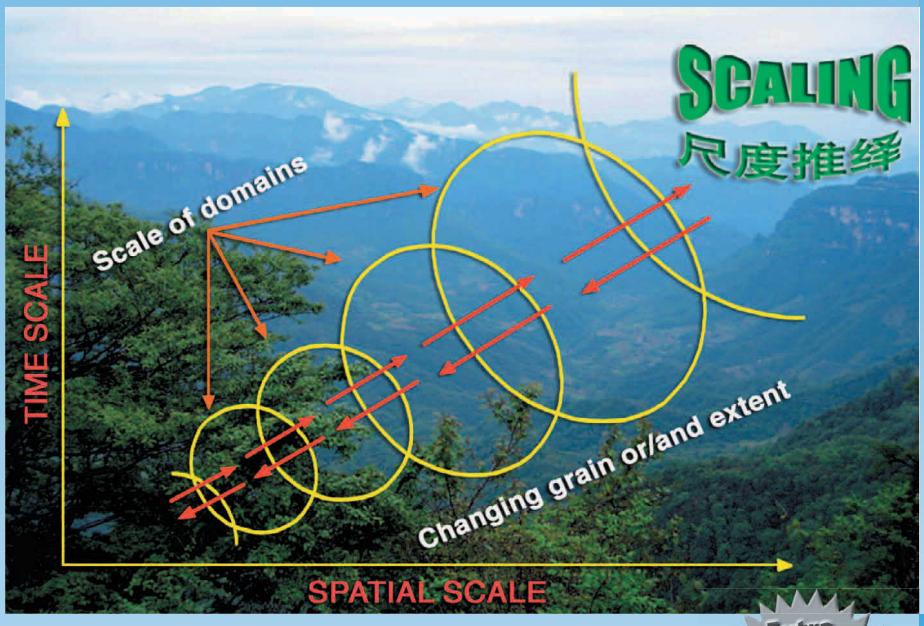
Scaling and Uncertainty Analysis in Ecology

Methods and Applications

Edited by
Jianguo Wu, K. Bruce Jones, Harbin Li
and Orie L. Loucks



Extra Materials
extras.springer.com

Springer

SCALING AND UNCERTAINTY ANALYSIS IN ECOLOGY

Scaling and Uncertainty Analysis in Ecology

Methods and Applications

Edited by

JIANGUO WU
Arizona State University, Tempe, AZ, U.S.A.

K. BRUCE JONES
US Environmental Protection Agency, Las Vegas, U.S.A.

HARBIN LI
USDA Forest Service Southern Research Station, Charleston, U.S.A.

and

ORIE L. LOUCKS
Miami University, Oxford, OH, U.S.A.

Springer

A C.I.P. Catalogue record for this book is available from the Library of Congress.

ISBN-10 1-4020-4664-2 (PB)
ISBN-13 978-1-4020-4664-3 (PB)
ISBN-10 1-4020-4662-6 (HB)
ISBN-13 978-1-4020-4662-9 (HB)
ISBN-10 1-4020-4663-4 (e-books)
ISBN-13 978-1-4020-4663-6 (e-books)

Published by Springer,
P.O. Box 17, 3300 AA Dordrecht, The Netherlands.

www.springer.com

Printed on acid-free paper

Printed in the Netherlands.

Dedication

This book is dedicated to landscape ecologists who have made outstanding contributions to our understanding of scaling issues and hierarchy theory in ecology and environmental science.

Contents

Preface

Scale is a unifying concept that cuts across all natural and social sciences. At the same time, scaling is a common challenge in both basic and applied research. Accordingly, scale and scaling have become two of the most widely used buzzwords in ecology today. Over the past two decades, more than a dozen books and many more journal papers have been published on the problems of scale and scaling in ecology and geophysical sciences. These publications, as reviewed in the chapters of this book, have contributed significantly to our current understanding of scale issues. A little more than 30 years ago, the noted geneticist and evolutionary biologist, Theodosius Dobzhansky, stated that "Nothing in biology makes sense except in the light of evolution" (The American Biology Teacher 35:125-129). Today, there seems a growing consensus in ecology that pattern and process make little sense without consideration of scale.

While scale issues are widely recognized, a comprehensive understanding of scaling theory and methods still is missing. In this book we make several observations on the status of research on scale in ecology. First, while ecologists have played an active role in the application of scale-related theories such as hierarchy, self-similarity, and self-organized criticality, a number of pragmatic scaling methods have developed in geophysical disciplines. Many of them may be quite appropriate for a range of ecological problems, but are yet to be fully explored in ecology. Second, some of the most frequently mentioned scaling theories are often seen as being at odds with each other. For example, hierarchy theory implies scale-multiplicity and thresholds, while self-similarity and self-organized criticality suggest scale invariance. A full understanding of the relationships among different scaling theories is needed, and this requires critical examination of recent theoretical and empirical studies. Third, most scaling studies in ecology have either ignored or inadequately addressed the issues of uncertainty and error propagation, which should be an integral part of scaling. We argue that scaling, without considering uncertainty, is easy but relatively trivial; scaling with known uncertainty is challenging but essential. Fourth, scaling often requires field-based data from multiple spatial and temporal scales, but these data rarely exist for many ecosystems. Such inadequacies of data further elevate the demand for effective scaling

approaches. Finally, scaling theories and methods have seldom been applied explicitly in the contexts of environmental management, planning, and decision-making processes, where the scale of social, economic, political, and ecological processes may clash with each other. A pluralistic and interdisciplinary approach is needed to resolve scaling problems in such complex situations.

To address these problems, a workshop entitled "Scaling and Uncertainty Analysis in Ecology: Methods and Application" was held during September 17-19, 2002 at Arizona State University, Tempe, U.S.A., supported through a grant from the United States Environmental Protection Agency (EPA). The major objectives of the workshop were to identify approaches and methods in scaling and uncertainty analysis, and to consider a series of case studies illustrating how scale issues are dealt with in various areas of research. More than 20 active researchers in scaling and uncertainty analysis were invited to participate in the workshop, many of whom were recipients of EPA's Science To Achieve Results (STAR) program (Regional Scale Analysis and Assessment). This book has evolved out of the scaling workshop, and is comprised primarily of the papers remaining after a critical external review process.

The book, therefore, presents a comprehensive and up-to-date review and synthesis of concepts, theories, methods and case studies in scaling and uncertainty analysis that are relevant to ecology. The series of case studies included here illustrate how scaling and uncertainty analysis are being conducted in ecology and environmental science, from population to ecosystem processes, from biodiversity to landscape patterns, and from basic research to multidisciplinary management and policy-making issues. The book explicitly considers uncertainty and error analysis as an integral part of scaling. While the theme of this book focuses primarily on spatial scaling, several chapters deal as well with aspects of temporal scaling. It is not intended to be a handbook of "scaling recipes," but we hope that it will help readers gain a fuller understanding of the state-of-the-science of scale issues. We expect that this book will be of interest to a wide range of audiences, including graduate students, academic professionals, and applied researchers and specialists in ecological, environmental, and earth sciences. It may be used as a text or reference book for graduate courses in ecology and related disciplines. This book should be particularly appealing to scientists and practitioners working on broad spatial scales. Also, the book can be useful to decision makers who are conscious about scale issues as they translate science into resource use policies.

We are most deeply indebted to the contributors of papers included in the book, whose enthusiasm and dedication have made this book a reality. Many other individuals also were instrumental to the completion of the book. We especially thank the following people for providing valuable reviews of book chapters: Dennis Baldocchi, Klaus Butterbach-Bahl, Mark Castro, Jiquan Chen, Mark R. T. Dale, Dean Gesch, Phil A. Graniero, John Harte, Geoffrey J. Hay, Louis R. Iverson, James R. Karr, Madhu Katti, Richard G. Lathrop, Helene Muller-Landau, John Ludwig, James R. Meadowcroft, Garry Peterson, Geoffrey C. Poole, Edward B. Rastetter, Helen Regan, Christine Ribic, Steven W. Running, Santiago Saura, Matthew Williams, and Xinyuan Wu. We are extremely grateful to Chuck Redman (Director), Nikol Grant, and Shirley Stapleton at the Center for Environmental Studies of Arizona State University who provided wonderful logistic support during the scaling workshop in Tempe. We also thank Barbara Levinson and Jonathan Smith at EPA

for their support for the scaling workshop in Tempe. Last, but not least, we express our sincere appreciation to Dr. Catherine Cotton (Publishing Editor) and Ms. Ria Kanters at Springer for their wonderful guidance and assistance during the production of the book.

Finally, we should note that several chapters originally had color images which later were converted to grayscale. We have made these color figures available online at a web site specifically for this book, which also contains the abstracts of all chapters and additional information on scaling and uncertainty analysis. The web address can be freely accessed at: http://LEML.asu.edu/ScalingBook/.

