuick, 1992; Schneider, 1994b).

Subsequently, West et al. (1997) offered a theoretical explanation for an exact exponent of volume$^{3/4}$. This theory replaces the theory of surface-limited flux with a scaling based on structured supply through a (fractal) vascular system. Banavar et al. (1999) showed that a ¾ scaling arises in any supply network, regardless of whether it is fractal. The scaling is $D/(D + 1)$, that is, ¾ scaling of supply to a volume and ⅔ scaling of supply to a plane. These authors extended the supply rate theory to inanimate delivery systems, notably the dendritic structure of river systems. West et al. (1999) then generalized their theory to a range of exponents. For a vascular system with surface area $A = L^{2+\varepsilon}$ and active tissue volume $V = L^{3+\varepsilon+\zeta}$, delivery scales as $V^{(2+\varepsilon)/(3+\varepsilon+\zeta)}$. Delivery is maximized at $\varepsilon = 1$ and $\zeta = 0$, hence delivery scales with an exponent near $V^{3/4}$ and hence Mass$^{3/4}$. The channeled supply rate theory (West et al., 1997; Banavar et al., 1999) has generated a substantial literature (see Volume 18 of *Functional Ecology*, Volume 85(7) of *Ecology*, and Volume 208 of *Journal of Experimental Biology*).

16.3 Biodiversity Scaling

Traditional treatment of the scaling of biodiversity has focused largely on a single measure, the number of species (species richness) in relation to area. Nearly a century ago Arrhenius (1921) showed that species number does not scale in a 1:1 fashion with area. As a rough rule of thumb, a tenfold increase in area will increase the number of plant species by a factor of 2 rather than 10. Thus as a first approximation, the exponent that scales species number to area is ln(2)/ln(10), or roughly $1/3$. A graphical compilation of studies from the first half of the 20th century (Williams, 1964) showed that the exponent that scales species number to area itself depends on spatial scale. The scaling of species number with area is relatively steep at the scale of plots (attributable to the vagaries of sampling), less steep at the scale of biogeographic regions (attributable to habitat diversity), and then steeper again at the scale of continents (attributable to evolutionary history). Whittaker (1960) expressed the idea that taxonomic diversity depends on scale by drawing the distinction between alpha (within habitat) and gamma (among habitat) diversity. Whittaker (1977) then extended the concept to the smaller scale of point diversity within a habitat and to the larger scale of regional (epsilon) diversity. Whittaker (1977) presented these as convenient labels with approximate spatial scales. Subsequent work has shown that species number increases, but not in a 1:1 fashion, as more time is spent watching an area, as more organisms are collected, and as a wider range of body sizes is accumulated.

One of the major advances in biodiversity scaling has been the development of biodiversity distributions (Preston, 1962; May, 1975). A *biodiversity distribution* is a frequency distribution showing the number of classes (usually species) that have 1, 2, 3 (and so on) individuals per class (per species). Because biodiversity distributions can use any classification (not just taxonomic groups), they can be constructed for habitat and genetic diversity. This is consistent with current practice, which is to define *biodiversity* as including habitat and genetic diversity along with taxonomic diversity. Section 16.3.1 covers the topic of measurement of biodiversity, from single-value measures to the use of the full frequency distribution. Section 16.3.2 briefly reviews theoretical development of biodiversity scaling.

16.3.1 Measuring Biodiversity

Diversity studies typically employ two quantities (N = number of organisms, A = area) and sometimes a third (T = time or duration of measurement). A collection of N organisms from a unit area A_o during a unit period T_o, is sorted into s groups (typically species), as shown in Figure 16.1. Any taxonomic level can be used in sorting, but for the sake of the clarity that comes with clear examples, species will be used to illustrate taxonomic diversity distributions. The number of organisms n_i is recorded for each species (Table 16.1). A familiar presentation of this information is the curve of abundance across species, ranked from high to low (Figure 16.1). The information in this *rank-abundance curve* can be reexpressed as a *taxonomic frequency distribution* $s(n_i = n_k)$, defined in Table 16.1. The symbol $s(n_i = n_k)$ represents the number of species at abundances ranging from $n_k = 1$ upward to the abundance n_k of the commonest species in the collection. This frequency distribution shows all the information on diversity in a

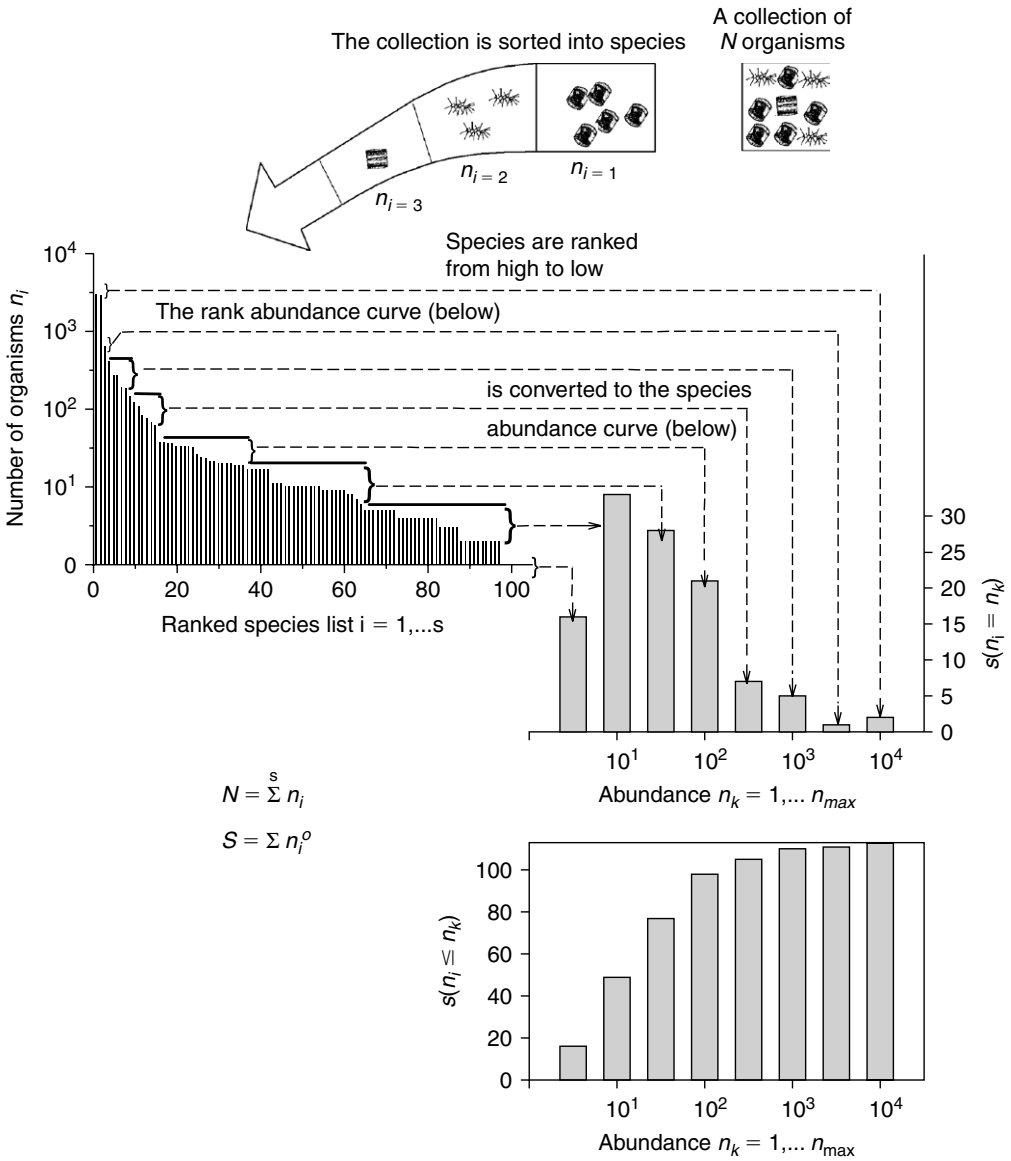


FIGURE 16.1 Construction of the Cumulative Diversity Distribution. Arrows show how the taxonomic diversity distribution $s(n_i = n_k)$ is constructed from the rank abundance curve, for a collection of organisms. The cumulative diversity distribution $s(n_i < n_k)$ is then constructed from the diversity distribution. *N* is number of organisms in the collection, *s* is the number of species. Data from Patrick (1968). Redrawn from Schneider 2001c.

single collection of $\Sigma n_i = N$ organisms, taken from unit area A_o in unit time T_o. If our interest is in the effects of scaling, we will need to work with the frequency distribution $s(n_i = n_k)$, which is on a ratio type of scale, rather than with the rank abundance curve (Figure 16.1), which is on an ordinal type of scale.

The information in the species abundance distribution $s(n_i = n_k)$ is often reduced to a single number, a *diversity index*. The most common index of taxonomic diversity is

Table 16.1 Notation for Biodiversity

N	Collection size; number of organisms in single collection unit
A_o	Unit collection area; square area from which N is obtained
T_o	Unit collection time; period during which collection N is obtained
s	Number of classes (usually, species) into which collection N is sorted
n_i	Number of organisms in each class, labeled $i = 1$ to s
Taxonomic Diversity	
$s(n_i = n_k)$	Taxonomic diversity distribution; number of classes (usually species) at abundances ranging from low ($n_k = 1$) to high ($n_k = Nmax$); $Nmax$ is the number of organisms in the most populous class
$s(n_i \leq n_k)$	Cumulative taxonomic diversity distribution; number of classes with n_k or fewer organisms
$S(n_i = n_k)$	Taxonomic diversity distribution for a community; usually estimated from multiple collections, labelled $j = 1$ to jt
Genetic Diversity	
q	Genotype frequency; number of classes into which collection N is sorted
$q(n_i = n_k)$	Genetic diversity distribution of collection N; number of classes with n_k organisms
$q(n_i \leq n_k)$	Cumulative genetic diversity distribution of collection N; number of classes with n_k or fewer organisms
$Q(n_i = n_k)$	Genetic diversity distribution of population; estimated from multiple collections
Habitat Diversity	
A_o	Unit area; smallest unit of area within a larger area of interest
A_i	Habitat area; area of each habitat within area ΣA_i
h	Habitat diversity; number of habitats in area ΣA_i
$h(A_i = A_k)$	Habitat diversity distribution within area ΣA_i; number of unit areas of habitat i within larger area ΣA_i
$h(A_i \leq A_k)$	Cumulative habitat diversity distribution; number of unit areas with A_k or less area
$H(A_i = A_k)$	Habitat diversity distribution of an ecosystem; estimated from several completely surveyed areas ΣA_i

the total number of species s (species richness). An example is the number of cichlid fish species found in each of six African lakes, as reported by Ricklefs and Schluter (1993a):

Procedural Statement	**Name**	**Symbol**	**Numbers**	**· Units**
Ricklefs and Schluter (1993a, p. 358)	Species number	$s_j =$	$\begin{bmatrix} 200 \\ 136 \\ 200 \\ 7 \\ 9 \\ 40 \end{bmatrix}$	$\cdot$ # lake^{-1}

In this example the lakes are listed in order from large (Lake Victoria = 69,484 km^2) to small (Lake Edward = 2150 km^2). Biodiversity indices such as the Shannon Weaver or Simpson index (Magurran, 1988) contain more information than just the number of species s. Yet as with any index, they fall short of describing all the information in the species diversity distribution $s(n_i = n_k)$ for the collection N.

If we obtain a second collection, again from an area A_o during a period T_o, we expect a somewhat different number of organisms $n_{j=2}$, a new set of species $s_{j=2}$ (many the same as in the first collection), and a new diversity distribution $s_2(n_i = n_k)$. With several collections $s_j(n)$, labeled $j = 1$ through jt, we have a fuller characterization of diversity. We can construct the combined distribution $S(n_i = n_k)$ across some number of collections, jt. The total number of species S across the combined collection will exceed the number of species s_j in any one collection.

If the collections were drawn always by the same procedural statement, they can be viewed as samples from a set of all possible collections. The diversity distribution for this hypothetical set is $S(n_i = n_k)$. This is estimated from the observed frequency distribution over many collections, provided that the collections are representative. As with any sampling effort, a representative sample is more surely achieved by samples that are random and have an equal (or at least known) chance of appearing in the sample than by haphazard sampling. The total number of species S and total number of organisms N are computed respectively as $S = \Sigma\ s(n_i = n_k)$ and $N_{tot} = \Sigma\ n_j$. Following May (1975) we can compare an observed distribution $s(n_i = n_k)$ to a mathematical function, allowing us to compute species numbers at unobserved values of abundance n_k. An impressive number of methods (Bunge and Fitzpatrick, 1993) have been devised to estimate the number of species present on the landscape, from collections $s_j(n_i = n_k)$ taken to be samples.

When making statistical estimates from taxonomic distributions, it often proves convenient to use the cumulative distribution $s(n_i \leq n_k)$, which is shown in Figure 16.1 immediately beneath the distribution $s(n_i = n_k)$. The cumulative distribution rises monotonically from left to right as it records the number of species with n_k or fewer organisms per species. Cumulative distributions are useful in estimating the fit of a distribution to a functional expression because they eliminate the presence of zeroes in classes on the upper end of the distribution. Cumulative distributions are necessary for spatial and temporal scaling, which depends on taking ratios, as described in the following discussion.

Genetic variability within a species can be characterized in the same fashion as taxonomic diversity. For a collection of N individualis the usual single locus measure is the proportional presence of each type of allele $q_i = n_i/2N$. This information can be expressed as a rank-abundance distribution of allelic proportions from common (q_i large) to rare (q_i small). For convenience, the information in this distribution is reduced to an index such as number of classes (alleles) or expected heterozygosity ($H_e = 1 - \Sigma q_i^2$). The information in several collections can be expressed as a rank-abundance curve for each collection and a rank abundance curve for the combined collection. This information can in turn be reduced to an index of population structure, such as F_{st}, which compares heterozygosity across collections to the heterozygosity of the entire population. This approach, applied to genetic diversity at the single-locus level, can be extended to coarser classifications, such as recognizable genotypes in a population (Mallet, p. 16, in Gaston, 1996). If we are interested in the effects of scale on genetic diversity, the logic of working with ratio scale quantities will compel us to reexpress the rank-abundance curve as the genetic diversity distribution $q(n_i = n_k)$, which is on a ratio type of scale. This is the frequency of genotypic classes for which there was one individual ($n_k = 1$), two individuals ($n_k = 2$), and so on. This can be converted to a cumulative frequency distribution $q(n_i \leq n_k)$ to make comparisons.

We can use a diversity distribution to quantify habitat diversity. Instead of classifying individuals, we classify spatial units. From this we can construct the usual

rank-abundance curve, running from the commonest habitat (n_i large) to the rarest (n_i small). The same information can be plotted on a ratio scale as a habitat diversity distribution $h(n_i = n_k)$.

Biodiversity does not scale in a 1:1 fashion with collection size N, collection area A, or collection duration T. We can use any of several scaling maneuvers (Figure 7.1) to examine dependence on collection size, area, or duration. We can use rating, as when we examine dependence of diversity on area across units that differ in area (Arrhenius, 1921). We can examine diversity relative to effort, measured as the accumulated number of collections, with a "collector's curve." Figure 16.2a shows the increase in diatom species number with increase in effort (number of slides). Collector's curves can be scaled to the increase in area or increase in time as the number of collections increases. We can examine diversity as a function of separation, using the lagging maneuver (Figure 7.1). If we have a large and fully censused area we can use coarse graining, where we construct a diversity distribution from a large area, then take the average diversity distribution for each half of the area, each quarter, and so on. Because scaling relations for biodiversity are obtained from a variety of scaling maneuvers, care is required in comparing exponents and in applying scaling functions. A spatial scaling exponent obtained by one maneuver is not freely substitutable for an exponent by another maneuver, even for a single quantity. Similarly, a temporal scaling exponent by one maneuver is not necessarily convertible to a scaling exponent by another maneuver.

Scaling relations for biodiversity are generally based on *incomplete similarity* (Barenblatt, 1996), which relates some property $Q(x)$ to some measure x according to a power law (Chapter 6).

$$Q(x) = cx^{\alpha} \tag{16.3a}$$

This relation says that when x is rescaled (say, by a factor of 2), the property $Q(x)$ is still proportional to x^{α} (by a factor of 2^{α}, not just by 2). For species number scaled to area, we have:

$$s(A) = cA^{z} \tag{16.3b}$$

where z is the traditional symbol for the exponent. In the example of African lakes (Ricklefs and Schluter, 1993a), how does fish species number scale with area? A rough scaling is obtained by applying the principle of homogeneity of scope, which requires that all terms in an equation have the same scope. For the cichlid fish, the scope in species number is (200 species lake^{-1})/(40 species lake^{-1}) = 5. The scope for area is (69484 km^2/2150 km^2) = 32. An exponent rescales the scope for area to the scope for diversity:

$$\frac{s(A_{big})}{s(A_{small})} = \left(\frac{A_{big}}{A_{small}}\right)^{z} \tag{16.4a}$$

$$\frac{200}{40} = \left(\frac{69484}{2150}\right)^{z} \tag{16.4b}$$

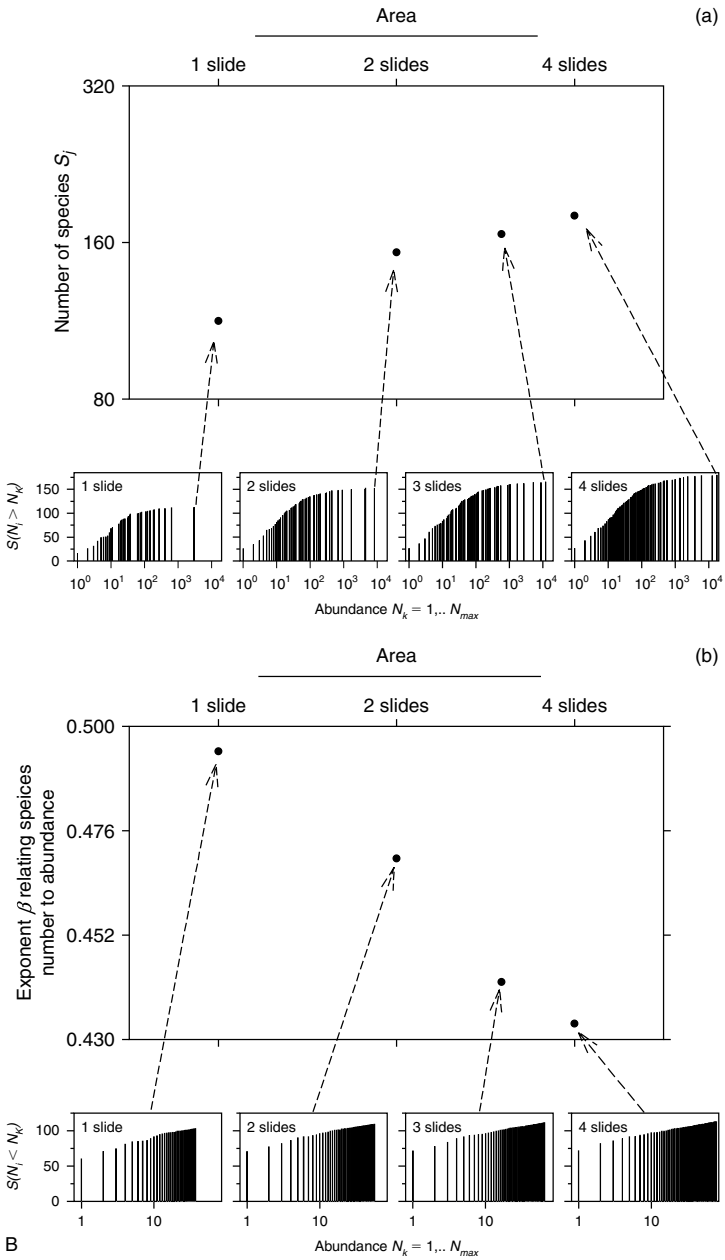


FIGURE 16.2 Species Number, Taxonomic Diversity, and Cumulative Area. Data from Figure 16.1; (a) Relation of species number S_j to cumulative area (number of collections of equal area); (b) Relation taxonomic diversity distribution $S(N_i < N_k)$ to cumulative area.

Solving for the exponent, we obtain:

$$z_{cichlids} = \log(200/40)/\log(69484/2150) = 0.463 \tag{16.5a}$$

Based on this rough estimate of the scaling exponent, a doubling in lake area from, say, 2150 km^2 to 4300 km^2 is expected to amplify cichlid species number by $2^{0.463} = 1.4$. The

expected amplification is less than the factor of 2 expected from isometric scaling. For noncichlid species, the scope in species number is somewhat more than for the highly speciose cichlids:

$$(111 \text{ species lake}^{-1})/(17 \text{ species lake}^{-1}) = 6.53 \tag{16.6a}$$

The scope for area remains the same, so the scaling exponent is also somewhat more than for cichlids:

$$z_{noncichlids} = \log(111/17)/\log(69484/2150) = 0.54 \tag{16.6b}$$

These are rough estimates compared to the regression estimates, which use data from all six lakes.

Incomplete similarity applies to cumulative frequency distributions (Figures 16.1 and 16.2) as well as to single indices such as species number. The scaling based on incomplete similarity is:

$$P[X \leq x] = c \cdot x^{\beta} \tag{16.7}$$

where x is again some measure of interest, such as length, area, or time. $P[X \leq x]$ is the proportion of cases with values X less than or equal to x. An example of such a proportion is the taxonomic distribution as a proportion $s(n_i \leq n_k)/N$. Figure 16.2b shows four scaling relations such as Equation 16.7. When plotted on a log/log scale, the cumulative frequency distribution is approximately a straight line, and hence the frequency distribution can be summarized by a power law, as in Equation 16.7. For a single slide, the exponent of the scaling relation was $\beta = 0.494$ (Figure 16.2b). For two slides combined (a doubling of effort and area), the exponent drops to $\beta = 0.463$. A further doubling of effort (and area) drops the exponent to $\beta = 0.4336$. The decrease falls along a straight line on the log/log plot (note that the intervals on both the X and Y axes are logarithmic, not linear).

Power law frequency distributions (Equation 16.7) not only describe taxonomic diversity distributions (Figure 16.7b), they also describe a variety of geophysical phenomena, such as earthquakes and flood frequencies (Rodriguez-Iturbe and Rinaldo, 1997). It is interesting to note (Figure 16.2b) that the exponent of the power law taxonomic diversity distribution (which characterizes the full frequency distribution) scales in a simple way with cumulative effort (1 slide, 2, slides, 4 slides). For comparison, the scaling of the total species with cumulative effort is shown in Figure 16.2a.

Phenomena that show incomplete scaling, such as the diversity distributions in Figure 16.2, are also examined via renormalization and coarse graining (Barenblatt, 1996). Renormalization has been used to investigate the dynamics of landscapes (Rodriguez-Iturbe and Rinaldo, 1997) and population interaction (Levin and Pacala, p. 271, in Tilman and Kareiva, 1997). Renormalization is an appropriate technique for examining the dynamics of biodiversity, which do not proceed by evenly paced transitions but instead jump and lurch by episodes of invasion, local extinction, speciation, and anthropogenic extinction.

The idea of incomplete similarity has already started to take its place in conservation planning. An example is the "single large *vs*. several small" debate over refuge size,

a debate that recognizes that species number does not scale isometrically with area and recognizes that rates (such as extinction) are not independent of area. Milne (1997) provides several examples of incomplete similarity on conservation biology. What is needed are incomplete scaling exponents that are both reliable and based on biological theory rather than empirical coefficients. One particularly exciting topic, in need of theoretical development, is the scaling of taxonomic diversity to habitat diversity. The conceptual foundation in biological theory already exists. Species lists increase by invasion or speciation, for which adaptation to habitat via natural selection is important. Species lists shrink by local or global extinction, for which change in habitat is important. There is thus the biological basis for quantitative scaling of taxonomic diversity with habitat diversity. It is a promising topic for theoretical development. It is a topic where reliable computations are needed to address questions of local and planet scale change in biodiversity that cannot be ignored.

16.3.2 Theoretical Development of Biodiversity Scaling

The development of scaling relations for biodiversity has been vigorous for taxonomic diversity. It has just begun for habitat diversity. For taxonomic diversity, the history of theoretical development can be summarized briefly as a sequence of scaling relations. The first relation is that species number departs from an isometric scaling with area:

$$S \cong A^{\beta} \tag{16.8}$$

where $\beta \neq 1$ and hence species number scales incompletely with area. Nonisometric scaling of species number with area is one of the oldest and best-known scaling relations in ecology, with a rich history (Ricklefs and Schluter, 1993b; Rosenzweig, 1995). The scaling exponent was recognized to be less than 1 for most situations, but values were completely empirical and could not be generalized beyond the data used to estimate the coefficient. Preston (1962) proposed that species number scales with collection size according to a ¼ power law:

$$S \cong N^{1/4} \tag{16.9}$$

This relation was discussed in detail by May (1975), who provides a list of models. The canonical value (¼) depends on the assumption that the taxonomic diversity distribution has a lognormal form.

In general, we expect numbers to scale with area isometrically:

$$N \cong A^{1} \tag{16.10a}$$

$$N = c \cdot A^{1} \tag{16.10b}$$

where the value of c depends on productivity. This scaling will always hold for coarse graining; we can always compute the average numbers in smaller areas from numbers in larger areas. This scaling will hold on average for accumulation if we take multiple sequences of organism counts from the same area. However, this scaling is often untrue for any one sequence of accruals because of patchiness. More often than otherwise, the average (like the variance in Figure 10.3d) will repeatedly lurch upward as area is added, because dense sites are rare and so more often than not will be sampled later rather than

earlier during accrual. Isometric scaling of numbers N with area A (Equation 16.10a) is best considered a first approximation for islands or other isolated units.

Combining Equations 16.9 and 16.10, we obtain scaling of species number with area according to a ¼ power law:

$$S \cong N^{1/4} \cong A^{1/4} \tag{16.11}$$

Preston's canonical theory (Preston, 1962) and MacArthur and Wilson's theory for isolated communities (MacArthur and Wilson, 1967) attribute the scaling to equilibrium conditions (May, 1975), but it is now known that scaling laws such as Equation 16.11 can arise in nonequilibrium conditions (Barenblatt, 1996). If the scaling exponent in Equation 16.9 depends on ecosystem productivity, we can expect deviation from the scaling in Equation 16.11 (Hubbell, 2001; Storch et al., 2005).

Rosenzweig (1995) describes several well-established patterns in the value of the scaling exponent in Equation 16.8, depending on degree of isolation. Biogeographic provinces, isolated at evolutionary time scales (Rosenzweig, 1995), have scaling exponents that exceed blocks within a province:

$$\beta_{\text{province}} > \beta_{\text{block}} \tag{16.12}$$

Estimates of β_{province} fall closer to unity (Rosenzweig, 1995) than to typical values for β_{block}. The scaling exponent is steeper for islands ($0.25 < \beta_{\text{island}} < 0.33$) than for isolated blocks of land on continents ($0.13 < \beta_{\text{block}} < 0.18$):

$$\beta_{\text{island}} > \beta_{\text{block}} \tag{16.13}$$

Rescue effects, where small populations are maintained by frequent migration from surrounding areas on continents, readily explain the shallow scaling for continental communities (Rosenzweig, 1995). Milne (1997) describes a correction factor to account for scaling effects in comparing area of Euclidean blocks to area of islands with fractal shapes.

Finally, we expect taxonomic diversity to be a function of habitat diversity:

$$S(A_i \leq A_k) = f(H(A_i \leq A_k)) \tag{16.14}$$

Substantial qualitative support exists as correlations of species richness with several habitat variables (Williamson, 1988; Wright et al., p. 73, in Ricklefs and Schluter, 1993b; Rosenzweig, 1995; Gaston, 1996), including fractal measures of habitat complexity (Milne, 1997). An analytic review, similar to that of May (1975), is needed to integrate the taxonomic diversity literature with the literature on the structure and dynamics of landscapes (Rodriguez-Iturbe and Rinaldo, 1997).