Editors
Jianguo (Jingle) Wu
K. Bruce Jones
Harbin Li
Orie L. Loucks

List of Contributors

Ronald D. Bassar, Montana Cooperative Wildlife Research Unit, University of Montana, Missoula, MT 59812

C. Ann Bateson, Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, CO 80309-0216

Richard A. Berk, Department of Statistics, University of California, Los Angeles, CA 90095

Mark A. Bradford, Institute of Ecology, University of Georgia, Athens, GA 30602-2602

John Campbell, College of Forestry, Oregon State University, Corvallis, OR 97331

Warren B. Cohen, College of Forestry, Oregon State University, Corvallis, OR 97331

Chad L. Cross, U.S. Environmental Protection Agency, Las Vegas, NV 89193-3478

Jan de Leeuw, Department of Statistics, University of California, Los Angeles, CA 90095-1554

Curtis M. Edmonds, U.S. Environmental Protection Agency, Las Vegas, NV 89193-3478

Robert P. Gibbens, USDA ARS, Jornada Experimental Range, Las Cruces, NM 88003-0003

Peter M. Groffman, Institute of Ecosystem Studies, Box AB, Millbrook, NY 12545-0129

Michael Guzy, College of Forestry, Oregon State University, Corvallis, OR 97331

Melissa M. Hart, Montana Cooperative Wildlife Research Unit, University of Montana, Missoula, MT 59812

Kris M. Havstad, USDA ARS, Jornada Experimental Range, Las Cruces, NM 88003-0003

Fangliang He, Department of Renewable Resources, University of Alberta, Edmonton, Alberta, Canada T6G 2H1

Jeffrey E. Herrick, USDA ARS, Jornada Experimental Range, Las Cruces, NM 88003-0003

Thomas P. Hollenhorst, Natural Resources Research Institute, University of Minnesota, Duluth, MN 55811-1442

George E. Host, Natural Resources Research Institute, University of Minnesota, Duluth, MN 55811-1442

Laura F. Huenneke, College of Engineering and Natural Sciences, Northern Arizona University, Flagstaff, AZ 86011

K. Bruce Jones, U.S. Environmental Protection Agency, Las Vegas, NV 89193-3478

Lucinda B. Johnson, Natural Resources Research Institute, University of Minnesota, Duluth, MN 55811-1442

Carol A. Johnston, Center for Biocomplexity Studies, South Dakota State University, Brookings, SD 57007-0896

Bruce M. Kahn, Department of Zoology, Miami University, Oxford, OH 45056

Ute Langner, Montana Cooperative Wildlife Research Unit, University of Montana, Missoula, MT 59812

Beverly E. Law, College of Forestry, Oregon State University, Corvallis, OR 97331-5752

Michael Lefsky, College of Forestry, Oregon State University, Corvallis, OR 97331

Harbin Li, USDA Forest Service Southern Research Station, Center for Forested Wetlands Research, Charleston, SC 29414

Penn Lloyd, DST/NRF Centre of Excellence, Percy FitzPatrick Institute, University of Cape Town, Rondebosch, 7701, South Africa

Orie L. Loucks, Department of Zoology, Miami University, Oxford, OH 45056

Thomas E. Martin, Montana Cooperative Wildlife Research Unit, University of Montana, Missoula, MT 59812

Robert I. McDonald, Nicholas School of Environment & Earth Sciences, Duke University, Durham, NC 27708

Emily S. Minor, Nicholas School of Environment & Earth Sciences, Duke University, Durham, NC 27708

Maliha S. Nash, U.S. Environmental Protection Agency, Las Vegas, NV 89193-3478

Anne C. Neale, U.S. Environmental Protection Agency, Las Vegas, NV 89193-3478

Robert V. O'Neill, Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831

Debra P.C. Peters, USDA ARS, Jornada Experimental Range, Las Cruces, NM 88003-0003

Christopher S. Potter, NASA Ames Research Center, Moffett Field, CA 94035-1000

Albert Rango, USDA ARS, Jornada Experimental Range, Las Cruces, NM 88003-0003

Roland L. Redmond, Montana Cooperative Wildlife Research Unit, University of Montana, Missoula, MT 59812

William Reed, Department of Mathematics and Statistics, University of Victoria, Victoria, BC, Canada V8W 3P4

James F. Reynolds, Department of Biology and Nicholas School of Environment & Earth Sciences, Duke University, Durham, NC 27708

Kurt H. Riitters, U.S. Forest Service Southern Research Station, Research Triangle Park, NC 27709

William H. Schlesinger, Nicholas School of Environment & Earth Sciences, Duke University, Durham, NC 27708

Boris A. Shmagin, Water Resources Institute, South Dakota State University, Brookings, South Dakota 57007

Elizabeth R. Smith, U.S. Environmental Protection Agency (E243-05), National Exposure Research Laboratory, Research Triangle Park, NC 27711

Harry J. Stone, Department of Zoology, Miami University, Oxford, OH 45056

Osbert Sun, College of Forestry, Oregon State University, Corvallis, OR 97331-5752

Eric A. Treml, Nicholas School of Environment & Earth Sciences, Duke University, Durham, NC 27708

Dave Turner, College of Forestry, Oregon State University, Corvallis, OR 97331

Dean L. Urban, Nicholas School of Environment & Earth Sciences, Duke University, Durham, NC 27708-0328

Rick D. Van Remortel, Lockheed Martin Environmental Services, Las Vegas, NV 89119

Steve Van Tuyl, College of Forestry, Oregon State University, Corvallis, OR 97331

Rodney T. Venterea, Institute of Ecosystem Studies, Box AB, Millbrook, NY 12545

Louis V. Verchot, Institute of Ecosystem Studies, Box AB, Millbrook, NY 12545

Timothy G. Wade, U.S. Environmental Protection Agency (E243-05), National Exposure Research Laboratory, Research Triangle Park, NC 27711

Carol A. Wessman, Cooperative Institute for Research in Environmental Sciences and Department of Ecology and Evolutionary Biology, University of Colorado, Boulder, CO 80309-0216

James D. Wickham, U.S. Environmental Protection Agency (E243-05), National Exposure Research Laboratory, Research Triangle Park, NC 27711

Jianguo (Jingle) Wu, School of Life Sciences and Global Institute of Sustainability, Arizona State University, Tempe, AZ 85287-4501

Jin Yao, Department of Biology and Earth Sciences, Adams State College, Alamosa, CO 81102

PART I

CONCEPTS AND METHODS

CHAPTER 1

CONCEPTS OF SCALE AND SCALING

JIANGUO WU AND HARBIN LI

1.1 INTRODUCTION

The relationship between pattern and process is of great interest in all natural and social sciences, and scale is an integral part of this relationship. It is now well documented that biophysical and socioeconomic patterns and processes operate on a wide range of spatial and temporal scales. In particular, the scale multiplicity and scale dependence of pattern, process, and their relationships have become a central topic in ecology (Levin 1992, Wu and Loucks 1995, Peterson and Parker 1998). Perspectives centering on scale and scaling began to surge in the mid-1980's and are pervasive in all areas of ecology today (Figure 1.1). A similar trend of increasing emphasis on scale and scaling is also evident in other natural and social sciences (e.g., Blöschl and Sivapalan 1995, Marceau 1999, Meadowcroft 2002).

Scale usually refers to the spatial or temporal dimension of a phenomenon, and scaling is the transfer of information between scales (more detail below). Three distinctive but interrelated issues of scale have frequently been discussed in the literature: (1) characteristic scales, (2) scale effects, and (3) scaling (and associated uncertainty analysis and accuracy assessment). The concept of *characteristic scale* implies that many, if not most, natural phenomena have their own distinctive scales (or ranges of scales) that characterize their behavior (e.g., typical spatial extent or event frequency). Characteristic scales are intrinsic to the phenomena of concern, but detected characteristic scales with the involvement of the observer may be tinted with subjectivity (Wu 1999). Conceptually, characteristic scales may be perceived as the levels in a *hierarchy*, and associated with *scale breaks* (O'Neill et al. 1991, Wu 1999). Ecological patterns and processes have been shown to have distinctive characteristic scales on which their dynamics can be most effectively studied (Clark 1985, Delcourt and Delcourt 1988, Wu 1999). Thus, identifying characteristic scales provides a key to profound understanding and enlightened scaling.

Scale effects usually refer to the changes in the result of a study due to a change in the scale at which the study is conducted. Effects of changing scale on sampling

J. Wu, K.B. Jones, H. Li, and O.L. Loucks (eds.),
Scaling and Uncertainty Analysis in Ecology: Methods and Applications, 3–15.

and experimental design, statistical analyses, and modeling have been well documented in ecology and geography (e.g., Turner et al. 1989b, White and Running 1994, Wu and Levin 1994, Pierce and Running 1995, Jelinski and Wu 1996, Dungan et al. 2002, Wu 2004). In geography, scale effects have been studied for several decades in the context of the modifiable areal unit problem or *MAUP* (Openshaw 1984, Jelinski and Wu 1996, Marceau 1999). Scale effects may be explained in terms of scale-multiplicity, characteristic scales, and hierarchy, but may also be artifacts due to errors in sampling and measurements, distortions in data resampling, and flaws in statistical analysis and modeling (Jelinski and Wu 1996, Wu 2004). Characteristic scales and scale effects are inherently related to the issue of scaling. While characteristic scales provide a conceptual basis and practical guidelines for scaling, quantitative descriptions of scale effects can directly lead to scaling relations (Wu 2004).

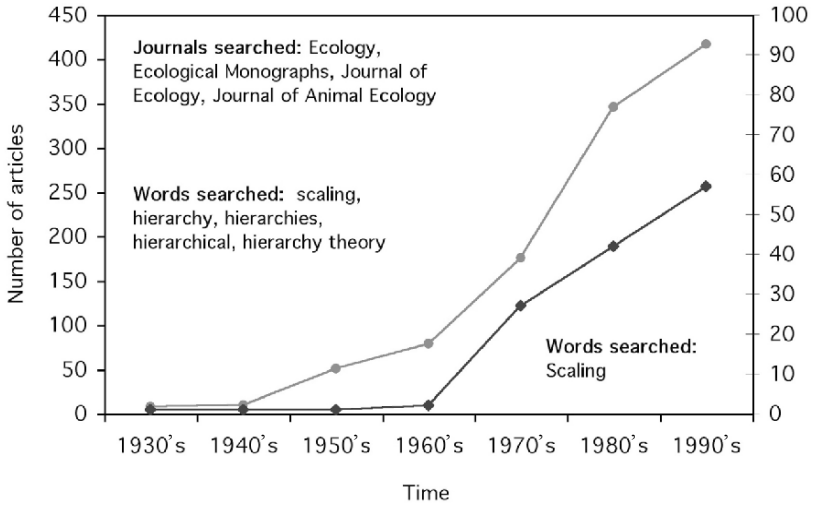


Figure 1.1. *Rapid increase in the use of terms related to scale in the ecological literature. Based on an internet search using JSTOR (http://www.jstor.org/), the number of articles containing words (scaling, hierarchy, hierarchies, hierarchical, hierarchy theory) shows a great increase in four major ecology journals in the last seven decades (gray line). The trend for the word scaling alone is similar (black line). The four journals are: Ecology and Ecological Monographs published by Ecological Society of America, and Journal of Ecology and Journal of Animal Ecology published by British Ecological Society. Note that the number of years for the 1990's was only seven (1990-1996).*

With the recent burst of interest in the issues of scale, the terms *scale* and *scaling* have become buzzwords in ecology. However, because these terms have been used in diverse disciplines, both have acquired a number of different connotations and expressions. Good science starts with clear definitions. The development of a science of scale or scaling may be hampered if the concepts of scale and scaling are

used without any consistency. In this section, we review the main usages of these terms, propose a three-tiered scale conceptualization framework, and discuss their relevance to the issue of ecological scaling.

1.2 CONCEPT OF SCALE

We propose a three-tiered conceptualization of scale, which organizes scale definitions into a conceptual hierarchy that consists of the *dimensions*, *kinds*, and *components* of scale (Figure 1.2). Dimensions of scale are most general, components of scale are most specific, and kinds of scale are in between. This three-tiered structure seems to provide a clearer picture of how various scale concepts differ from or relate to each other.

1.2.1 Dimensions of Scale

We distinguish three primary dimensions of scale: *space*, *time*, and *organizational level*. Note that Dungan et al.'s (2002) three dimensions of scale (sampling, analysis, and phenomena) are commensurable with what we here call the *kinds* of scale (see below). Space and time are the two fundamental axes of scale, whereas organizational hierarchies are usually constructed by the observer (Figure 1.2a). *Scale* has been commonly defined in terms of time or space. In recent decades, the relationship between temporal and spatial scales has received increasing attention. It is well documented that the characteristic scales of many physical and ecological phenomena are related in space versus time, such that the ratio between spatial and temporal scales tends to be relatively invariant over a range of scales. This ratio is termed the *characteristic velocity* (Blöschl and Sivapalan 1995). The idea that spatial and temporal scales are fundamentally linked so that complex systems can be decomposed in time and space simultaneously is essential to hierarchy theory (Courtois 1985, Wu 1999). This *space-time correspondence principle* has been supported by a number of empirically constructed space-time scale diagrams (or Stommel diagrams) in the past two decades (Stommel 1963, Clark 1985, Urban et al. 1987, Delcourt and Delcourt 1988, Blöschl and Sivapalan 1995, Wu 1999). These studies have shown that, for a variety of physical, ecological, and socioeconomic phenomena, large-sized events tend to have slower rates and lower frequencies, whereas small things are faster and more frequent. However, one must recognize that not all natural phenomena strictly obey the space-time correspondence principle. Many temporally cyclic events, for example, take place over a wide range of spatial scales with a relatively constant frequency. In some other cases, scale variability of different sources may overwhelm the signal of scale correspondence. Furthermore, the space-time scale ratio of most ecological phenomena can surely be altered drastically by human modifications.

For the purpose of scaling, levels of organization or integration are most useful when they are consistent with spatial and temporal scales. Hierarchy theory states that higher levels are larger and slower than lower levels, which is consistent with the space-time principle. This is generally true for *nested hierarchies* (i.e., systems

in which small entities are contained by larger entities which are in turn contained by even larger entities), but not for non-nested hierarchies (Wu 1999). In this view, the three dimensions of scale – space, time and organizational or integrative levels – can be related to each other. When moving up the ladder of hierarchical levels, the characteristic scales of entities or events in both space and time also tend to change accordingly.

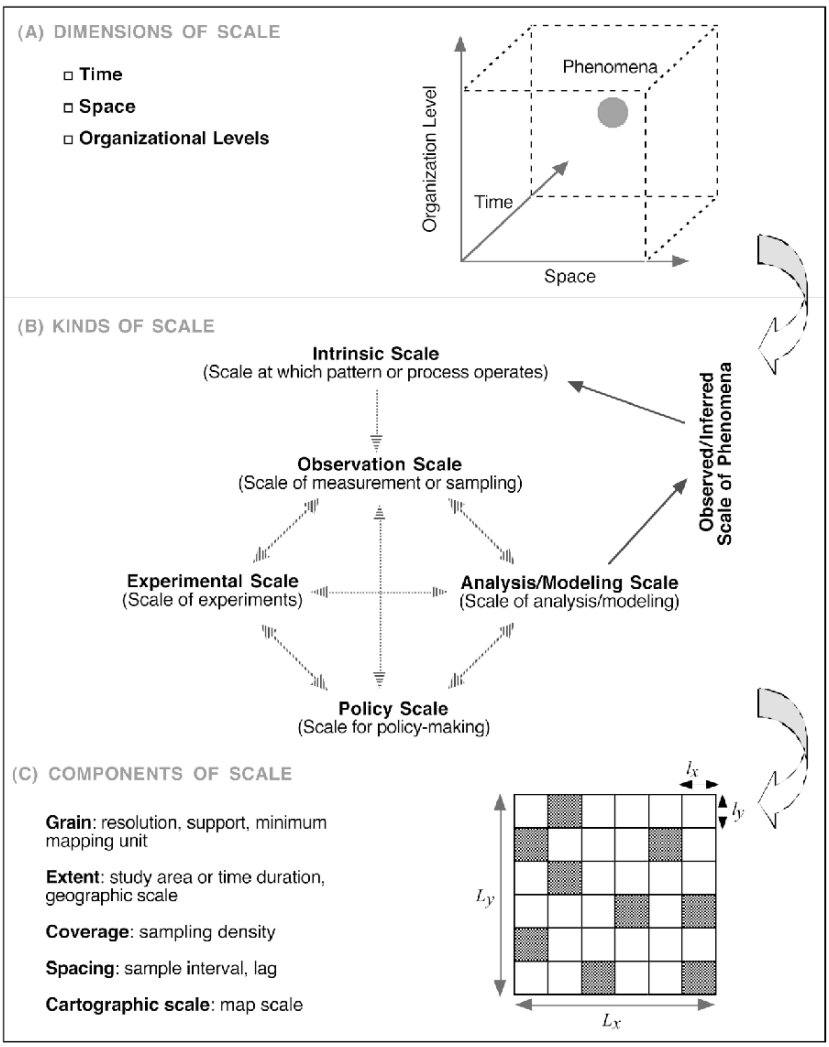


Figure 1.2. *A hierarchy of scale concepts: (A) dimensions of scale, (B) kinds of scale, and (C) components of scale (A was modified from Dungan et al. 2002; B and C were based on Bierkens et al. 2000).*

1.2.2 Kinds of Scale

Several kinds of scale can be distinguished based on any of the three dimensions of scale (Figure 1.2b). *Intrinsic scale* refers to the scale on which a pattern or process actually operates, which is similar to, but broader than, the concept of *process scale*, a term frequently used in earth sciences (e.g., Blöschl and Sivapalan 1995). Some may argue that there is no *intrinsic* scale in nature, and that scales or hierarchical levels are merely epistemological consequences of the observer (Allen and Starr 1992). We believe that the observed scale of a given phenomenon is the result of the interaction between the observer and the inherent scale of the phenomenon. Although the existence of intrinsic scales does not mean that they are always readily observable, a suite of methods, including spectral analysis, fractal analysis, wavelet analysis, scale variance, geostatistics, and multiscale object-specific analysis (e.g., Turner et al. 1991, Wu et al. 2000, Hay et al. 2001, Dale et al. 2002, Hall et al. 2004), have been used in detecting characteristic scales or scale breaks. Effective scale detection requires that the scale of analysis be commensurate with the intrinsic scale of the phenomenon under study (Blöschl and Sivapalan 1995, Wu and Loucks 1995, Dungan et al. 2002, Legendre et al. 2002). Because the latter is unknown *a priori*, multiple observation sets at different scales usually are necessary (Allen et al. 1984, Wu 1999).