16.4 Spatial and Temporal Scaling

Power laws that scale a quantity with body size often arise from theory. In contrast, power laws for spatial and temporal scaling (Figure 2.5) often do not meet Hempel's (1964) criterion—that a theory must state the conditions leading to a result that can be tested against data. One theory that does meet Hempel's criterion is that power laws arise

from local interactions in disorganized physical systems when they near some critical state (Widom, 1965; Wilson, 1971). Power laws are thus expected in systems that are repeatedly forced away from equilibrium conditions (Barenblatt and Sivashinsky, 1969; Barenblatt, 1996) or that lurch rather than move smoothly toward some equilibrium (Bak et al., 1988). This theory, called *universal scaling* or *complexity theory*, applies to animate as well as inanimate systems (Barenblatt and Monin, 1983; Stanley et al., 1996; Milne, 1998). Spatially heterogeneous systems that tend to diverge from equilibrium toward a critical state are called *self-organizing* (Bak et al., 1988). An ecological example of self-organized criticality is the appearance of power law distribution of gaps in the rainforest at Barro Colorado Island in Panama (Sole and Manrubia, 1995). Trees eventually collapse when their size exceeds their mechanical strength; the gaps created by the collapse of a large tree extend to larger areas because vines bind trees to one another. The result is canopy collapse that propagates to scales larger that individual trees and a power law distribution in gap size. In population biology, Levin (2000) summarized examples of large-scale patterns emerging from locally coupled interactions, as in the spread of disease. In geomorphology, Rodriguez-Iturbe and Rinaldo (1997) summarize power law characterizations of landscape structure results from episodic erosion by rivers and streams. Power law (fractal) scalings in landscape data (Burrough, 1981) emerge when erosive dynamics exhibit self-organized criticality (Rodriguez-Iturbe and Rinaldo, 1997).

When and where can we expect power laws emerge? In general, warring exponential rates result in power laws when loss and production occur at different time scales. This can occur as sudden production and slow loss or as slow production and sudden loss. An example is when loss acts on production at a lag rather than simultaneously. An exponential production rate r acting on an initial quantity Q_0 yields an increase relative to Q_0 of:

$$\frac{Q_{produced}}{Q_0} = e^{r \cdot t} \tag{16.15}$$

An exponential loss rate of z acting on Q_0 yields a decrease of:

$$\frac{Q_{remaining}}{Q_0} = e^{-z \cdot t} \tag{16.16}$$

If loss acts with no lag on current stock of the quantity Q, then Q_0 in Equation 16.15 is the same as Q_0 in Equation 16.16, and hence:

$$Q = Q_0 e^{(r-z) \cdot t} \tag{16.17}$$

However, if loss acts on production at some time lag, then Q_0 in Equation 16.15 is not the same as Q_0 in Equation 16.16. For discrete time intervals:

$$\frac{1}{r} \cdot \ln\left(\frac{Q_{produced}}{Q_0}\right) = t = \frac{1}{-z} \cdot \ln\left(\frac{Q_{remaining}}{Q_0'}\right) \tag{16.18a}$$

$$\frac{Q_{produced}}{Q_0} = \left(\frac{Q_{remaining}}{Q_0'}\right)^{-r/z} \tag{16.18b}$$

A large number of familiar problems in ecology are susceptible to the development of theoretical scaling relations using the classical methods described in Chapter 6 and Section 13.8. However, many phenomena in ecology are due to episodically warring rates, which give rise to power laws that are not susceptible to analysis and development of theoretical scaling relations via classical methods (Barenblatt, 1996). An example is the number of species in isolated systems (islands, lakes). Island biogeography (MacArthur and Wilson, 1967) offered an equilibrium theory of loss versus colonization in isolated ecosystems. However, the scaling exponents in MacArthur and Wilson (1967) may be due to nonequilibrium dynamics whereby loss due to extinction acts episodically on species that accumulate via colonization events or evolutionary change. The dynamics of complexity offer a more inclusive theory of island biogeography, whereby either equilibrium or complex behavior pertains, depending on whether or not losses act on gains with a lag.

Another potential example is metapopulation analysis (Levins, 1969), which focuses on the critical point at which recolonization offsets the probability of local extinction. If recolonization and extinction act at the same time scales, an equilibrium theory suffices. If extinction occurs suddenly relative to recolonization (or vice versa), then complexity offers an inclusive theory of the emergence of power law behavior at the scale of semi-isolated patches of habitat.

Landscape ecology focuses on ecological processes against a background of spatial structure such as ecotones (Gardner et al., 2001). The power law (fractal) structure of ecotones suggests that episodically warring rates are at work. Milne (1997) gives an example of an ecotone maintained by gopher versus rabbit interaction. In general, one can expect to find antagonistic rates acting episodically whenever a self-similar power law (such as a fractal) describes habitat structure.

At evolutionary time scales, changes in species number can be analyzed as the outcome of sudden extinction acting on slower speciation (Jablonski, 1989; Bak, 1996). Power law behavior is expected in the record of species number through time.

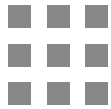
Complexity and power law behavior generated by the interaction of fast with slow processes bring together many of the scaling concepts that emerged in ecology in the late 20th century. Distinguishable concepts identified in Chapter 2 were dependence of pattern on scale of analysis; the absence of any single "characteristic" scale of phenomena; the use of fractal geometry to quantify environmental features; the response of organisms to environmental change at multiple time scales; the interaction of larger scale with local processes to maintain diversity; the propagation of large-scale events to smaller scales through dissipation of energy by fluid processes in the earth's mantle, in the sea, and in the atmosphere; and the extraction of kinetic energy from large-scale fluid motions by mobile organisms to generate local spatial and temporal variability. Multiscale analysis, including fractal concepts, is needed to characterize phenomena that are smeared across a range of scales by the interaction of fast with slow processes. The interaction of fast with slow rates often plays out as the interaction of local with larger-scale processes.

Scaling has become a central concept in ecology (Levin, 1992; Wiens, 1999). Could it become a theoretically unifying concept? The history of power laws in ecology (particularly in body-size scaling) shows that classical scaling methods can go a long way. But the history of power laws in ecology also suggests that scaling based on nonequilibrium dynamics will be needed in order for scaling to become a theoretically unifying concept. Power laws are expected to emerge from antagonistic rates when the history of a

system matters, and exponential losses or production occur episodically. These *cross-scale dynamics* (Levandowsky and White, 1977; Holling, 1992; Carpenter and Turner, 2000) and their resulting power laws (Bak et al., 1988; Barenblatt, 1996; Stanley et al., 1996; Milne, 1998; Schneider, 2001b) have the potential to bring theoretical unity to ecology.

Defined Concepts and Terms for Review and Future Reference

____ allometry
____ biodiversity
____ biodiversity distribution
____ complexity theory (universal scaling)
____ cross scale dynamics
____ diversity index
____ incomplete similarity
____ rank-abundance curve
____ taxonomic frequency distribution



References

Aidley, D.J., 1981. Animal Migration. Cambridge University Press, Cambridge.

Akaike, H., 1973. Information theory as an extension of the maximum likelihood principle. In: Petrov, B.N., Csaki, F. (Eds.), Second International Symposium on Information Theory. Akademiai Kiado, Budapest, pp. 267–281.

Alados, C.L., Emlen, J.M., Wachocki, B., Freeman, D.C., 1998. Instability of development and fractal architecture in dryland plants as an index of grazing pressure. J. Arid Environ. 38, 63–76.

Alerstam, T., 1990. Bird Migration. Cambridge University Press, Cambridge.

Alerstam, T., 1981. The course and timing of bird migration. In: Aidley, D.J. (Ed.), Animal Migration. Cambridge University Press, Cambridge.

Alexander, R.M., 1989. Dynamics of Dinosaurs and Other Extinct Giants. Columbia University Press, New York.

Alexander, R.M., Jayes, A.S., Maloiy, G.M.O., Wathuta, E.M., 1979. Allometry of the limb bones of mammals from shrews (*Sorex*) to elephant (*Loxodonta*). J. Zool. Soc. London 189, 305–314.

Almquist-Jacobson, H., Foster, D.R., 1995. Toward an integrated model for raised-bog development: Theory and field evidence. Ecology 76, 2503–2516.

Allen, T.F.H., Hoekstra, T.W., 1992. Toward a Unified Ecology. Columbia University Press, New York.

Allen, T.F.H., Starr, T.B., 1982. Hierarchy. University of Chicago Press, Chicago.

Allen, R.G.D., 1939. The assumptions of linear regression. Economica 6, 191–201.

Anderson, R.M., Gordon, D.M., Crawley, M.J., Hassell, M.P., 1982. Variability and the abundance of animal and plant species. Nature 296, 245–248.

Anderson, A.R.A., Young, I.M., Sleeman, B.D., Griffiths, B.S., Robertson, W.M., 1997. Nematode movement along a chemical gradient in a structurally heterogeneous environment, 1. Experiment. Fundam. Appl. Nematol. 20, 157–163.

Andrews, D.F., Herzberg, A.M., 1985. Data: A Collection of Problems from Many Fields for the Student and Research Worker. Springer-Verlag, New York.

Anscombe, F.J., 1948. The transformation of Poisson, binomial and negative-binomial data. Biometrika 35, 246–254.

Arnold, G.P., Cook, P.H., et al., 1984. Fish migration by selective tidal stream transport: First results with a computer simulation model for the European continental shelf. In: McCleave, J.D. (Ed.), Mechanisms of Migration in Fishes. Plenum Press, New York, pp. 227–261.

Arrhenius, O., 1921. Species and area. J. Ecol. 9, 95–99.

Auble, G.T., Hamilton, D.B., Roelle, J.E., Andrews, A.K., 1995. Simulation modeling in a workshop format. In: Patten, B.C. (Ed.), Complex Ecology. Prentice-Hall, New York, pp. 311–334.

Avnir, D., Biham, O., Lidar, D., Malcai, O., 1998. Is the geometry of nature fractal? Science 279, 39–40.

Bak, P., 1996. How Nature Works: The Science of Self-Organized Criticality. Springer, New York.

Bak, P., Chen, K., 1995. Fractal dynamics of earthquakes. In: Barton, C.C., LaPointe, P.R. (Eds.), Fractals in the Earth Sciences. Plenum Press, New York, pp. 227–236.

Bak, P., Tang, C., Wiesenfeld, K., 1988. Self-organized criticality. Phys. Rev. A 38, 364–374.

Baker, M.C., 1974. Foraging behavior of black-bellied plovers (*Pluvialis squatarola*). Ecology 55, 162–167.

Banavar, J.R., Maritan, A., Rinaldo, A., 1999. Size and form in efficient transportation networks. Nature 399, 130–132.

Banse, K., Mosher, S., 1980. Adult body mass and annual production/biomass relationships of field populations. Ecol. Monog. 50, 355–379.

Barenblatt, G.I., 1996. Scaling, Self-Similarity, and Intermediate Asymptotics. Cambridge University Press, Cambridge.

Barenblatt, G.I., Monin, A.S., 1982. Similarity principles for the biology of pelagic animals. Proc. Nat. Acad. Sci. US 80, 3540–3542.

Barenblatt, G.I., Sivashinsky, G.I., 1969. Self-similar solutions of the second kind in nonlinear filtration. Appl. Math. Mech. (PMM) 33, 836–845.

Barry, J.P., Dayton, P.K., 1991. Physical heterogeneity and the organization of marine communities. In: Kolasa, J., Pickett, S.T.A. (Eds.), Ecological Heterogeneity. Springer-Verlag, New York, pp. 270–320.

Bascompte, J., Vilà, C., 1997. Fractals and search paths in mammals. Landscape Ecol. 12, 213–221.

Batchelor, G.K., 1967. An Introduction to Fluid Mechanics. Cambridge University Press, Cambridge.

Baudry, J., 1993. Landscaping dynamics and farming systems: Problems of relating patterns and predicting ecological changes. In: Bunce, R.G.H., Ryszkowski, L., Paoletti, M.G. (Eds.), Landscape Ecology and Agroecosystems. Lewis Publishers, London, pp. 21–40.

Beck, M.W., 1998. Comparison of the measurement and effects of habitat structure on gastropods in rocky intertidal and mangrove habitats. Mar. Ecol.-Prog. Ser. 169, 165–178.

Bell, D.T., 1997. Eighteen years of change in an Illinois streamside deciduous forest. J. Torrey Bot. Soc. 124, 174–188.

Bell, S.S., McCoy, E.D., Mushinsky, H.R., 1991. Habitat Structure. Chapman and Hall, London.

Bennett, A.F., Denman, K.L., 1985. Phytoplankton patchiness: Inferences from particle statistics. J. Mar. Res. 43, 307–335.

Bidder, G.P., 1931. The biological importance of Brownian movement (with notes on sponges and Protista). Proc. Linn. Soc. London 143, 82–96.

Blackman, G.E., 1942. Statistical and ecological studies in the distribution of species in plant communities: I. Dispersion as a factor in the study of changes in plant populations. Ann. Bot. 6, 351–370.

Blem, C.R., 1980. The energetics of migration. In: Gauthreaux, S.R. (Ed.), Animal Migration, Orientation, and Navigation. Academic Press, New York, pp. 175–224.

Bliss, C.I., 1941. Statistical problems in estimating populations of Japanese beetle larvae. J. Econ. Entomol. 34, 221–232.

Boehlert, G.W., Mundy, B.C., 1988. Roles of behavioral and physical factors in larval and juvenile fish recruitment to estuarine nursery areas. Am. Fish. Soc. Symp. 3, 51–67.

Bonner, J.T., 1965. Size and Cycle. Princeton University Press, Princeton, New Jersey.

Bormann, F.H., Likens, G.E., 1979. Pattern and Process in a Forested Ecosystem. Springer-Verlag, New York.

Bradbury, R.H., Reichelt, R.E., Green, D.G., 1984. Fractals in ecology: methods and interpretation. Mar. Ecol.-Prog. Ser. 14, 295–296.

Brattey, J., Cadigan, N.D., Lilly, G.R., Murphy, E.F., Shelton, P.A., Stansbury, D.E. 1999. An Assessment of the Cod Stock in NAFO Subdivision 3 PS in October 1999. DFO Canadian Stock Assessment Secretariat Research Document 1999/161, Ottawa.

Bridgman, P.W., 1922. Dimensional Analysis. Yale University Press, New Haven.

Brody, S., 1945. Bioenergetics and Growth. Reinhold, New York.

Brussard, P.F., 1984. Geographic patterns and environmental gradients: The central-marginal model in *Drosophila* revisited. Ann. Rev. Ecol. Syst. 15, 25–64.

Bult, T.P., Haedrich, R.L., Schneider, D.C., 1998. New technique describing spatial scaling and habitat selection in riverine habitats. Regul. Riv.: Res. Mgmt. 14, 107–118.

Bunge, J., Fitzpatrick, M., 1993. Estimating the number of species: A review. J. Am. Stat. Assoc. 88, 364–373.

Burger, A.E., 1993. Estimating the mortality of seabirds following oil spills: Effects of spill volume. Mar. Pollut. Bull. 26, 140–143.

Burnham, K.P., Anderson, D.R., 1998. Model Selection and Inference: A Practical Information-Theoretic Approach. Springer-Verlag, New York.

Burrough, P.A., 1981. Fractal dimensions of landscapes and other environmental data. Nature 294, 240–242.

Byerly, K.F., Gutierrez, A.P., Jones, R.E., Luck, R.F., 1978. Comparison of sampling methods for some arthropod populations in cotton. Hilgardia 46, 257–282.

Cajori, F., 1929. A History of Mathematical Notations. The Open Court Publishing Company, Chicago.

Calder, W.A., 1983. Ecological scaling: mammals and birds. Ann. Rev. Ecol. Syst. 14, 213–230.

Calder, W.A., 1984. Size, Function, and Life History. Harvard University Press, Cambridge.

Campbell, N.R., 1942. Dimensions and the facts of measurement. Philosophical Magazine 33, 761–771.

Cantero, J.J., Leon, R., Cisneros, J.M., Cantero, A., 1998. Habitat structure and vegetation relationships in central Argentina salt marsh landscapes. Plant Ecol. 137, 79–100.

Carpenter, S.R., 1990. Large-scale perturbations: opportunities for innovation. Ecology 71, 2038–2043.

Carpenter, S.R., Turner, M.G., 2000. Hares and tortoises: interactions of fast and slow variables in ecosystems. Ecosystems 3, 495–497.

Case, T., 2000. An Illustrated Guide to Theoretical Ecology. Cambridge University Press, Cambridge.

Charnov, E.L., 1976. Optimal foraging: the marginal value theorem. Theor. Popul. Biol. 9, 126–136.

Chen, C.C., Petersen, J.E., Kemp, W.M., 1997. Spatial and temporal scaling of periphyton growth on walls of estuarine mesocosms. Mar. Ecol.-Prog. Ser. 155, 1–15.

Clapham, A.R., 1936. Over-dispersion in grassland communities and the use of statistical methods in plant ecology. J. Ecol. 24, 232–251.

Claussen, C.L, Finkler, M.S., Smith, M.M., 1997. Thread trailing of turtles: methods for evaluating spatial movements and pathway structure. Can. J. Zool. 75, 2120–2128.

Cliff, A.D., Ord, J.K., 1974. Spatial Autocorrelation. Pion Limited, London.

Cochran, W.G., Cox, G.M., 1957. Experimental Designs reissued 1992. John Wiley & Sons, New York.

Cochran, W.G., 1977. Sampling Techniques, third ed. Wiley & Sons, New York.

Cochran, W.G., 1947. Some consequences when the assumptions for the analysis of variance are not satisfied. Biometrics 3, 22–38.

Collins, J.W., 1884. Notes on the habits and methods of capture of various species of sea birds that occur on the fishing banks off the eastern coast of North America, and which are used as bait for catching codfish by New England fisherman. Rep. Comm. Fish and Fisheries for 1882 13, 311–335.

Crawford, S.M., Sanford, L.P., 2001. Boundary shear velocities and fluxes in the MEERC experimental ecosystems. Mar. Ecol.-Prog. Ser. 210, 1–12.

Crawley, M., 1993. GLIM for Ecologists. Blackwell, London.

Cressie, N.A.C., 1991. Statistics for Spatial Data. John Wiley & Sons, New York.

Cronin Jr., T.W., Forward, R.B., 1979. Tidal vertical migration: an endogenous rhythm in estuarine crab larvae. Science 205, 1020–1022.

Csanady, G., 1982. Circulation in the Coastal Ocean. D. Reidel Publishing Company, Dordrecht.

Cushman, J.H., 1986. On measurement, scale, and scaling. Water Resour. Res. 22, 129–134.

Damuth, J., 1981. Population density and body size in mammals. Nature 290, 699–700.

Darnell, R.M., 1971. Organism and Environment; A Manual of Quantitative Ecology. W. H. Freeman, San Francisco.

Davenport, J., Pugh, P.J.A., McKechnie, J., 1996. Mixed fractals and anisotropy in subantarctic marine macroalgae from South Georgia: Implications for epifaunal biomass and abundance. Mar. Ecol.-Prog. Ser. 136, 245–255.

Davis, J.C., 1986. Statistics and Data Analysis in Geology, second ed. John Wiley & Sons, New York.

Davis, R.B., Anderson, D.S., 1991. The eccentric bogs of Maine. In: Maine Agricultural Experiment Station Technical Bulletin, 146. University of Maine, Orono, pp. 1–151.

Dayton, P.D., Tegner, M.J., 1984. The importance of scale in community ecology: A kelp forest example with terrestrial analogs. In: Price, P.W., Slobodchikoff, C.M., Gaud, W.S. (Eds.), A New Ecology: Novel Approaches to Interactive Systems. John Wiley & Sons, New York, pp. 457–481.

Delcourt, H.R., Delcourt, P.A., Webb, T., 1983. Dynamic plant ecology: the spectrum of vegetational change in space and time. Quat. Sci. Rev. 1, 153–175.

Despland, E., Collett, M., Simpson, S.J., 2000. Small-scale processes in desert locust swarm formation: how vegetation patterns influence gregarization. Oikos 88, 652–662.

Deutschman, D.H., Bradshaw, G.A., Childress, W.M., Daly, K.L., Grünbaum, D., Pascual, M., Schumaker, N.H., Wu, J., 1993. Mechanisms of Patch formation. In: Steele, J.H., Levin, S., Powell, T. (Eds.), Patch Dynamics. Springer-Verlag, Berlin, pp. 184–209.

Diamond, J.M., 1975. The island dilemma: lessons of modern biogeographic studies for the design of nature reserves. Biol. Conserv. 7, 129–146.

Dicke, M., Burrough, P.A., 1988. Using fractal dimensions for characterizing tortuosity of animal trails. Phys. Entomol. 13, 393–398.

Dickie, L., Kerr, S.R., Boudreau, P., 1987. Size-dependent processes underlying regularities in ecosystem structure. Ecology 57, 233–250.

Dickinson, R.E., 1988. Atmospheric systems and global change. In: Risser, P.G., Woodmansee, R.G., Rosswall, T. (Eds.), Scales and Global Change. John Wiley & Sons, New York, pp. 57–80.

Dobzhansky, T., 1948. Genetics of natural populations, XVI, Altitudinal and seasonal changes produced by natural selection in natural populations of *Drosophila pseudoobscura* and *Drosophila persimilis*. Genetics 33, 158–176.

Downing, J.A., Osenberg, C.W., Sarnelle, O., 1999. Meta-analysis of marine nutrient-enrichment experiments: variation in the magnitude of nutrient limitation. Ecology 80, 1157–1167.

Draper, N.R., Smith, H., 1981. Applied Regression Analysis. John Wiley, New York.

Duffy, D.C., Schneider, D.C., 1994. Seabird-fishery interactions: a manager's guide. Birdlife Conser. Ser. 1, 26–38.

Dungan, J.L., 2001. Scaling up and scaling down: the relevance of the support effect on remote sensing of vegetation. In: Tate, N.J., Atkinson, P.M. (Eds.), Modeling Scale in Geographical Information Science. John Wiley, New York, pp. 221–235.

Dutilleul, P., 1993. Spatial heterogeneity and the design of ecological field experiments. Ecology 74, 1646–1658.

Dutton, J.A., 1975. The Ceaseless Wind. McGraw-Hill, New York.

Eberhardt, L.L., Thomas, J.M., 1991. Designing environmental field studies. Ecol. Monog. 61, 53–73.

Egler, F.E., 1942. Vegetation as an object of study. Philos. Sci. 9, 245–260.

Eisenhart, C., 1947. The assumptions underlying the analysis of variance. Biometrics 3, 1–21.

Ellis, B., 1966. Basic Concepts of Measurement. Cambridge University Press, Cambridge.

Ellis, J.I., Schneider, D.C., 1997. Evaluation of a gradient sampling design for environment impact assessment. Environ. Monit. Assess. 48, 157–172.

Ellis, J.I., Schneider, D.C., 2008. Spatial and temporal scaling in benthic ecology. J. Exper. Mar. Biol. Ecol. 366, 92–98.

Elton, C., 1924. Periodic fluctuations in the numbers of animals: their causes and effects. Br. J. Exp. Biol. 2, 119–163.

Enright, J.T., 1965. The search for rhythmicity in biological time-series. J. Theor. Biol. 8, 426–468.

Erlandsson, J., Kostylev, V., 1995. Trail following, speed and fractal dimension of movement in a marine prosobranch, *Littorina littorea*, during a mating and a non-mating season. Mar. Biol. 122, 87–94.

Erwin, R.M., 1977. Foraging and breeding adaptations to different food regimes in three seabirds: The Common Tern, *Sterna hirundo*, Royal Tern, *Sterna maxima*, and Black Skimmer, *Rynchops niger*. Ecology 58, 389–397.

Escós, J., Alados, C.L., Emlen, J.M., 1997. The impact of grazing on plant fractal architecture and fitness of a Mediterranean shrub *Anthyllis cytisoides* L. Funct. Ecol. 11, 66–78.

Etzenhouser, M.J., Owens, M.K., Spalinger, D.E., Murden, S.B., 1998. Foraging behavior of browsing ruminants in an heterogeneous landscape. Landscape Ecol. 13, 55–64.

Falconer, K.J., 1985. The Geometry of Fractal Sets. Cambridge University Press, Cambridge.

Fasham, M.J.R., 1978. The statistical and mathematical analysis of plankton patchiness. Oceanogr. Mar. Biol. Ann. Rev. 16, 43–79.

Fee, E.J., 1979. A relation between lake morphometry and primary productivity and its use in interpreting whole-lake eutrophication experiments. Limnol. Oceanogr. 24, 401–416.

Ferguson, S.H., Taylor, M.K., Born, E.W., Messier, F., 1998. Fractals, sea-ice landscape and spatial patterns of polar bears. J. Biogeogr. 25, 1081–1092.

Firbank, L.G., 1991. The implications of scale on the ecology and management of weeds. In: Bunce, R.G.H., Ryszkowski, L., Paoletti, M.G. (Eds.) Landscape Ecology and Agroecosystems. Lewis Publishers, Boca Raton, pp. 91–103.

Fisher, R.A., 1954. Statistical Methods for Research Workers, Twelfth ed. Oliver and Boyd, Edinburgh.

Fisher, R.A., Thornton, H.G., Mackenzie, W.A., 1922. The accuracy of the plating method of estimating the density of bacterial populations. Ann. Appl. Biol. 9, 325–359.

Fonteyn, P.J., Mahall, B.E., 1981. An experimental analysis of structure in a desert plant community. J. Ecol. 69, 883–896.

Fourier, J.B.J., 1822. The Analytical Theory of Heat, Translated 1878 by A. Freeman. Dover edition 1955.

Frank, K.T., Leggett, W.C., 1983. Multispecies larval fish associations: accident or adaptation? Can. J. Fish. Aquat. Sci. 40, 754–762.

Frontier, S., 1987. Applications of fractal theory to ecology. In: Legendre, P., Legendre, L. (Eds.), Developments in Numerical Ecology. Springer-Verlag, Berlin, pp. 335–378.

Fuller, D.O., 1998. Trends in NDVI time series and their relation to rangeland and crop production in Senegal, 1987–1993. Int. J. Remote. Sens. 19, 2013–2018.

Gadgil, M., 1996. Genes, memes, and artefacts. In: di Castri, F., Younes, T. (Eds.), Biodiversity, Science and Development. CAB International, pp. 80–90.

Galileo, 1638, Dialogues Concerning Two New Sciences, Translated 1914 by H. Crew and A. de Salvio, Dover edition 1954.

Gallegos, C.L., Platt, T., 1985. Vertical advection of phytoplankton and productivity estimates: a dimensional analysis. Mar. Ecol.-Progr. Ser. 26, 125–134.

Gardner, R.H., Kemp, W.M., Peterson, J., Kennedy, V. (Eds.), 2001. Scaling Relations in Experimental Ecology. Columbia University Press, New York.

Garland, T., 1983. Scaling the ecological cost of transport to body mass in terrestrial mammals. Am. Nat. 121, 571–587.

Gaston, K.J. (Ed.), 1996. Biodiversity. Blackwell, Oxford.

Gause, G., 1934. The Struggle for Existence. Haeffner, New York.

Gee, J.M., Warwick, R.M., 1994. Body-size distribution in a marine metazoan community and the fractal dimensions of macroalgae. J. Exp. Mar. Bio. Ecol. 178, 247–259.

Gleason, H.A., 1922. On the relation between species and area. Ecology 3, 158–162.

Gold, H.J., 1977. Mathematical Modeling of Biological Systems—An Introductory Guidebook. John Wiley & Sons, New York.

Golley, F.B., 1993. A History of the Ecosystem Concept in Ecology. Yale University Press, New Haven.

Gould, S.J., 1966. Allometry and size in ontogeny and phylogeny. Biol. Rev. 41, 587–640.

Gould, S.J., 1971. Geometric similarity in allometric growth: a contribution to the problem of scaling in the evolution of size. Am. Nat. 105, 113–136.

Gower, J.C., 1987. Introduction to ordination techniques. In: Legendre, P., Legendre, L. (Eds.), Developments in Numerical Ecology. Springer-Verlag, Berlin, pp. 3–64.