There are several other kinds of scale that are not intrinsic to the phenomenon of interest. *Observational scale* is the scale at which sampling or measurement is taken (also referred to as *sampling scale* or *measurement scale*). In experimentation, the spatial and temporal dimensions of the experimental system represent the *experimental scale*, which is the primary criterion for distinguishing among micro-, meso-, and macro-scale experiments. Similarly, the resolution and extent in space and time of statistical analyses and dynamic models define the *analysis scale* or *modeling scale*. In the context of environmental management and planning, local, regional, and national laws and regulations introduce another kind of scale – the *policy scale*, which is influenced by a suite of economic, political, and social factors.

These different kinds of scales are related to each other in various ways (Figure 1.2b). In general, only when the scales of observation and analysis are properly chosen, may the characteristic scale of the phenomenon of interest be detected correctly; only when the scales of experiments and models are appropriate, may the results of experiments and models be relevant; only when the scale of implementation of policies is commensurate with the intrinsic scale of the problem under consideration, may the policies be effective. In reality, different kinds of scales may differ even for the same phenomenon, resulting in the problem of scale mismatch (or scale discordance). To rectify such scale mismatch or to relate one type of scale to the other usually involves scale transfer or scaling (Bierkens et al. 2000). An adequate understanding of the relationship among the different kinds of scale needs to invoke the definitions of scale components.

1.2.3 Components of Scale

Dimensions of scale and kinds of scale are useful general concepts, but more specific and measurable definitions are required in order to quantify scale and develop scaling relations. These are the *components of scale*, including cartographic scale, grain, extent, coverage, and spacing (Figure 1.2c). The traditional *cartographic scale* (or *map scale*) is the ratio of map distance to actual distance on the earth surface. A so-called large-scale map usually covers a smaller area with greater detail. Cartographic scale is essential for the creation and use of maps, but inadequate for studying the scale-dependent relationships between pattern and process in ecology because of its intended rigid connotation (Jenerette and Wu 2000).

In ecology and other earth sciences, scale most frequently refers to grain and extent – two primary components of scale. *Grain* is the finest resolution of a phenomenon or a data set in space or time within which homogeneity is assumed, whereas extent is the total spatial or temporal expanse of a study (Turner et al. 1989a, Wiens 1989). Grain may be considered as the pixel size for raster data, or the minimum mapping unit for vector data. A frequently used geostatistical term, *support*, refers to the smallest area or volume over which the average value of a variable is derived (Dungan et al. 2002). In most cases, grain and support have quite similar meanings, and thus have often been used interchangeably. However, support may differ from grain because support itself includes not only the size of an n-dimensional volume, but also its geometrical shape, size and orientation (Dungan et al. 2002). When the linear or areal dimension of grain is referred to, *grain element* or *grain unit* can be used, which corresponds to *support unit* in the literature. Note that soil scientists and hydrologists frequently use scale only to refer to support (e.g., Bierkens et al. 2000).

On the other hand, the concept of *extent* is less diversified than grain. A term equivalent to extent is *geographic scale*, which was defined by Lam and Quattrochi (1992) as the size of a particular map. Both grain and extent are of great importance to the study of heterogeneous landscapes (Turner 1989). Besides grain and extent, coverage and spacing, which are associated particularly with sampling, are also important in scaling. *Coverage*, not to be confused with extent, refers to sampling intensity in space or time (Bierkens et al. 2000), while *spacing* is the interval between two adjacent samples or lag. Spatial coverage can be represented as the ratio of the sampled area to the extent of a study, and spacing may be fixed or variable depending on the sampling scheme (Figure 1.2c). Support, extent, and spacing are sometimes called the *scale triplet* in hydrological literature, which highlights the importance of these three components in scaling (Blöschl and Sivapalan 1995).

The relationship between intrinsic scale and other kinds of scales can be further elaborated in terms of scale components. Hierarchy theory suggests that the scale of observation must be commensurate with the scale of the phenomenon under consideration if the phenomenon is to be properly observed (Simon 1973, Allen et al. 1984, O'Neill et al. 1986, Wu 1999). On the one hand, processes larger than the extent of observation appear as trends or constants in the observation set; on the

other hand, processes smaller than the grain size of observation become noise in the data. Thus, the choice of a particular scale for observation, analysis and modeling in terms of grain size and extent directly influences whether or not the intrinsic pattern and scale of a phenomenon can be eventually revealed in the final analysis. The significance of the choice of scale has long been recognized in plant ecology (e.g. Greig-Smith 1983) and human geography (Openshaw 1984, Jelinski and Wu 1996). In general, the grain size of sampling or observation should be smaller than the spatial or temporal dimension of the structures or patterns of interest, whereas it is desirable to have the sampling extent at least as large as the extent of the phenomenon under study (Dungan et al. 2002).

In addition, the concept of *relative scale* can be rather useful for comparative studies and scaling across different ecosystems or landscapes. Meentemeyer (1989) defined relative scale as the relationship between the smallest distinguishable unit and the extent of the map, which can be expressed simply as a ratio between grain and extent. Schneider (2001) used *range* to refer to extent, and defined *scope* as the ratio of the range to the resolution of a research design, a model, or a process. In principle, different phenomena and research designs can be compared on the basis of their scopes. Relative scale can also be defined by directly incorporating the ecological pattern and process under consideration. Such definition is rooted in the conceptualization of *relative* versus *absolute space* (Meentemeyer 1989, Marceau 1999). For example, Turner et al. (1989b) considered relative scale as "a transformation of absolute scale to a scale that describes the relative distance, direction, or geometry based on some functional relationship (e.g., the relative distance between two locations based on the effort required by an organism to move between them)."

1.3 CONCEPT OF SCALING

Scaling has been defined differently in various fields of study, and its meanings can be quite disparate. Scaling has long been associated with measurement that is "the assignment of numerals to objects or events according to rules" (Stevens 1946). In this case, scaling is a way of measuring the "unmeasurable" (Torgerson 1958). In multivariate statistics, scaling usually refers to a set of techniques for data reduction and detection of underlying relationships between variables. Multivariate statistical methods, such as polar ordination, multidimensional scaling, principal component analysis, and correspondence analysis, have been used extensively in vegetation classification and ordination to organize field plots (or community types) into some order according to their similarities (or dissimilarities) on the basis of species composition. *Multidimensional scaling*, in particular, is used to represent similarities among objects of interest through visual representation of Euclidean space-based patterns, and has been widely used to analyze subjective evaluations of pairwise similarities of entities in a wide range of fields, including psychology, marketing, sociology, political science, and biology (Young and Hamer 1994). These multivariate statistical methods can be useful for relating patterns and processes across scales (e.g., multiscale ordination; ver Hoef and Glenn-Lewin 1989). However, the concept of scaling as either the assignment of numerical values to

qualitative variables or the reduction and ordination of data is not directly relevant to scaling as defined below.

In physical sciences, *scaling* usually refers to the study of how the structure and behavior of a system vary with its size, and this often amounts to the derivation of a *power-law* relationship. This notion of scaling has often been related to the concepts of similarity, fractals, or scale-invariance, all of which are associated with power laws. For example, a phenomenon or process is said to exhibit "scaling" if it does not have any characteristic length scale; that is, its behavior is independent of scale – i.e., a power law relationship (Wood 1998). This definition of scaling has long been adopted by biologists in terms of *allometry* that primarily correlates the size of organisms with biological form and process (Wu and Li, Chapter 2). In this context, scale refers to "the proportion that a representation of an object or system bears to the prototype of the object or system" (Niklas 1994), and ecological scaling then becomes "the study of the influence of body size on form and function" (LaBarbera 1989). Thus, to some, *ecological scaling* is simply some form of biological allometry (e.g., Calder 1983, Schmidt-Nielsen 1984, LaBarbera 1989, Brown and West 2000).