Grassle, J.F., Sanders, H.L., Hessler, R.R., Rowe, G.T., McLellan, T., 1975. Pattern of zonation: A study of the bathyal megafauna using the submersible Alvin. Deep-Sea Res. 22, 457–481.

Green, R.H., 1979. Sampling Design and Statistical Methods for Environmental Biologists. John Wiley, New York.

Greenwood, M., Yule, G.U., 1920. An inquiry into the nature of frequency distributions representative of multiple happenings with particular reference to the occurrence of multiple attacks of disease or of repeated accidents. J. R. Stat. Soc. 83, 255–279.

Greig-Smith, P., 1952. The use of random and contiguous quadrats in the study of the structure of plant communities. Ann. Bot. 16, 293–316.

Greig-Smith, P., 1983. Quantitative Plant Ecology, third ed. Blackwell, London.

Greig-Smith, P., Chadwick, M.J., 1965. Data on pattern within plant communities, III. Acacia-Capparis semi-desert scrub in the Sudan. J. Ecol. 53, 465–474.

Gunther, B., 1975. Dimensional analysis and theory of biological similarity. Physiol. Rev. 55, 659–698.

Gutierrez, A.P., Summers, C.G., Baumgaertner, J., 1980. The phenology and distribution of aphids in California alfalfa as modified by ladybird beetle predation. Can. Entomol. 112, 489–495.

Hairston, N.G., 1989. Ecological Experiments. Purpose, Design, and Execution. Cambridge University Press, Cambridge.

Hale, W.G., 1980. Waders. Collins, London.

Hall, F.G., Strebel, D.E., Sellers, P.J., 1988. Linking knowledge among spatial and temporal scales: Vegetation, atmosphere, climate and remote sensing. Landscape Ecol. 2, 3–22.

Hamner, W.M., Schneider, D.C., 1986. Regularly spaced rows of medusae in the Bering Sea: Role of Langmuir circulation. Limnol. Oceanogr. 31, 171–177.

Hancock, J. (2000). Along Shore Movement of Juvenile Cod Determined by Mark-Recapture. B.Sc. Honours Thesis, St. John's, Canada: Memorial University of Newfoundland.

Harden-Jones, F.R., Walker, M.G., Arnold, G.P., 1978. Tactics of fish movement in relation to migration strategy and water circulation. In: Charnock, H., Deacon, G. (Eds.), Advances in Oceanography. Plenum Press, New York, pp. 185–207.

Harris, G.P., 1980. Temporal and spatial scales in phytoplankton ecology, Mechanisms, methods, models, and management. Can. J. Fish. Aquat. Sci. 37, 877–900.

Hart, I.B., 1923. Makers of Science. Macmillan, London.

Hassell, M.P., 1978. The Dynamics of Arthropod Predator-Prey Systems. Princeton University Press, Princeton.

Hassell, M.P., Comins, H.N., May, R.M., 1991. Spatial structure and chaos in insect population dynamics. Nature 353, 255–258.

Hastings, H.M., Sugihara, G., 1993. Fractals: a User's Guide for the Natural Sciences. Oxford University Press, Oxford.

Hatcher, A.I., Frith, C.A., 1985. The control of nitrate and ammonium concentrations in a coral reef lagoon. Coral Reefs 4, 101–110.

Haury, L.R., McGowan, J.S., Wiebe, P., 1978. Patterns and processes in the time-space scales of plankton distributions. In: Steele, J. (Ed.), Spatial Pattern in Plankton Communities. Plenum Press, New York, pp. 277–327.

Hausdorf, F., 1919. Dimension und äußeres Mass. Math. Ann. 79, 157–179.

Hempel, C.G., 1964. Philosophy of Natural Science. Prentice-Hall, Englewood Cliffs, New Jersey.

Hengeveld, H., 1990. Dynamic Biogeography. Cambridge University Press, Cambridge.

Hill, A.V., 1950. The dimensions of animals and their muscular dynamics. Sci. Prog. 38, 209–229.

Hill, M.O., 1973. The intensity of spatial pattern in plant communities. J. Ecol. 61, 225–235.

Hobbs, R.J., 1998. Managing ecological systems and processes. In: Peterson, D.L., Parker, V.T. (Eds.), Ecological Scale: Theory and Applications. Columbia University Press, New York, pp. 459–484.

Hoffman, W., Heinemann, D., Wiens, J.A., 1981. The ecology of seabird feeding flocks in Alaska. Auk 98, 437–456.

Holling, C.S., 1959. The components of predation as revealed by a study of small mammal predation of the European pine sawfly. Can. Entomol. 91, 293–320.

Holling, C.S., 1992. Cross-scale morphology, geometry, and dynamics of ecosystems. Ecol. Monog. 62, 447–502.

Horne, J.K., Schneider, D.C., 1995. Spatial variance in ecology. Oikos 74, 321–329.

Horne, J.K., Schneider, D.C., 1997. Spatial variance of mobile marine organisms: Capelin and cod in Newfoundland coastal waters. Philos. Trans. R. Soc. London 352, 633–642.

Horton, R.E., 1945. Erosional development of streams and their drainage basins: Hydrophysical approach to quantitative morphology. Bull. Geol. Soc. Am. 56, 275–370.

Horwood, J.W., Cushing, D.H., 1978. Spatial distributions and ecology of pelagic fish. In: Steele, J.H. (Ed.), Spatial Pattern in Plankton Communities. Plenum Press, New York, pp. 355–382.

Hotelling, H., 1960. The behaviour of some standard statistical tests under non-standard conditions. Proc., 4th Berkeley Symp. Math. Stat. Probab. 1, 319–360.

Houser, A., Dunn, J.E., 1967. Estimating the size of the thread-fin shad population in Bull Shoal reservoir from midwater trawl catches. Trans. Am. Fish. Soc. 96, 176–184.

Hubbell, S.P., 2001. The Unified Neutral Theory of Biodiversity and Biogeography. Princeton University Press, Princeton.

Hurlbert, S.H., 1984. Pseudoreplication and the design of ecological field experiments. Ecol. Monog. 54, 187–211.

Hutchinson, G.E., 1951. Copepodology for the ornithologist. Ecology 32, 571–577.

Hutchinson, G.E., 1971. Banquet address: Scale effects in ecology. In: Patil, G.P., Pielou, E.C., Waters, W.E. (Eds.), Spatial Patterns and Statistical Distribution (Statistical Ecology, Vol. I). The Pennsylvania State University Press, University Park, pp. xvii–xxii.

Huxley, J.S., 1932. Problems of Relative Growth. Methuen, New York.

Innes, J.L., 1998. Measuring environmental change. In: Peterson, D.L., Parker, V.T. (Eds.), Ecological Scale, Theory and Applications. Columbia University Press, New York, pp. 429–457.

Ivlev, V.S., 1961. Experimental Ecology of the Feeding of Fishes, Translated from the Russian by D. Scott. Yale University Press, New Haven.

Iwao, S., 1968. A new regression method for analyzing the aggregation patterns of animal populations. Res. Pop. Ecol. 10, 1–20.

Jablonski, D., 1989. The biology of mass extinction: a palaeontological view. Philos. Trans. R. Soc. of London 325, 357–368.

Jagers op Akkerhuis, G.A.J.M., 2008. Analysing hierarchy in the organization of biological and physical systems. Biol. Rev. 83, 1–12.

Jardine, N., Sibson, R., 1971. Mathematical Taxonomy. John Wiley & Sons, New York.

Jassby, A.D., Powell, T.M., 1990. Detecting changes in ecological time series. Ecology 71, 2044–2052.

Johnson, A.R., Milne, B.T., Wiens, J.A., 1992. Diffusion in fractal landscapes: Simulations and experimental studies of tenebrionid beetle movements. Ecology 73, 1968–1983.

Johnstone, J., 1908. Conditions of Life in the Sea: A Short Account of Quantitative Marine Biological Research. Cambridge University Press, Cambridge Reprinted in 1977 by Arno Press, New York.

Jones, J.R., Hoyer, M.V., 1982. Sportfish harvest predicted by summer chlorophyll: a concentration in midwestern lakes and reservoirs. Trans. Am. Fish. Soc. 111, 176–179.

Jorge, L.A.B., Garcia, G.J., 1997. A study of habitat fragmentation in southeastern Brazil using remote sensing and geographic information systems (GIS). For. Ecol. Manage. 98, 35–47.

Journel, A.G., Huijbregts, C.J., 1978. Mining Geostatistics. Academic Press, London.

Kaandorp, J.A., 1991. Modelling growth forms of the sponge *Haliclona oculata* (*Porifera, Demospongiae*) using fractal techniques. Mar. Biol. 110, 203–215.

Kachanoski, R.G., 1988. Processes in soils: from pedon to landscape. In: Rosswall, T., Woodmansee, R.G., Risser, P.G. (Eds.), Scales and Global Change. John Wiley & Sons, New York, pp. 153–177.

Kareiva, P., 1989. Renewing the dialogue between theory and experiments in population ecology. In: Roughgarden, J., May, R., Levin, S.A. (Eds.), Perspectives in Ecological Theory. Princeton University Press, Princeton, New Jersey, pp. 68–88.

Kareiva, P., Anderson, M., 1988. Spatial aspects of species interactions: the wedding of models and experiments. In: Hastings, A. (Ed.), Community Ecology, Lecture Notes in Biomathematics 77. Springer-Verlag, Berlin, pp. 35–50.

Kemp, W.M., Petersen, J.E., Gardner, R.H., 2001. Scale-dependence and the problem of extrapolation. In: Gardner, R.H., Kemp, W.M., Kennedy, V.S., Petersen, J.E. (Eds.), Scaling Relation in Experimental Ecology. Columbia University Press, New York, pp. 3–57.

Kermack, K.A., Haldane, J.B.S., 1950. Organic correlation and allometry. Biometrika 37, 30–41.

Kershaw, K.A., 1957. The use of cover and frequency in the detection of pattern in plant communities. Ecology 38, 291–299.

Kershaw, K.A., Looney, J.H.H., 1985. Quantitative and Dynamic Plant Ecology. Edward Arnold, London.

Kierstead, H., Slobodkin, L.B., 1953. The size of water masses containing plankton blooms. J. Mar. Res. 12, 141–147.

Kilpatrick, A.M., Ives, A.R., 2003. Species interactions can explain Taylor's power law for ecological time series. Nature 422, 65–68.

King, D.A., 1991. Correlations between biomass allocation, relative growth rate and light environment in tropical forest saplings. Funct. Ecol. 5, 485–492.

King, J.R., 1974. Seasonal allocation of time and energy resources in birds. In: Paynter, R.A. (Ed.), Avian Energetics. Nuttall Ornithological Club, Cambridge, Massachusetts, pp. 4–85.

Kleiber, M., 1947. Body size and metabolic rate. Physiol. Rev. 27, 511–541.

Kline, S.J., 1965. Similitude and Approximation Theory. McGraw-Hill, New York.

Kolmogorov, A.N., 1941. Dissipation of energy in a locally isotropic turbulence. Dokl. Akad. Nauk SSSR 30, 301–305, 1991 English translation in *Proceedings Royal Society London* A: 434: 9–13.

Kolmogorov, A.N., 1962. A refinement of previous hypotheses concerning the local structure of turbulence in a viscous incompressible fluid at high Reynolds number. J. Fluid Mech. 13, 82–85.

Kooyman, G.L., Davis, R.W., Croxall, J.P., Costa, D.P., 1982. Diving depths and energy requirements of King Penguins. Science 217, 726–727.

Korvin, G., 1992. Fractal Models in the Earth Sciences. Elsevier, Amsterdam.

Kovacs, K.M., Lavigne, D.M., 1985. Neonatal growth and organ allometry of Northwest Atlantic harp seals (*Phoca groenlandica*). Can. J. Zool. 63, 2793–2799.

Krantz, D.H., Luce, R.D., Suppes, P., Tversky, A., 1971. Foundations of Measurement. Academic Press, New York.

Krebs, C.J., 1972. Ecology: The Experimental Analysis of Distribution and Abundance. Harper and Row, New York.

Krebs, C.J., 1999. Ecological Methodology. Addison Wesley Longman, New York.

Krummel, J.R., Gardner, R.H., Sugihara, G., O'Neill, R.V., Coleman, P.R., 1987. Landscape patterns in a disturbed environment. Oikos 48, 321–324.

Kullback, S., Leibler, R.A., 1951. On information and sufficiency. Ann. Math. Stat. 22, 79–86.

Kurlansky, M., 1997. Cod: A Biography of the Fish That Changed the World. Alfred A. Knopf, Toronto.

Kyburg, H.E., 1984. Theory and Measurement. Cambridge University Press, Cambridge.

Lande, R., 1993. Risks of population extinction from demographic and environmental stochasticity, and random catastrophes. Am. Nat. 142, 911–927.

Langhaar, H.L., 1951. Dimensional Analysis and the Theory of Models. John Wiley & Sons, New York.

Langton, R.W., Auster, P.J., Schneider, D.C., 1995. A spatial and temporal perspective on research and management of groundfish in the northwest Atlantic. Rev. Fish. Sci. 3, 201–229.

Legendre, P., Fortin, M.-J., 1989. Spatial pattern and ecological analysis. Vegatatio 80, 107–138.

Legendre, P., Legendre, L., 1998. Numerical Ecology. Elsevier, Amsterdam.

Legendre, P., Thrush, S.F., Cummings, V.J., Dayton, P.K., Grant, J., Hewitt, J.E., Hines, A.H., McArdle, B.H., Pridmore, R.D., Schneider, D.C., Turner, S.J., Whitlatch, R.B., Wilkinson, M.R., 1997. Spatial structure of bivalves in a sandflat: scale and generating processes. J. Exp. Mar. Biol. Ecol. 216, 99–128.

Lems, K., 1960. Floristic botany of the Canary Islands. Sarracenia 5, 1–94.

Levandowsky, M., White, B.S., 1977. Randomness, time scales and the evolution of biological communities. In: Evolutionary Biology, 10. Plenum Press, New York, pp. 69–161.

Levin, S.A., 1976. Population dynamic models in heterogeneous environments. Ann. Rev. Ecol. Syst. 7, 287–311.

Levin, S.A., 1981. Mechanisms for the generation and maintenance of diversity. In: Hiorns, R.W., Cooke, D. (Eds.), The Mathematical Theory of the Dynamics of Biological Populations. Academic Press, London, pp. 173–194.

Levin, S.A., 1992. The problem of pattern and scale in ecology. Ecology 73, 1943–1967.

Levin, S.A., 2000. Multiple scales and the maintenance of biodiversity. Ecosystems 3, 498–506.

Levin, S.A., Morin, A., Powell, T.M., 1989. Patterns and processes in the distribution and dynamics of Antarctic krill. Select. Sci. Pap., Convent. Antarctic Mar. Living Res. VII/BG/20, 281–296.

Levin, S.A., Paine, R.T., 1974. Disturbance, patch formation, and community structure. Proc. Nat. Acad. Sci. US 71, 2744–2747.

Levins, R., 1969. Some demographic and genetic consequences of environmental heterogeneity for biological control. Bull. Entomol. Soc. Am. 15, 237–240.

Lewis, M.R., Cullen, J.J., Platt, T., 1984. Relationships between vertical mixing and photoadaptation of phytoplankton: similarity criteria. Mar. Ecol.-Prog. Ser. 15, 141–149.

Lewontin, R.C., 1965. Selection for colonizing ability. In: Baker, H.G., Stebbins, G.L. (Eds.), The Genetics of Colonizing Species. Academic Press, New York, pp. 77–94.

Liddicker, W.Z., 2008. Levels of organization in biology: on the nature and nomenclature of ecology's fourth level. Biol.Rev. 83, 71–78.

Likens, G.E., Bormann, F.H., Johnson, N.M., Fisher, D.W., Pierce, R.S., 1970. Effects of forest cutting and herbicide treatment on nutrient budgets in the Hubbard Brook watershed-ecosystem. Ecol. Monog. 40, 23–47.

Lloyd, M., 1967. Mean Crowding. J. Anim. Ecol. 36, 1–30.

Lodge, D.M., Blumenshine, S.C., Vadeboncoeur, Y., 1998. Insights and applications of large-scale, long-term ecological observations and experiments. In: Resetarits, W.J., Bernardo, J. (Eds.), Experimental Ecology. Oxford Univ. Press, Oxford, pp. 202–235.

Loehle, C., 1983. The fractal dimension and ecology. Speculations Sci. Technol. 6, 131–142.

Logerwell, E.A., Hewitt, R.P., Demer, D.A., 1998. Scale-dependent spatial variance patterns and correlations of seabirds and prey in the southeastern Bering Sea as revealed by spectral analysis. Ecography 21, 212–223.

Lovejoy, S., 1982. Area-perimeter relation for rain and cloud areas. Science 216, 185–187.

Lovejoy, S., Schertzer, D., 1986. Scale invariance, symmetries, fractals, and stochastic simulations of atmospheric phenomena. Bull. Am. Meteorol. Soc. 67, 21–32.

Lovejoy, S., Schertzer, D. (Eds.), 1991. Scaling, Fractals and Non-linear Variability in Geophysics. Kluwer Academic Publishers, Norwell, Massachusetts.

Luce, R.D., Narens, L., 1987. Measurement scales on the continuum. Science 236, 1527–1532.

Lugo, A.E., 1996. Tropical forests: Their future and our future. In: Lugo, A.E., Lowe, C. (Eds.), Tropical Forests: Management and Ecology. Springer Verlag, New York, pp. 3–17.

MacArthur, R.H., 1969. Patterns of communities in the tropics. Biol. J. Linn. Soc. Lond. 1, 19–30.

MacArthur, R.H., Wilson, E.O., 1967. Theory of Island Biogeography. Princeton University Press, Princeton.

Mackas, D.L., Denman, K.L., Abbott, M.R., 1985. Plankton patchiness: biology in the physical vernacular. Bull. Mar. Sci. 37, 652–674.

McMahon, T.A., 1973. Size and shape in biology. Science 179, 1201–1204.

McMahon, T.A., 1975. The mechanical design of trees. Sci. Am. 233, 93–102.

Magnuson, J.J., 1990. Long-term ecological research and the invisible present. BioScience 40, 495–501.

Magurran, A.E., 1988. Ecological Diversity and Its Measurement. Princeton University Press, Princeton.

Mandelbrot, B.B., 1977/1983. Fractals: Form, Chance, and Dimension. Freeman, San Francisco.

Manly, B.F.J., 1991. Randomization and Monte Carlo Methods in Biology. Chapman and Hall, London.

Mann, K.H., Lazier, J.R.N., 1991. Dynamics of Marine Ecosystems. Blackwell Scientific Publications, Boston.

Marquet, P.A., Fortin, M.-J., Pineda, J., Wallin, D.O., Clark, J., Wu, Y., Bollens, S., Jacobi, C.M., Holt, R.D., 1993. Ecological and evolutionary consequences of patchiness: a marine-terrestrial perspective. In: Levin, S.A., Powell, T.A., Steele, J.H. (Eds.), Patch Dynamics. Springer-Verlag, New York, pp. 277–304.

Marten, G.G., 1972. Censusing mouse populations by means of tracking. Ecology 53, 859–867.

Massey, B.S., 1986. Measures in Science and Engineering. Halstead (John Wiley), New York.

May, R.M., 1975. Patterns of species abundance and diversity. In: Cody, M.L., Diamond, J.M. (Eds.), Ecology and Evolution of Communities. Belknap Press, Cambridge, USA, pp. 81–120.

May, R.M., 1991. The role of ecological theory in planning reintroduction of endangered species. Proc. Zool. Soc. London 62, 145–163.

Mayor, S.J., Schaefer, J.A., Schneider, D.C., Mahoney, S.P., 2007. Spectrum of selection: new approaches to detecting the scale-dependent response to habitat. Ecology 88, 1634–1640.

McClaran, M.P., 1995. Interpreting explanatory processes for time series patterns: Lessons from three time series. In: Powell, T.M., Steele, J.H. (Eds.), Ecological Time Series. Chapman and Hall, New York, pp. 465–482.

McCullagh, P., Nelder, J.A., 1989. Generalised Linear Models. Chapman and Hall, London.

McDowell, P.F., Webb, T., Bartlein, P.J., 1995. Long-term ecological change. In: Powell, T.M., Steele, J.H. (Eds.), Ecological Time Series. Chapman and Hall, New York, pp. 327–370.

McGarigal, K., Cushman, S., Stafford, S.G., 2000. Multivariate Statistics for Wildlife and Ecology Research. Springer-Verlag, New York.

Mead, R., 1974. A test for spatial pattern at several scales using data from a grid of contiguous quadrats. Biometrics 30, 295–307.

Meentemeyer, V., Box, E.O., 1987. Scale effects in landscape studies. In: Turner, M.G. (Ed.), Landscape Heterogeneity and Disturbance. Springer-Verlag, Berlin, pp. 15–34.

Meltzer, M.I., Hastings, H.M., 1992. The use of fractals to assess the ecological impact of increased cattle population: case study from Runde communal land, Zimbabwe. J. Appl. Ecol. 29, 635–646.

Menge, B.A., 1976. Organization of the New England rocky intertidal community: role of predation, competition, and environmental heterogeneity. Ecol. Monog. 46, 355–393.

Menge, B., Olson, A.M., 1990. Role of scale and environmental factors in regulation of community structure. Trends Res. Ecol. Evol. 5, 52–57.

Mercer, W.B., Hall, A.D., 1911. The experimental error of field trials. J. Agr. Sci. 4, 107–132.

Mesterton-Gibbons, M., Dugatkin, L.A., 1995. Toward a theory of dominance hierarchies: effects of assessment, group size, and variation in fighting ability. Behav. Ecol. 6, 416–423.

Meyer, J.S., Irwin, L.L., Boyce, M.S., 1998. Influence of habitat abundance and fragmentation on northern spotted owls in western Oregon. Wildl. Monog. 139, 1–51.

Miller, J.N., Brooks, R.P., Croonquist, M.J., 1997. Effects of landscape patterns on biotic communities. Landscape Ecol. 12, 137–153.

Miller, D.C., Jumars, P.A., Nowell, A.R.M., 1984. Effects of sediment transport on deposit feeding: Scaling arguments. Limnol. Oceanogr. 29, 1202–1217.

Milne, B.T., 1992. Spatial aggregation and neutral models in fractal landscapes. Am. Nat. 139, 32–57.

Milne, B.T., 1997. Applications of fractal geometry in wildlife biology. In: Bissonette, J.A. (Ed.), Wildlife and Landscape Ecology: Effects of Pattern and Scale. Springer-Verlag, New York, pp. 32–69.

Milne, B.T., 1998. Motivation and benefits of complex systems approaches in ecology. Ecosystems 1, 449–456.

Morisita, M., 1954. Estimation of population density by spacing method. Mem. Fac. Sci. Kyushu Univ. Ser. E Biol. 1, 187–197.

Morisita, M., 1959. Measuring of the dispersion of individuals and analysis of the distributional patterns. Mem. Fac. Sci. Kyushu Univ. Ser. E Biol. 2, 215–235.

Morse, D.R., Lawton, J.H., Dodson, M.M., Williamson, M.H., 1985. Fractal dimension of vegetation and the distribution of arthropod body lengths. Nature 314, 731–732.

Mueller-Dombois, D.R., Ellenberg, H., 1974. Aims and Methods of Vegetation Ecology. Wiley, New York.

Murphy, R.C., 1914. Observations on birds of the south Atlantic. Auk 31, 439–457.

NASA, 1988. Earth System Science: A Closer View. NASA, Washington, D.C..

Neis, B., Schneider, D.C., Felt, L., Haedrich, R.L., Fischer, J., Hutchings, J.A., 1999. Stock assessment: what can be learned from interviewing resource users? Can. J. Fish. Aquat. Sci. 56, 1949–1963.

Nelder, J.A., Wedderburn, R.W.M., 1972. Generalised linear models. J. R. Stat. Soc. A 135, 370–384.

Nevitt, G.A., 2000. Olfactory foraging by Antarctic *procellariiform* seabirds: Life at high Reynolds numbers. Biol. Bull. 198, 245–253.

Newton, I. (1686). Philosophiae Naturalis Principia Mathematica, English translation 1729, Issued 1962 by University of California Press. Dynamical similarity in Book II, Sec. 7, Prop. 32.

Nikora, V.I., Pearson, C.P., Shankar, U., 1999. Scaling properties in landscape patterns: New Zealand experience. Landscape Ecol. 14, 17–33.

Nowell, A.R.M., Jumars, P.A., 1984. Flow environments of aquatic benthos. Ann. Rev. Ecol. Syst. 15, 303–328.

O'Brien, J.J., Wroblewski, J.S., 1973. On advection in phytoplankton models. J. Theor. Biol. 38, 197–202.

O'Driscoll, R.L., Schneider, D.C., Rose, G.A., Lilly, G.R., 2000. Potential contact: an ecologically interpretable method for measuring scale-dependent spatial pattern and association. Can. J. Fish. Aquat. Sci. 57, 1355–1368.

O'Neill, R.V., DeAngelis, D.L., Waide, J.B., Allen, T.F.H., 1986. A Hierarchical Concept of Ecosystems. Princeton University Press, Princeton.

O'Neill, R.V., King, A.W., 1998. Homage to St. Michael; or, Why are there so many books on scale?. In: Peterson, D.L., Parker, V.T. (Eds.), Ecological Scale. Theory and Applications. Columbia University Press, New York, pp. 3–15.

O'Neill, R.V., Krummel, J.R., Gardner, R.H., Sugihara, G., Jackson, B., DeAngelis, D.L., Milne, B.T., Turner, M.G., Zygmont, B., Christensen, S.W., Dale, V.H., Graham, R.L., 1983. Indices of landscape pattern. Landscape Ecol. 1, 153–162.

Odum, H.T., 1957. Trophic structure and productivity of Silver Springs, Florida. Ecol. Monog. 27, 55–112.

Odum, H.T., 1959. Fundamentals of Ecology, second ed. W. B. Saunders and Company, New York.

Ogilvie, A.E. (2000). Susceptibility of Marked Age 0 Atlantic cod (*Gadus morhua*) to Predation during Mark-recapture Studies, B.Sc. Honours Thesis, St. John's Canada: Memorial University.

Olea, R.A., 1999. Geostatistics for Engineers and Earth Scientists. Kluwer Academic Publishers, Norwell, MA.

Otto, R.D., 1996. An evaluation of forest landscape spatial pattern and wildlife community structure. For. Ecol. Manage. 89, 139–147.

Pascual, M., Ascioti, F.A., Caswell, H., 1995. Intermittency in the plankton: a multifractal analysis of zooplankton biomass variability. J. Plankton Res. 17, 1209–1232.

Patil, G.P., Stiteler, W.M., 1974. Concepts of aggregation and their quantification: a critical review with some new results and applications. Res. Pop. Ecol. 15, 238–254.

Patrick, R., 1968. The structure of diatom communities in similar ecological conditions. Am. Nat. 102, 173–183.

Pearson, K., Lee, A., 1903. On the laws of inheritance in man, I. Inheritance of physical characters. Biometrika 2, 357–462.

Pedley, T. (Ed.),, 1977. Scale Effects in Animal Locomotion. Academic Press, New York.

Pedlosky, J., 1979. Geophysical Fluid Dynamics. Springer-Verlag, Berlin.

Pennycuick, C.J., 1992. Newton Rules Biology. Cambridge University Press, Cambridge.

Pennycuick, C.J., Kline, K.C., 1986. Units of measurement of fractal extent, applied to the coastal distribution of bald eagle nests in the Aleutian Islands. Oecologia 68, 254–258.

Perry, J.N., Woiwood, I.P., 1992. Fitting Taylor's power law. Oikos 65, 538–542.

Peters, R.H., 1983. The Ecological Implications of Body Size. Cambridge University Press, Cambridge.

Petersen, J.E., Englund, G., 2005. Dimensional approaches to designing better experimental ecosystems: a practitioners guide with examples. Oecologia 145, 216–224.

Petersen, J.E., Hastings, A., 2001. Dimensional approaches to scaling experimental ecosystems: designing mousetraps to catch elephants. Am. Nat. 157, 324–333.