However, a more general and widely accepted definition of *scaling* in ecology and earth sciences is the translation of information between or across spatial and emporal scales or organizational levels (Turner et al. 1989a, King 1991, Blöschl and Sivapalan 1995, Stewart et al. 1996, van Gardingen et al. 1998, Wu 1999, Bierkens et al. 2000, Gardner et al. 2001). In some cases, this across-scale translation of information can be done through explicit mathematical expressions and statistical relationships (scaling equations), whereas in many other cases process-based mulation models are necessary. This definition of scaling is also referred to as *scale transfer* or *scale transformation* (Blöschl and Sivapalan 1995, Bierkens et al. 2000). This broadly defined scaling concept neither implies that scaling relations must be power-laws, nor that ecological patterns and processes must show scale-independent properties in order to "scale" or to be "scaled." In this case, allometric scaling is but only one special case of scaling. Based on the directionality of the scaling operation, two kinds of scaling can be further distinguished: (1) *scaling up* or *upscaling* which is translating information from finer scales (smaller grain sizes or extents) to broader scales (large grain sizes and extents), and (2) *scaling down* or *downscaling* which is translating information from broader scales to finer scales.

Several other terms are closely related to, but not the same as, scaling. These terms are associated with three basic scaling operations: changing extent, changing grain size, and changing coverage. *Extrapolation* is transferring information from smaller to larger extents, *coarse-graining* transferring information with increasing grain size, and *fine-graining* transferring information with decreasing grain size. Sometimes, upscaling and downscaling refer specifically to coarse-graining and fine-graining, respectively (e.g., Bierkens et al. 2000). When dealing with spatial data that do not have 100% coverage, one may need to estimate the values of unmeasured spatial locations using information from measured sites – a process called *interpolation*. The reverse process of interpolation is *sampling*. In practice, the three basic operations may all be needed in a single study. That is, different

methods for interpolation, sampling, coarse-graining, fine-graining, and extrapolation may be used together to achieve the overall goal of scaling. In general, to make the concept of scale operational, one needs to be specific about the scale components (e.g., grain, extent, coverage, spacing). To put the concept of scaling into action, one has to invoke specific scaling operations (e.g., extrapolation, coarse-graining, fine-graining, interpolation). Any spatial scaling approach or method will inevitably involve one or more of the basic scaling operations.

Note that the definition of extrapolation given above is quite specific and unequivocal. However, in the literature, extrapolation in space has been used in at least four distinct ways: (1) using known data acquired from certain locations to estimate unknown values or draw inferences at other locations, (2) estimating values or drawing inferences about things that fall outside the study area, (3) transferring information from one scale to another in terms of either extent or grain, and (4) transferring information between different systems at the same spatial scale (Turner et al. 1989a, Wu 1999). The multiple meanings of extrapolation may cause confusions. For example, the first usage is simply spatial interpolation. The second is consistent with the definition of spatial extrapolation as information transfer with increasing extent. The third is extremely broad and may refer to coarse-graining, fine-graining, or scaling in general. The fourth usage makes sense with regard to the literal meaning of the word, but it does not fit the definition of scaling because scaling has to involve at least two or more scales. Hence, the term extrapolation should be used with caution.

1.4 WHY SCALING AND HOW?

Simply put, scaling is the essence of prediction and understanding, and is at the heart of ecological theory and application (Levin 1992, Levin and Pacala 1997, Wu 1999, Chave and Levin 2003). More specifically, two main reasons are commonly recognized. First, scaling is inevitable in research and practice whenever predictions need to be made at a scale that is different from the scale where data are acquired. In general, whenever information is averaged over space or time, scaling is at work. For example, the sampling plots that ecologists usually use for determining the distribution of organisms or the stocks and fluxes of materials are only a small portion of the spatial extent of ecological systems of interest. Thus, system-level descriptions dictate the translation of information from these small plots to much larger areas. Also, while most ecological studies traditionally have been conducted on local scales, environmental and resource management problems often have to be dealt with on much broader scales (i.e., landscapes, regions, or the entire globe). To bridge such scale gaps requires scaling.

Second, because ecological phenomena occur over a wide range of scales and because there are often hierarchical linkages among them, relating information across scales as well as levels of biological organization is an essential part of ecological understanding. For example, the dynamics of sub-watershed units and their interactions are crucial to understanding the hydrological and biogeochemical cycles of the whole watershed ecosystem (Wickham et al., Chapter 12). The dynamics of local populations and their interpatch interactions are crucial to

understanding population dynamics at the landscape scale. In a similar vein, understanding the primary productivity of the whole ecosystem requires knowledge of photosynthesis at the individual leaf level.

While it is imperative in almost all ecological studies, spatial scaling can also be extremely challenging in theory and practice. Spatial heterogeneity can greatly complicate the scaling process. Spatial heterogeneity may manifest itself in terms of various patterns of land use and land cover, topography, hydrology, soils, climatic conditions, and biological factors. For example, extrapolation of plot-scale data to the landscape or regional scale is a trivial matter in a spatially homogeneous (uniform or random) environment. In a heterogeneous landscape, however, simply multiplying the plot-scale average with the total study area usually provides a rather poor estimate at the landscape scale (Li and Wu, Chapter 3). When ecological relationships are translated across scales in heterogeneous environments, they often become distorted – a phenomenon known as "spatial transmutation" (*sensu* O'Neill 1979, King et al. 1991, Wu and Levin 1994).

Also, as scale changes, new patterns and processes may emerge, and controlling factors may shift even for the same phenomena. Thus, observations made at fine scales may miss important patterns and processes operating on broader scales. Conversely, broad-scale observations may not have enough details necessary to understand fine-scale dynamics. In addition, nonlinear interactions, time delays, feedbacks, and legacies in ecological systems may impose formidable challenges for translating information across scales or levels of organization (O'Neill and Rust 1979, Wu 1999). Therefore, on the one hand, spatial heterogeneity, scale multiplicity, and nonlinearity are important sources of biodiversity and ecological complexity; on the other hand, they are major hurdles for successful scaling.

Given the various obstacles, how should we proceed with scaling? This is the focus of our next chapter, where we will discuss two general scaling approaches: similarity-based and dynamic model-based scaling. A dozen specific scaling methods will also be examined in terms of their assumptions, ways of dealing with spatial heterogeneity and nonlinear interactions, and accuracy of scaling results. No matter which approach is used, an important concept in scaling up and down is *scaling threshold* or *scaling break*, which signifies a narrow range of scale around which scaling relations change abruptly. A scaling threshold may also be understood as a critical scale of a phenomenon where emergent properties due to nonlinear interactions and spatial heterogeneity come into effect. Thus, scaling thresholds, when properly identified, may reflect fundamental shifts in underlying processes or controlling factors, and can be used to define the domains of applicability of specific scaling methods.

1.5 DISCUSSION

In this chapter, we have discussed and clarified a number of concepts related to scale and scaling as used in a variety of fields of study. We propose a hierarchical framework in which the different connotations of scale can be organized with clarity and consistency. The three-tiered definitional hierarchy, consisting of the dimensions, kinds, and components of scale, shows both the diversity and interrelatedness of the

concepts of scale. In the practice of scaling, the components of scale (most frequently extent, grain, and coverage) must be invoked. Indeed, scaling methods are often designed to capture and deal with the change in these scale components singularly or in concert (see Wu and Li, Chapter 2 for details).

Clarification of key concepts is the first step towards a science of scale. The three-tiered definitional hierarchy seems to serve this purpose well even though it is only one of many possible ways of organizing these concepts. It is crucial for ecologists to recognize the different usages of scale and scaling, and to adopt a system of definitions that are consistent, clear, and accommodating to the development of quantitative methods. The science of scale will certainly benefit from clear concepts and definitions, which are essential for the development of effective methods and sound theories of scaling.

ACKNOWLEDGEMENTS

We would like to thank Geoffrey Hay, Fangliang He, Bruce Jones, and Simon Levin for their comments on an earlier version of the chapter. JW's research on scaling has been supported in part by grants from U.S. Environmental Protection Agency's Science to Achieve Results (STAR) Program (R827676-01-0) and National Science Foundation (DEB 9714833 for Central Arizona-Phoenix Long-Term Ecological Research).

REFERENCES

Allen, T. F. H., and T. B. Starr. 1982. Hierarchy: Perspectives for Ecological Complexity. University of Chicago Press, Chicago.

Allen, T. F. H., R. V. O'Neill, and T. W. Hoekstra. 1984. Interlevel relations in ecological research and management: some working principles from hierarchy theory. USDA Forest Service General Tech Report RM-110, Rocky Mountain Forest and Range Experiment Station.

Bierkens, M. F. P., P. A. Finke, and P. de Willigen. 2000. Upscaling and Downscaling Methods for Environmental Research. Kluwer Academic Publishers, Dordrecht.

Blöschl, G., and M. Sivapalan. 1995. Scale issues in hydrological modelling: a review. Hydrological Processes 9:251-290.

Brown, J. H., and G. B. West, editors. 2000. Scaling in Biology. Oxford University Press, New York.

Calder, W. A. 1983. Ecological scaling: mammals and birds. Annual Reviews Ecology & Systematics 14:213-230.

Chave, J., and S. A. Levin. 2003. Scale and scaling in ecological and economic systems. Environmental and Resource Economics 26:527-557.

Clark, W. C. 1985. Scales of climate impacts. Climatic Change 7:5-27.

Courtois, P. J. 1985. On time and space decomposition of complex structures. Communications of the ACM 28:590-603.