Petersen, J.E., Kennedy, V.C., Dennison, W.C., Kemp, W.M., 2009. Enclosed Experimental Ecosystems and Scale. Springer+Business, New York.

Peterson, D.L., Parker, V.T., 1998. Dimensions of scale in ecology, resource management, and society. In: Peterson, D.L., Parker, V.T. (Eds.), Ecological Scale: Theory and Applications. Columbia University Press, New York, pp. 499–522.

Piatt, J.F., 1990. The aggregative responses of Common Murres and Atlantic Puffins to schools of capelin. Stud. Avian Biol. 14, 36–51.

Pielou, E.C., 1969. An Introduction to Mathematical Ecology. Wiley-Interscience, New York.

Platt, J.R., 1964. Strong inference. Science 164, 347–353.

Platt, T.R., 1981. Thinking in terms of scale: Introduction to dimensional analysis. In: Mathematical Models in Biological Oceanography. UNESCO Press, Paris, pp. 112–121.

Platt, T.R., Denman, K.L., 1975. Spectral analysis in ecology. Ann. Rev. Ecol. Syst. 6, 189–210.

Platt, T.R., Denman, K.L., 1978. The structure of pelagic marine ecosystems. J. du Cons. Int. pour l'Explor. de la Mer 173, 60–65.

Platt, T.R., Silvert, W., 1981. Ecology, physiology, allometry, and dimensionality. J. Theor. Biol. 93, 855–860.

Poole, R.W., 1974. An Introduction to Quantitative Ecology. McGraw-Hill, New York.

Porter, E.T., Sanford, L.P., Suttles, S.E., 2000. Gypsum dissolution is not a universal integrator of "water motion". Limnol. Oceanogr. 45, 145–158.

Powell, T.M., 1989. Physical and biological scales of variability in lakes, estuaries, and the coastal ocean. In: Roughgarden, J., May, R.M., Levin, S.A. (Eds.), Perspectives in Ecological Theory. Princeton University Press, Princeton, New Jersey, pp. 157–176.

Powell, T.M., Steele, J.H., 1995. Ecological Time Series. Chapman and Hall, New York.

Preston, F.W., 1962. The canonical distribution of commonness and rarity. Parts I, II. Ecology 43 (185–215), 410–432.

Pridmore, R.D., Thrush, S.F., Wilcock, R.J., Smith, T.J., Hewitt, J.E., Cummings, V.J., 1991. Effect of the organochlorine pesticide technical chlordane on the population structure of suspension and deposit feeding bivalves. Mar. Ecol.-Prog. Ser. 76, 261–271.

Rahel, F.J., 1990. The hierarchical nature of community persistence: A problem of scale. Am. Nat. 136, 328–344.

Rainey, R.C., 1989. Migration and Meteorology. Oxford University Press, New York.

Rastetter, E.B., King, A.W., Cosby, B.J., Hornberger, G.M., O'Neill, R.V., Hobbie, J.E., 1992. Aggregating fine-scale ecological knowledge to model coarser-scale attributes of ecosystems. Ecol. Appl. 2, 55–70.

Ray, G.C., Hayden, B.P., 1992. Coastal zone ecotones. In: Hansen, A.J., di Castri, F. (Eds.), Landscape Boundaries: Consequences for Biotic Diversity and Ecological Flows. Springer-Verlag, New York, pp. 403–420.

Reckhow, K.H., 1990. Bayesian inference in non-replicated ecological studies. Ecology 71, 2053–2059.

Resetarits, W.J., Bernardo, J. (Eds.), 1998. Experimental ecology: Issues and perspectives. Oxford University Press, Oxford.

Richardson, W.J., 1978. Timing and amount of bird migration in relation to the weather: a review. Oikos 30, 224–272.

Ricker, W.E., 1958. Handbook of computation for biological statistics of fish populations. Fish. Res. Board Can. Bull. 119, 1–300.

Ricklefs, R.E., 1990. Scaling pattern and process in marine ecosystems. In: Sherman, K., Alexander, L.M., Gold, B.D. (Eds.), Large Marine Ecosystems. American Association for the Advancement of Science, Washington, D.C., pp. 169–178.

Ricklefs, R.E., Miller, G.L., 2000. Ecology, fourth ed. Freeman, New York.

Ricklefs, R.E., Schluter, D. 1993a. Species diversity: Regional and historical influences, pp. 350–363 in Ricklefs and Schluter, 1993b.

Ricklefs, R.E., Schluter, D. (Eds.), 1993b. Species Diversity in Ecological Communities. University of Chicago Press, Chicago.

Riggs, D.S., 1963. The Mathematical Approach to Physiological Problems. Williams and Wilkins, Baltimore 1977 reprint by M.I.T. Press, Cambridge, Massachusetts.

Ripley, B.D., 1981. Spatial Statistics. John Wiley & Sons, New York.

Ritchie, M.E., Olff, H., 1999. Spatial scaling laws yield a synthetic theory of biodiversity. Nature 400, 557–560.

Robbins, C.S., Bruun, B., Zim, H.S., 1983. Birds of North America: A Guide to Field Identification. Golden Press, New York.

Rodriguez-Iturbe, I., Rinaldo, A., 1997. Fractal River Basins: Chance and Self-Organization. Cambridge University Press, Cambridge.

Rogers, A., 1974. Statistical Analysis of Spatial Dispersion: The Quadrat Method. Pion, London.

Roper, D.S., Pridmore, R.D., Thrush, S.F., 1992. Population biology of *Tellina liliana* in Manukau harbour, New Zealand. N. Z. J. Mar. Freshw. Res. 26, 385–392.

Rose, G.A., Leggett, W.C., 1990. The importance of scale to predator-prey spatial correlations: an example of Atlantic fishes. Ecology 71, 33–43.

Rosen, R., 1989. Similitude, similarity, and scaling. Landscape Ecol. 3, 207–216.

Rosenzweig, M.L., 1995. Species Diversity in Space and Time. Cambridge University Press, Cambridge.

Roughgarden, J., 1979. Theory of Population Genetics and Evolutionary Ecology: An Introduction. Macmillan, New York.

Rubner, M., 1883. Über die Einflus der Körpergrösse auf Stoff und Kraftwechsel. Z. fur Biol. 19, 535–562.

Russell, B., 1937. Science and Philosophy. Macmillan, London.

Ryder, R.A., 1965. A method for estimating the potential fish production of north-temperate lakes. Trans. Am. Fish. Soc. 94, 214–218.

Safina, C., Burger, J., 1985. Common Tern foraging: seasonal trends in prey fish densities and competition with bluefish. Ecology 66, 1457–1463.

Sanford, L.P., 1997. Turbulent mixing in experimental ecosystem studies. Mar. Ecol.-Prog. Ser. 161, 265–293.

Sarrus, R., Rameaux, N., 1839. Mathématique appliquée à la physiologie. Bull. de l'Académie R. de Médecine 3, 1094–1100.

Satoh, K., 1990. Single and multiarmed spiral patterns in a cellular automaton model for an ecosystem. J. Phys. Soc. Jpn. 59, 4202–4207.

Scheffe, H., 1959. The Analysis of Variance. John Wiley & Sons, New York.

Schindler, D.W., 1987. Detecting ecosystem response to anthropogenic stress. Can. J. Fish. Aquat. Sci. 44, 6–25.

Schindler, J.E., 1988. Freshwater ecosystems: A perspective. In: Pomeroy, L.R., Alberts, J.J. (Eds.), Concepts of Ecosystem Ecology: A Comparative View. Springer-Verlag, New York, pp. 57–74.

Schmidt-Nielsen, K., 1984. Scaling. Why Is Animal Size So Important? Cambridge University Press, Cambridge.

Schneider, D., 1978. Equalisation of prey numbers by migratory shorebirds. Nature 271, 353–354.

Schneider, D.C., 1985. Migratory shorebirds: resource depletion in the tropics? Am. Ornithological Union Monog. 36, 546–558.

Schneider, D.C., 1989. Identifying the spatial scale of density-dependent interaction of predators with schooling fish in the southern Labrador Current. J. Fish Biol. 35, 109–115.

Schneider, D.C., 1990. Spatial autocorrelation in marine birds. Polar Res. 8, 89–97.

Schneider, D.C., 1991. The role of fluid dynamics in the ecology of marine birds. Oceanogr. Mar. Biol. Ann. Rev. 29, 487–521.

Schneider, D.C., 1992. The thinning and clearing of prey by predators. Am. Nat. 139, 148–160.

Schneider, D.C., 1994a. Scale-dependent patterns and species interactions in marine nekton. In: Giller, P., Rafaelli, D., Hildrew, A. (Eds.), Aquatic Ecology: Scale, Pattern, and Process. Blackwell, London, pp. 441–467.

Schneider, D.C., 1994b. Quantitative Ecology: Spatial and Temporal Scaling. Academic Press, San Diego.

Schneider, D.C., 2001a. Spatial allometry: Theory and application to experimental and natural aquatic ecosystems. In: Gardner, R.H., Kemp, W.M., Peterson, J., Kennedy, V. (Eds.), Scaling Relations in Experimental Ecology. Columbia University Press, New York, pp. 113–148.

Schneider, D.C., 2001b. The rise of the concept of scale in ecology. BioScience 51, 545–553.

Schneider, D.C., 2001c. Scale: Concept and effects. In: Levin, S.A. (Ed.), Encyclopedia of Biodiversity, vol. 5. Academic Press, San Diego, pp. 245–254.

Schneider, D.C., 2002. Scaling theory: application to marine ornithology. Ecosystems 5, 736–748.

Schneider, D.C., Bajdik, C.D., 1992. Decay of zooplankton patchiness generated at the sea surface. J. Plankton Res. 14, 531–543.

Schneider, D.C., Bult, T., Gregory, R.S., Methven, D.A., Ings, D.W., Gotceitas, V., 1999. Mortality, movement, and body size: critical scales for Atlantic cod *Gadus morhua* in the northwest Atlantic. Can. J. Fish. Aquat. Sci. 46 (supplement), 180–187.

Schneider, D.C., Duffy, D.C., 1988. Historical variation in guano production from the Peruvian and Benguela upwelling ecosystems. Clim. Change 13, 309–316.

Schneider, D.C., Gagnon, J.-M., Gilkinson, K.D., 1987. Patchiness of epibenthic megafauna on the outer Grand Banks of Newfoundland. Mar. Ecol.-Prog. Ser. 39, 1–13.

Schneider, D.C., Haedrich, R.L., 1989. Prediction limits of allometric equations: a reanalysis of Ryder's morphoedaphic index. Can. J. Fish. Aquat. Sci. 46, 503–508.

Schneider, D.C., Piatt, J.F., 1986. Scale-dependent correlation of seabirds with schooling fish in a coastal ecosystem. Mar. Ecol.-Prog. Ser. 32, 237–246.

Schneider, D.C., Walters, R., Thrush, S., Dayton, P.K., 1997. Scale-up of ecological experiments: Density variation in the mobile bivalve *Macomona liliana*. J. Exp. Mar. Biol. Ecol. 216, 129–152.

Schneider, W.J., Ayer, G.R., 1961. Effect of reforestation on streamflow in central New York. U.S. Geol. Surv. Water-Supply Pap. 1602, 1–61.

Schoener, T.W., 1968. Sizes of feeding territories among birds. Ecology 49, 123–131.

Schoonmaker, P.K., 1998. Paleoecological perspectives on ecological scale. In: Peterson, D.L., Parker, V.T. (Eds.), Ecological Scale. Columbia University Press, New York, pp. 79–103.

Seal, H.L., 1964. Multivariate Statistical Analysis for Biologists. John Wiley & Sons, New York.

Seber, G.A.F., 1980. The Linear Hypothesis: A General Theory. Charles Griffin & Company Limited, London.

Shugart, H.H. (Ed.), 1978. Time Series and Ecological Processes. Society for Industrial and Applied Mathematics, Philadelphia.

Shugart, H.H., 1984. A Theory of Forest Dynamics. Springer-Verlag, New York.

Shugart, H.H., 1998. Terrestrial Ecosystems in Changing Environments. Cambridge University Press, New York.

Shugart, H.H., Michaels, P.J., Smith, T.M., Weinstein, D.A., Rastetter, E.A., 1988. Simulation models of forest succession. In: Rosswall, T., Woodmansee, R.G., Risser, P.G. (Eds.), Scales and Global Change. John Wiley & Sons, London, pp. 125–151.

Shugart, H.H., West, D.C., 1981. Long-term dynamics of forest ecosystems. Am. Sci. 69, 647–652.

Skellam, J.G., 1951. Random dispersal in theoretical populations. Biometrika 38, 196–218.

Skellam, J.G., 1952. Studies in statistical ecology, I. Spatial pattern. Biometrika 39, 346–362.

Smith, H.F., 1938. An empirical law describing heterogeneity in the yields of agricultural crops. J. Agricul. Sci. 28, 1–23.

Smith, P.E., 1973. The mortality and dispersal of sardine eggs and larvae. J. du Cons. Int. pour l'Expl. de la Mer 164, 282–292.

Smith, P.E., 1978. Biological effects of ocean variability: time and space scales of biological response. J. du Cons. Int. pour l'Expl. de la Mer 173, 117–127.

Sneath, P.H.A., Sokal, R.R., 1973. Numerical Taxonomy. Freeman, San Francisco.

Snedecor, G.W., Cochran, W.G., 1980. Statistical Methods, seventh ed.. Iowa State University Press, Ames, Iowa.

Sokal, R.R., 1979. Ecological parameters inferred from spatial correlograms. In: Patil, G.P., Rosenzwig, M.L. (Eds.), Contemporary Quantitative Ecology and Related Ecometrics. International Co-operative Publishing House, Fairland, MD, pp. 167–196.

Sokal, R.R., Rohlf, F.J., 1995. Biometry. Freeman, San Francisco.

Solé, R.V., Manrubia, S.C., 1995. Are rainforests self-organized in a critical state? J. Theor. Biol. 173, 31–40.

Stahl, W.R, 1961. Dimensional analysis in mathematical biology, I. General discussion. Bull. Math. Biophys. 23, 355–376.