Dale, M. R. T., P. Dixon, M.-J. Fortin, P. Legendre, D. E. Myers, and M. S. Rosenberg. 2002. Conceptual and mathematical relationships among methods for spatial analysis. Ecography 25:558-577.

Delcourt, H. R., and P. A. Delcourt. 1988. Quaternary landscape ecology: relevant scales in space and time. Landscape Ecology 2:23-44.

Dungan, J. L., J. N. Perry, M. R. T. Dale, P. Legendre, S. Citron-Pousty, M.-J. Fortin, A. Jakomulska, M. Miriti, and M. S. Rosenberg. 2002. A balanced view of scale in spatial statistical analysis. Ecography 25:626-640.

Gardner, R. H., W. M. Kemp, V. S. Kennedy, and J. E. Petersen, editors. 2001. Scaling Relations in Experimental Ecology. Columbia University Press, New York.

Greig-Smith, P. 1983. Quantitative Plant Ecology. University of California Press, Berkeley.

Hall, O., G. J. Hay, A. Bouchard, and D. J. Marceau. 2004. Detecting dominant landscape objects through multiple scales: an integration of object-specific methods and watershed segmentation. Landscape Ecology 19:59-76.

Hay, G., D. J. Marceau, P. Dubé, and A. Bouchard. 2001. A multiscale framework for landscape analysis: object-specific analysis and upscaling. Landscape Ecology 16:471-490.

Jelinski, D. E., and J. Wu. 1996. The modifiable areal unit problem and implications for landscape ecology. Landscape Ecology 11:129-140.

Jenerette, G. D., and J. Wu. 2000. On the definitions of scale. Bulletin of the Ecological Society of America 81:104-105.

King, A. W. 1991. Translating models across scales in the landscape. Pages 479-517 *in* M. G. Turner and R. H. Gardner, editors. Quantitative Methods in Landscape Ecology. Springer-Verlag, New York.

King, A. W., A. R. Johnson, and R. V. O'Neill. 1991. Transmutation and functional representation of heterogeneous landscapes. Landscape Ecology 5:239-253.

LaBarbera, M. 1989. Analyzing body size as a factor in ecology and evolution. Annual Reviews Ecology and Systematics 20:97-117.

Lam, N. S.-N., and D. A. Quattrochi. 1992. On the issues of scale, resolution, and fractal analysis in the mapping sciences. Professional Geographer 44:88-98.

Legendre, P., M. R. T. Dale, M.-J. Fortin, J. Gurevitch, M. Hohn, and D. Myers. 2002. The consequences of spatial structure for the design and analysis of ecological field surveys. Ecography 25:601-615.

Levin, S. A. 1992. The problem of pattern and scale in ecology. Ecology 73:1943-1967.

Levin, S. A., and S. W. Pacala. 1997. Theories of simplification and sclaing of spatially distributed processes. Pages 271-295 *in* D. Tilman and P. Kareiva, editors. Spatial Ecology. Princeton University Press, Princeton.

Marceau, D. J. 1999. The scale issue in social and natural sciences. Canadian Journal of Remote Sensing 25:347-356.

Meadowcroft, J. 2002. Politics and scale: some implications for environmental governance. Landscape and Urban Planning 61:169-179.

Meentemeyer, V. 1989. Geographical perspectives of space, time, and scale. Landscape Ecology 3: 163-173.

Niklas, K. J. 1994. Plant Allometry: The Scaling of Form and Process. University of Chicago Press, Chicago.

O'Neill, R. V. 1979. Natural variability as a source of error in model predictions. Pages 23-32 *in* G. S. Innis and R. V. O'Neill, editors. Systems Analysis of Ecosystems. International Co-operative Publishing House, Fairland, Maryland.

O'Neill, R. V., and B. Rust. 1979. Aggregation error in ecological models. Ecological Modelling 7: 91-105.

O'Neill, R. V., D. L. DeAngelis, J. B. Waide, and T. F. H. Allen. 1986. A Hierarchical Concept of Ecosystems. Princeton University Press, Princeton.

O'Neill, R. V., R. H. Gardner, B. T. Milne, M. G. Turner, and B. Jackson. 1991. Heterogeneity and Spatial Hierarchies. Pages 85-96 *in* J. Kolasa and S. T. A. Pickett, editors. Ecological Heterogeneity. Springer-Verlag, New York.

Openshaw, S. 1984. The Modifiable Areal Unit Problem. Geo Books, Norwich.

Peterson, D. L., and V. T. Parker, editors. 1998. Ecological Scale: Theory and Applications. Columbia University Press, New York.

Pierce, L. L., and S. W. Running. 1995. The effects of aggregating sub-grid land surface variation on large-scale estimates of net primary production. Landscape Ecology 10:239-253.

Schmidt-Nielsen, K. 1984. Scaling: Why Is Animal Size So Important? Cambridge University Press, Cambridge.

Schneider, D. C. 2001. Spatial allometry. Pages 113-153 *in* R. H. Gardner, W. M. Kemp, V. S. Kennedy, and J. E. Petersen, editors. Scaling Relations in Experimental Ecology. Columbia University Press, New York.

Simon, H. A. 1973. The organization of complex systems. Pages 1-27 *in* H. H. Pattee, editor. Hierarchy Theory: The Challenge of Complex Systems. George Braziller, New York.

Stevens, S. 1946. On the theory of scales of measurement. Science 103:677-680.

Stewart, J. B., E. T. Engman, R. A. Feddes, and Y. Kerr, editors. 1996. Scaling up in Hydrology Using Remote Sensing. Wiley, Chichester.

Stommel, H. 1963. Varieties of oceanographic experience. Science 139:572-576.

Torgerson, W. S., editor. 1958. Theory and Methods of Scaling. John Wiley, New York.
Turner, M. G. 1989. Landscape ecology: the effect of pattern on process. Annual Review of Ecology and Systematics 20:171-197.
Turner, M. G., V. H. Dale, and R. H. Gardner. 1989a. Predicting across scales: theory development and testing. Landscape Ecology 3:245-252.
Turner, M. G., R. V. O'Neill, R. H. Gardner, and B. T. Milne. 1989b. Effects of changing spatial scale on the analysis of landscape pattern. Landscape Ecology 3:153-162.
Turner, S. J., R. V. O'Neill, W. Conley, M. R. Conley, and H. C. Humphries. 1991. Pattern and scale: statistics for landscape ecology. Pages 17-49 *in* M. G. Turner and R. H. Gardner, editors. Quantitative Methods in Landscape Ecology. Springer-Verlag, New York.
Urban, D. L., R. V. O'Neill, and H. H. Shugart. 1987. Landscape ecology: a hierarchical perspective can help scientists understand spatial patterns. BioScience 37:119-127.
van Gardingen, P. R., G. M. Foody, and P. J. Curran, editors. 1997. Scaling-Up: From Cell to Landscape. Cambridge University Press, Cambridge.
ver Hoef, J. M., and D. C. Glenn-Lewin. 1989. Multiscale ordination: a method for detecting pattern at several scales. Vegetatio 82:59-67.
White, J. D., and S. Running. 1994. Testing scale dependent assumptions in regional ecosystem simulations. Journal of Vegetation Science 5:687-702.
Wiens, J. A. 1989. Spatial scaling in ecology. Functional Ecology 3:385-397.
Wood, E. F. 1998. Scale analyses for land-surface hydrology. Pages 1-29 *in* G. Sposito, editor. Scale Dependence and Scale Invariance in Hydrology. Cambridge University Press, Cambridge.
Wu, J. 1999. Hierarchy and scaling: extrapolating information along a scaling ladder. Canadian Journal of Remote Sensing 25:367-380.
Wu, J. 2004. Effects of changing scale on landscape pattern analysis: scaling relations. Landscape Ecology 19:125-138.
Wu, J., and S. A. Levin. 1994. A spatial patch dynamic modeling approach to pattern and process in an annual grassland. Ecological Monographs 64(4):447-464.
Wu, J., and O. L. Loucks. 1995. From balance-of-nature to hierarchical patch dynamics: a paradigm shift in ecology. Quarterly Review of Biology 70:439-466.
Wu, J., D. Jelinski, M. Luck, and P. Tueller. 2000. Multiscale analysis of landscape heterogeneity. Geographic Information Sciences 6:6-19.
Young, F. W., and R. M. Hamer. 1994. Theory and Applications of Multidimensional Scaling. Eribaum Associates., Hillsdale, New Jersey.

CHAPTER 2

PERSPECTIVES AND METHODS OF SCALING

JIANGUO WU AND HARBIN LI

2.1 INTRODUCTION

Transferring information between scales or organizational levels is generally referred to as "*scaling*" (Wu and Li, Chapter 1), and is inevitable in both basic research and its applications. Scaling is the essence of prediction and understanding both of which require cross-scale translation of information, and is at the core of ecological theory and application (Levin 1992, Levin and Pacala 1997, Wu 1999). While the importance of scaling in ecology has been acutely recognized in recent decades, how to conduct scaling across heterogeneous ecosystems remains a grand challenge (Turner et al. 1989, Wu and Hobbs 2002).

A number of scaling approaches and methods have been developed and applied in different disciplines ranging from physics, engineering, biology, to social sciences. Two general scaling approaches can be distinguished: similarity-based scaling and dynamic model-based scaling methods (Blöschl and Sivapalan 1995). *Similarity-based scaling methods* are rooted in the concepts and principles of similarity and self-similarity and often characterized by relatively simple mathematical or statistical scaling functions, even though the underlying ecological processes of a phenomenon may be extremely complex. In contrast, *dynamic model-based scaling methods* use deterministic or stochastic models to simulate the processes of interest, and to transfer information across scales by either modifying the parameters and input variables of the same model or developing multiple-scaled models. In this case, information transfer between different scales is accomplished through manipulating the inputs, outputs, and formulations of dynamic models. In both approaches, it is important to properly identify scaling thresholds at which scaling relations often change abruptly, reflecting fundamental shifts in underlying processes or controlling factors and defining the domains of applicability of specific scaling methods.

While the previous chapter (Wu and Li, Chapter 1) discussed various concepts of scale and scaling, in this chapter we focus on the major characteristics of the two scaling approaches and several more specific upscaling and downscaling methods

J. Wu, K.B. Jones, H. Li, and O.L. Loucks (eds.),
Scaling and Uncertainty Analysis in Ecology: Methods and Applications, 17–44.

within each approach. The purpose of this chapter is not to provide a recipe for scaling. Rather, we shall review scaling perspectives and methods in different disciplines, and provide a synthesis based on a common conceptual framework. By so doing, we expect that a more comprehensive and cohesive understanding of ecological scaling will emerge.

2.2 SIMILARITY-BASED SCALING METHODS

2.2.1 The Concept of Similarity

The concept of *similarity* has been essential in scaling-related studies. In general, similarity exists between two systems whenever they share some properties that can be related across the systems by a simple conversion factor (Blöschl and Sivapalan 1995). LaBarbera (1989) summarized three types of similarity concepts applied in body size-oriented studies: geometric, physical, and functional similarities (also see Gunther 1975). *Geometric similarity* is characterized by the constancy in shape with changing size. In other words, geometric similarity assumes that "geometry and shape are size-independent properties" (Niklas 1994). For example, for different sized objects of the same shape and geometry, $S \propto L^2$, and $S \propto V^{2/3}$, where *L, S,* and *V* are the linear dimension, surface area, and volume of the objects. *Physical similarity* is defined based on the constancy of the ratios of different forces (also called *dynamic similarity*; see Blöschl and Sivapalan 1995). For example, two systems are said to have *hydrodynamic similarity* if they have the same Reynolds number (i.e., the ratio of inertial to viscous forces). Barenblatt (1996) stated that the concept of physical similarity is a natural generalization of that of geometric similarity in that two similar triangles differ only in the numerical values of side lengths, whereas two similar physical phenomena differ only in the numerical values of the dimensional governing parameters. *Functional similarity* refers to the constancy in changes of functional variables over a range of system sizes. For example, animal metabolic rates (R) change with body size or mass (M) following a power law (i.e., $R \propto M^b$). Similarly, if the primary productivity (P) of a group of ecosystems changes with their spatial extent (A) in a power-law fashion (i.e., $P \propto A^b$), then these ecosystems may be said to have functional similarity.

In recent decades, the concept of *self-similarity* has become a cornerstone of similarity-based studies. It refers to the phenomenon that the whole is composed of smaller parts that resemble the whole itself or that patterns remain similar at different scales. Self-similarity is the key idea in fractal geometry (Mandelbrot 1982, Hastings and Sugihara 1993), and is considered to be the unifying concept underlying fractals, chaos, and power laws (Schroeder 1991). While admitting that the terms, *fractal* and *multifractal*, still lack an agreed mathematical definition, Mandelbrot (1999) offered an informal definition of *fractal geometry* as "the systematic study of certain very irregular shapes, in either mathematics or nature, wherein each small part is very much like a reduced size image of the whole." Such irregular shapes, or fractals, exhibit properties of self-similarity which entails scale-invariance (i.e., patterns or relationships remain unchanged over a range of scales).

Commonly cited examples of fractals include coastlines, clouds, snowflakes, branching trees, and vegetation patches. However, not all self-similar objects are fractals because self-similarity is also found in Euclidean geometry.

Simple fractals exhibit scale-invariant patterns that can be characterized with only one scaling exponent, which is often interpreted as implying one single generating process. However, many fractal-like structures in nature are generated by a number of generating processes that operate at different scales. These are called *generalized fractals*, or *multifractals*, which are characterized by a spectrum of fractal dimensions that vary with scale. It has been suggested that additive processes tend to create *monofractals* (simple fractals), whereas multiplicative random processes generate multifractal structures (Stanley and Meakin 1988, Schroeder 1991). Multifractals have been used to describe the spatial distribution of people and minerals, energy dissipation in turbulence, and many other patterns and processes in nature. It is now widely recognized that many, if not most, fractal patterns and processes in nature show scale-invariance only over a limited range of scales. Hastings and Sugihara (1993) suggested that linear regression methods be used to distinguish between patterns with one scaling region (a single power law) and those with multiple scaling regions (separate power laws over separate regions). These authors asserted that multiscaling is detected if the slope of the regression line changes significantly over adjacent regions.

2.2.2 Dimensional Analysis and Similarity Analysis

The concepts of similarity are the foundation of dimensional analysis (Blöschl and Sivapalan 1995), and have long been used in engineering and physical sciences. Barenblatt (1996) indicated that the main idea behind dimensional analysis is that "physical laws do not depend on arbitrarily chosen basic units of measurement," and thus the functions expressing physical laws must possess some fundamental property (mathematically termed generalized homogeneity or symmetry) that allows the number of arguments in these functions to be reduced. *Dimensional analysis* aims to produce dimensionless ratio-based equations that can be applied at different scales for a phenomenon under study. In practice, dimensional analysis only applies in the framework of Euclidean geometry and Newtonian dynamics (Scheurer et al. 2001). Dimensional techniques have long been used to derive similarity relationships, establish scaling laws, reduce data volume, and help elucidate processes and mechanisms in physical and biological sciences (Gunther 1975, Blöschl and Sivapalan 1995). Like dimensional analysis, *similarity analysis* is also a simplification procedure to replace dimensional quantities required for describing a phenomenon with fewer dimensionless quantities; but unlike dimensional analysis, similarity analysis requires the governing equations of the phenomenon to be known (Blöschl and Sivapalan 1995).

Examples of similarity analysis are abundant in physical sciences. For example, similarity analysis in soil physics and hydrology started in the 1950s with the concept of *Miller-Miller similitude*, an intuitive depiction of structural similarities in porous media at fine spatial scales (Miller and Miller 1956). Miller-similar porous media have microscopic structures that look similar in the same way as triangles in

Euclidean geometry (Sposito 1998). Similarity analysis, then, involves the derivation of scale factors for soil-water transport coefficients on the basis of the fine-scale similar-media concept. Later studies extended the concept of the Miller-Miller similitude from microscopic to macroscopic scales using the idea of functional normalization (related to functional similarity) rather than dimensional techniques (Haverkamp et al. 1998). In recent decades, fractal and multifractal models of soil structure have been increasingly used in similarity analysis of hydrological processes and beyond. As Sposito (1998) noted, "fractal geometry has become the signature approach to both spatial-scale invariance and temporal-scale invariance, as epitomized by self-similarity in the patterns of hydrologic and other geophysical processes."

One of the most successful examples of using similarity analysis to deal with complex physical processes is *Monin-Obukhov similarity theory*. Atmospheric boundary-layer flows, though mostly turbulent, can be viewed as being dynamically similar such that the concepts of similarity can provide a powerful framework for analyzing empirical data and parameterizing models to represent these complex processes. In particular, Monin-Obukhov theory assumes that surface layers with the same ratio of the aerodynamic roughness length (z_0) to the Obukhov length (L) are dynamically similar, with z_0/L being considered as a dimensionless similarity parameter. In other words, the theory is based on the assumption of complete similarity of fluxes in terms of Reynolds number (Barenblatt 1996). The development of Monin-Obukhov theory follows the general procedures of similarity analysis: (1) identifying the atmospheric processes that conform the dynamic similarity principle, (2) characterizing these processes with dimensionless similarity parameters (e.g., Reynolds number), (3) determining a set of scaling parameters (e.g., scaling wind velocity, scaling temperature, scaling humidity) and non-dimensionalized dependent and independent variables, and (4) deriving a set of similarity laws that are valid over a broad range of scales (Barenblatt 1996). By so doing, Monin-Obukhov theory relates turbulent fluxes in the surface layer to mean vertical gradients of wind, temperature, and specific humidity (Wu 1990).