Stahl, W.R., 1962. Dimensional analysis in mathematical biology, II. Bull. Math. Biophys. 24, 81–108.

Stanley, H.E., Amaral, L.A.N., Buldyrev, S.V., Goldberger, A.L., Havlin, S., Leschhorn, H., Maass, P., Makse, H.A., Peng, C-K., Salinger, M.A., Stanley, M.H.R., Viswanathan, G.M., 1996. Scaling and universality in animate and inanimate systems. Phys. A 231, 20–48.

Steele, J.H., 1978. Some comments on plankton patches. In: Steele, J.H. (Ed.), Spatial Pattern in Plankton Communities. Plenum Press, New York, pp. 11–20.

Steele, J.H., 1991a. Marine ecosystem dynamics: comparison of scales. Ecol. Res. 6, 175–183.

Steele, J.H., 1991b. Can ecological theory cross the land-sea boundary? J. Theor. Biol. 153, 425–436.

Stephens, D.W., Dunbar, S.R., 1993. Dimensional analysis in behavioral ecology. Behav. Ecol. 4, 172–183.

Stevens, S.S., 1946. On the theory of scales of measurement. Science 103, 677–680.

Stevens, S.S., 1975. Psychophysics. John Wiley & Sons, New York.

Stewart-Oaten, A., Murdoch, W.W., 1986. Environmental impact assessment: "Pseudoreplication" in time? Ecology 67, 929–940.

Stohlgren, T.J., 1995. Planning long-term vegetation studies at landscape scales. In: Powell, T.M., Steele, J.H. (Eds.), Ecological Time Series. Chapman and Hall, New York, pp. 209–241.

Stommel, H., 1963. The varieties of oceanographic experience. Science 139, 572–576.

Storch, D., Evan, K.L., Gaston, K.J., 2005. The species-area-energy relationship. Ecol. Letters 8, 487–492.

Taylor, E.S., 1974. Dimensional Analysis for Engineers. Clarendon Press, Oxford.

Story, G.M., Kirwin, W.J., Widdowson, J.D.A. (Eds.), 1982. Dictionary of Newfoundland English. Breakwater Books, St. John's.

Strahler, A.N., 1952. Hypsometric (area-altitude) analysis of erosional topography. Bull. Geol. Soc. Am. 63, 1117–1142.

Student, S., 1907. On the error of counting with a haemacytometer. Biometrika 5, 351–364.

Sugihara, G., May, R.M., 1990. Applications of fractals in ecology. Trends Res. Ecol. Evol. 5, 79–87.

Sugihara, G., Grenfell, B., May, R.M., 1990. Distinguishing error from chaos in ecological time series. Philos. Trans. R. Soc. London B 330, 235–251.

Sutor, G.W., 1996. Abuse of hypothesis testing statistics in ecological risk assessment. Hum. Ecol. Risk Assess. 2, 331–347.

Swanson, F.J., Johnson, S.L., Gregory, S.V., Acker, S.A., 1992. Flood disturbance in a forested mountain landscape. BioScience 48, 681–689.

Taylor, C.C., 1953. Nature of variability in trawl catches. Fish. Bull. 54, 143–166.

Taylor, E.S., 1974. Dimensional Analysis for Engineers. Clarendon Press, Oxford.

Taylor, L.R., 1961. Aggregation, variance, and the mean. Nature 189, 732–735.

Taylor, L.R., Taylor, R.A.J., 1977. Aggregation, migration and population mechanics. Nature 265, 415–421.

Taylor, L.R., Woiwod, I.P., Perry, J.N., 1978. The density-dependence of spatial behaviour and the rarity of randomness. J. Anim. Ecol. 47, 383–406.

Teal, J.M., 1962. Energy flow in the salt marsh ecosystem of Georgia. Ecology 43, 614–624.

Thompson, D.W., 1917. On Growth and Form, 1961 abridged edition edited by J. T. Bonner. Cambridge University Press, Cambridge.

Thrush, S., 1991. Spatial patterns in soft-bottom communities. Trends Res. Ecol. Evol. 6, 75–79.

Thrush, S.F., Hewitt, J.E., Pridmore, R.D., 1989. Patterns in the spatial arrangements of polychaetes and bivalves in intertidal sandflats. Mar. Biol. 102, 529–535.

Thrush, S.F., Hewitt, J.E., Pridmore, R.D., Cummings, V.J., 1996. Adult/juvenile interactions of infaunal bivalves: contrasting outcomes in different habitats. Mar. Ecol.-Prog. Ser. 132, 83–92.

Thrush, S., Pridmore, R.D., Cummings, V.J., Dayton, P.K., Ford, R., Grant, J., Hewitt, J.E., Hines, A.H., Lawrie, S.M., McArdle, B.H., Schneider, D.C., Turner, S.J., Whitlatch, R.B., Wilkinson, M.R., 1997. Matching the outcome of small-scale density manipulation experiments with larger scale patterns: An example of bivalve adult/juvenile interactions. J. Exp. Mar. Biol. Ecol. 216, 153–169.

Thrush, S.F., Pridmore, R.D., Hewitt, J.E., Cummings, V.J., 1991. Impact of ray feeding disturbances on sandflat macrobenthos: do communities dominated by polychaetes or shellfish respond differently?. Mar. Ecol.-Prog. Ser. 69, 245–252.

Thrush, S.F., Pridmore, R.D., Hewitt, J.E., Cummings, V.J., 1992. Adult infauna as facilitators of colonization of intertidal sandflats. J. Exp. Mar. Biol. Ecol. 159, 253–265.

Thrush, S.F., Pridmore, R.D., Hewitt, J.E., Cummings, V.J., 1994. The importance of predators on a sandflat: interplay between seasonal changes in prey densities and predator effects. Mar. Ecol.-Prog. Ser. 107, 211–222.

Tilman, D., Kareiva, P., 1997. Spatial Ecology. Princeton University Press, Princeton.

Tolman, R.C., 1914. The principle of similitude. Phys. Rev. 4, 345–376.

Tufte, E., 1990. Envisioning Information. Graphics Press, Cheshire, Connecticut.

Tukey, J.W., 1977. Exploratory Data Analysis. Addison-Wesley, Reading, Massachusetts.

Turner, S.J., Hewitt, J.E., Wilkinson, M.R., Morrisey, D.J., Thrush, S.F., Cummings, V.J., Funnel, G., 1999. Seagrass patches and landscapes: the influence of wind-wave dynamics and hierarchical arrangements of spatial structure on macrofaunal seagrass communities. Estuaries 22, 1016–1032.

Turner, M.G., Gardner, R.H. (Eds.), 1991. Quantitative Methods in Landscape Ecology. Springer-Verlag, Berlin.

Uhlmann, D., 1985. Scaling of microcosms and the dimensional analysis of lakes. Int. Rev. der Gesamten Hydrobiologie 70, 47–62.

Usher, M.B., 1969. The relation between mean square and block size in the analysis of similar patterns. J. Ecol. 57, 505–514.

Usher, M.B., 1988. Biological invasions of nature reserves: a search for generalisations. Biol. Conserv. 44, 119–135.

Valentine, J.W., 1973. The Evolutionary Paleoecology of the Marine Biosphere. Prentice-Hall, New York.

van Dobben, W.H., Lowe-McConnell, R.H. (Eds.), 1975. Unifying Concepts in Ecology, Report of the plenary sessions of the first international congress of ecology. Dr. W. Junk B.V. Publishers, The Hague.

VandeCastle, J., 1998. Remote sensing applications in ecosystem analysis. In: Peterson, D.L., Parker, V.T. (Eds.), Ecological Scale. Theory and Applications. Columbia University Press, New York, pp. 271–288.

Varley, G.C., 1949. Population changes in German forest pests. J. Anim. Ecol. 18, 117–122.

Vedyushkin, M.A., 1994. Fractal properties of forest spatial structure. Vegetatio 113, 65–70.

Veitzer, S.A., Gupta, V.K., 2000. Random self-similar river networks and derivations of generalized Horton laws in terms of statistical simple scaling. Water Res. Res. 36, 1033–1048.

Venrick, E.L., 1995. Scales of variability in a stable environment: Phytoplankton in the central North Pacific. In: Powell, T.M., Steele, J.H. (Eds.), Ecological Time Series. Chapman and Hall, New York, pp. 150–180.

Vogel, S., 1981. Life in Moving Fluids: The Physical Biology of Flow. Princeton University Press, Princeton, New Jersey.

Walker, D.A., Walker, M.D., 1991. History and pattern of disturbance in Alaskan arctic terrestrial ecosystems: a hierarchical approach to analysing landscape change. J. Appl. Ecol. 28, 244–276.

Walsberg, G.L., 1983. Avian ecological energetics. In: Farner, D.S., King, J.R., Parkes, K.C. (Eds.), Avian Biol. Academic Press, New York, pp. 166–220.

Warner, S.C., Limburg, K.E., Ariso, A.H., Dodd, M., Dushoff, J., Stergiou, K.I., Potts, J., 1995. Time series compared across the land-sea gradient. In: Powell, T.M., Steele, J.H. (Eds.), Ecological Time Series. Chapman and Hall, New York, pp. 242–273.

Watt, K.E.F., 1968. Ecology and Resource Management: A Quantitative Approach. McGraw-Hill, New York.

Weber, L.H., El-Sayed, S.Z., Hampton, I., 1986. The variance spectra of phytoplankton, krill and water temperature in the Antarctic Ocean south of Africa. Deep-Sea Res. 33, 1327–1343.

Weibel, E.R., 1979. Stereological Methods. Academic Press, London.

Weissburg, M.J., 2000. The fluid dynamical context of chemosensory behavior. Biol. Bull. 198, 188–202.

Weissburg, M.J., Zimmer-Faust, R.K., 1993. Life and death in moving fluids: hydrodynamic effects on chemosensory-mediated predation. Ecology 74, 1428–1443.

Wessman, C., 1992. Spatial scales and global change: bridging the gap from plots to GCM grid cells. Ann. Rev. Ecol. Syst. 23, 175–200.

West, B.J., Schlesinger, M., 1990. The noise in natural phenomena. Am. Sci. 78, 40–45.

West, G.B., Brown, J.H., Enquist, B.J., 1997. A general model for the origin of allometric scaling laws in biology. Science 276, 122–126.

West, G.B., Brown, J.H., Enquist, B.J., 1999. The fourth dimension of life: fractal geometry and allometric scaling of organisms. Science 284, 1677–1679.

Westcott, D.A., Graham, D.L., 2000. Patterns of movement and seed dispersal of a tropical frugivore. Oecologia 122, 249–257.

Whitlatch, R.B., Wilkinson, M.R., 1997. Matching the outcome of small-scale density manipulation experiments with larger scale patterns: an example of adult-juvenile interactions. J. Exp. Mar. Biol. Ecol. 216, 153–169.

Whitney, H., 1968. The mathematics of physical quantities. Part 1. Mathematical models of measurement. Am. Math. Mon. 75, 115–138 Part II. Quantity structures and dimensional analysis, Am. Math. Mon. 75, 227–256.

Whittaker, R.H., 1960. Vegetation of the Siskiyou Mountains, Oregon and California. Ecol. Monog. 26, 1–80.

Whittaker, R.H., 1977. Evolution of species diversity in land communities. Evol. Biol. 10, 1–68.

Whittaker, R.J., Willis, K.J., Field, R., 2001. Scale and species richness: Towards a general, hierarchical theory of species diversity. J. Biogeography 28, 453–470.

Widom, B., 1965. Surface tension and molecular correlations near the critical point. J. Chem. Phys. 43, 3892–3898.

Wiens, J.A., 1973. Pattern and process in grassland bird communities. Ecol. Monog. 43, 237–270.

Wiens, J.A., 1976. Population responses to patchy environments. Ann. Rev. Ecol. Syst. 7, 81–120.

Wiens, J.A., 1989. Spatial scaling in ecology. Funct. Ecol. 3, 385–397.

Wiens, J.A., 1999. The science and practice of landscape ecology. In: Klopatch, J.M., Gardner, R.H. (Eds.), Landscape Ecological Analysis. Springer-Verlag, Berlin, pp. 371–383.

Wiens, J.A., Crist, T.O., With, K.A., Milne, B.T., 1995. Fractal patterns of insect movement in microlandscape mosaics. Ecology 76, 663–666.

Williams, C.B., 1964. Patterns in the Balance of Nature and Related Problems in Quantitative Ecology. Academic Press, London.

Williamson, M.H., 1988. Relationship of species number to area, distance and other variables. In: Myers, A.A., Giller, P.S. (Eds.), Analytical Biogeography. Chapman and Hall, Chicago, pp. 91–115.

Williamson, M.H., Lawton, J.H., 1991. Fractal geometry of ecological habitats. In: Bell, S.S., McCoy, E.D., Mushinsky, H.R. (Eds.), Habitat Structure. Chapman and Hall, London, pp. 69–86.

Wilson, K.G., 1971. Renormalization group and critical phenomena. Phys. Rev. B 4, 3174–3205.

Wilson, M.V., Shmida, A., 1984. Measuring beta diversity with presence-absence data. J. Ecol. 72, 1055–1064.

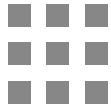
Winberg, G.G., 1971. Methods for the Estimation of Production of Aquatic Animals. Academic Press, New York.

With, K.A., 1994. Using fractal analysis to assess how species perceive landscape structure. Landscape Ecol. 9, 25–36.

With, K.A., Cadaret, S.J., Davis, C., 1999. Movement responses to patch structure in experimental fractal landscapes. Ecology 80, 1340–1353.

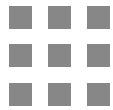
Withers, M.A., Meentemeyer, V., 1999. Concepts of scale in landscape ecology. In: Klopatch, J.M., Gardner, R.H. (Eds.), Landscape Ecological Analysis. Springer-Verlag, Berlin, pp. 205–252.

Yager, P.L., Nowell, A.R.M., Jumars, P.A., 1993. Enhanced deposition to pits: a local food source for benthos. J. Mar. Res. 51, 209–236.



List of Tables

List of Boxes



Index

Page numbers with a "b" indicate a box
Page numbers with an "f" indicate a figure
Page numbers with a "t" indicate a table