As an important part of similarity analysis, *renormalization group methods* (Wilson 1975) have been used for studying scaling behavior associated with critical phenomena and phase transitions in physical sciences, including turbulence, flows in porous media, fracture mechanics, flame propagation, atmospheric and oceanic processes (Binney et al. 1993, Barenblatt 1996). The general idea of renormalization groups is to simplify mathematically complex models that contain much fine-scale detail into simpler models and to develop scaling laws using similarity principles and techniques. The simpler models (or equations) consist only of essential information of the phenomenon under study, and are able to describe and predict coarse-scale patterns with explicit scaling relations. Renormalization group methods represent a fundamental concept and powerful technique in theoretical physics (Barenblatt 1996), which "make rigorous the scaling process through the derivation of equations for blocks of cells in terms of the units that make them up" (Levin and Pacala 1997). Critical phenomena and phase transitions are common in ecology, particularly with spatial problems (Gardner et al. 1987, Milne 1998), but only until

recently have renormalization group methods been applied in ecological studies (e.g., Levin and Pacala 1997, Milne 1998).

Gunther (1975) pointed out that "Structures and functions of all living beings, irrespective of their size, can be studied by means of some basic physical methods, viz., dimensional analysis and theories of similarity." Although it is unlikely that all structures and functions of the biological world (even at the organism level) can be effectively studied by using dimensional analysis and similarity analysis alone, there is little doubt that they will continue to play an important role in biological and ecological scaling. A great number of allometric studies in biology and ecology have further demonstrated the power and elegance of similarity-based methods. However, the applicability and accuracy of these methods may depend on the levels of biological organization and the variability of processes with scale. In the following, we turn our attention to some of the major issues in allometric scaling.

2.2.3 Biological Allometry

Gould (1966) defined *allometry* as "the study of size and its consequences." Similarly, Niklas (1994) described allometry as "the study of size-correlated variations in organic form and process." Among other definitions of allometry is any "departure from geometric similarity" (LaBarbera 1989). For several decades allometry has focused primarily on the body size (or mass) of organisms as the fundamental variable (Calder 1983, Peters 1983, Schmidt-Nielsen 1984). Niklas (1994) summarized three meanings of allometry: (1) a relationship between the growth of a part of an organism and the growth of the whole organism, (2) a relationship between organism size and biological form and process, and (3) a size-correlated relationship deviating from geometric similarity that is exhibited by objects of varying sizes with the same geometry and shape. Brown et al. (2000) noted that allometric studies in biology have been carried out at three levels of biological organizations: within individual organisms (e.g., animal circulatory networks and tree branching architecture), among individual organisms of different sizes (e.g., body-size related variations in biological pattern and process), and within populations or communities (e.g., allometric scaling of population density and community biomass).

Allometric scaling is rooted in the concepts of similarity and, as in physical sciences, allometric relations in biology usually take the form of a power law:

$$Y = Y_0 M^b \tag{2.1}$$

$$\text{or, } \log Y = \log Y_0 + b \log M \tag{2.2}$$

where Y is some biological variable, Y_0 is a normalization (or scaling) constant, M is some size-related variable (usually body mass), and b is the *scaling exponent*.

In Equation 2.1, if $b = 1$, the relationship becomes linear, and is called *isometric scaling*; if $b \neq 1$, then the relationship is either *geometric scaling* or *allometric*

scaling (including fractal scaling). Geometric (or Euclidean) scaling is based on *complete similarity*, whereas allometric scaling is based on *incomplete similarity* or self-similarity (Barenblatt 1996, Schneider 2001a). For example, based on the geometric similarity of Euclidean objects we can analytically derive the following relationships among volume (V), area (A), the length dimension (l), and mass (M): $A \propto l^2$, $V \propto l^3$, $M \propto V$, $l \propto M^{1/3}$, and $A \propto M^{2/3}$. These simple geometric scaling rules mean that, if objects of different sizes are completely similar, their liner dimensions and surface areas should be proportional to the 1/3 and 2/3 powers of their mass (assuming a constant density). In other words, if $b = 1/3$, Equation 2.1 suggests that a property of an object (Y) is dependent on the length dimension of the object (M); if $b = 2/3$, then Equation 2.1 suggests that Y is dependent on the surface area of the object. However, Brown et al. (2000, 2002), among others, argued that organisms do not seem to follow such simple geometric scaling rules; rather, they commonly exhibit "quarter-power scaling" relationships – i.e., the scaling exponent takes the value of simple multiples of 1/4. For example, $b = 3/4$ for the whole-organism metabolic rates of a variety of animals ranging from mice to elephants; $b = 1/4$ for the heart rates of animals; $b = -1/4$ for the life span of animal species; $b = 3/8$ for the radius of the aorta of animals and the trunks of trees; and $b = -3/4$ for the population density of animals (Brown et al. 2000, Schmid et al. 2000, Carbone and Gittleman 2002). While these scaling relations are general, variability can be substantial even for the same biological process. For instance, LaBarbera (1989) reported that, for scaling of home range area with body size of terrestrial mammals, $b = 1.18$ for herbivores, $b = 1.51$ for carnivores, $b = 0.97$ for omnivores, and $b =$ 0.74, 1.39, or 1.65 for all mammals depending on data sets used for calculation.

One of the best-known examples of allometric scaling in plant ecology is the self-thinning law in plants. In even-aged plant communities, the average biomass of individual plants (W) scales with plant density (D) following a power law: $W = cD^{-3/2}$, or $B = cD^{-1/2}$, where c is a scaling constant and B (=WD) is the stand biomass density. This means that plant population density scales with plant weight with a scaling exponent of −2/3 (i.e., $D \propto W^{-2/3}$) rather than −3/4 as in animals. This scaling relation was obtained from regression analysis based on empirical data as well as analytical studies based on geometric similarity – the so-called "surface area law" ($S \propto V^{2/3}$, where V is the volume and S is the surface area; Niklas 1994). While this biomass-density relation has been held as a "law" for decades, recent studies have found little empirical evidence to support its universality and consistency (Weller 1987, Zeide 1987, Lonsdale 1990). In particular, the scaling exponent is not a constant, but rather a variable that is influenced by the shade tolerance of plants under study and taxonomic groups of choice. Zeide (1987) concluded that "the law is neither precise nor accurate," and Lonsdale (1990) stated that, in the log-log plot of stand biomass vs. plant density, "straight lines are the exception rather than the rule."

Enquist et al. (1998) showed that whole-plant resource use scales as $W^{3/4}$ and that, accordingly, the scaling exponent for the biomass-density relation or the self-thinning law is −3/4 (i.e., $D \propto W^{-3/4}$), not −2/3 as previously reported. Thus, they

concluded that plants do not differ from animals in terms of scaling of population density with respect to body mass, confirming the prediction of their general mechanistic model of resource use in fractal-like branching networks (West et al. 1997). This model, however, has met an increasing number of criticisms claiming that it is mathematically flawed and empirically unwarranted (e.g., Magnani 1999, Bokma 2004, Cyr and Walker 2004, Kozlowski and Konarzewski 2004). Nevertheless, allometric scaling, as a general approach, remains useful, and its rule in spatial scaling is discussed below.

2.2.4 Spatial Allometry

While sharing common features of similarity-based scaling methods, biological allometry has focused primarily on body size. Most of the allometric equations do not directly address the problem of spatial scaling. However, allometry as a general method can be applied to spatial scaling when the independent variable is spatial scale instead of body mass. Such studies have been termed *spatial allometry* (Schneider 2001a, b). In this case, the similarity principles pertain to the spatially extended systems (e.g., habitats, landscapes) rather than the individual organisms. A general spatial allometric scaling relation can be written as follows:

$$Q(S) = Q(S_0)\left(\frac{S}{S_0}\right)^{\beta} \tag{2.3}$$

$$\text{or,}\quad Q(S) = kS^{\beta}, \text{ with } k = Q(S_0)S_0{}^{-\beta} \tag{2.4}$$

where $Q(S)$ and $Q(S_0)$ are the values of an ecological variable Q at spatial scales of S_0 and S, respectively, and β is the scaling exponent.

In Equations 2.3 and 2.4, S and S_0 may be expressed as extent or grain size. If S is extent and S_0 is grain size, then the ratio, S/S_0, defines the spatial (or temporal) scope (*sensu* Schneider 2001a), which is useful for comparing scaling studies among different systems. As with Equation 2.1, Equation 2.3 indicates *isometric scaling* when $\beta = 1$, and *geometric* (Euclidean) or *fractal scaling* when $\beta \neq 1$. Schneider (2001b) pointed out that geometric scaling results when β is "an integer or ratio of integers," whereas fractal scaling is indicated by a value of β that is "not an integer." In practice, however, it is not a trivial matter to distinguish between a "ratio of integers" and a "fractal" dimension. Thus, inferring the nature of similarity based merely on regression results, as often done in biological allometry, is not warranted.

Some allometric relations at the levels of populations and communities may be directly related to spatial scaling. For example, if population density scales with body mass as $D = D_0M^{-0.75}$, one can derive a scaling relation between the total number of animals (N) and habitat area (A): $N = D_0AM^{-0.75}$ or between the total biomass (B) of the animal species and habitat area: $B = DAM = D_0AM^{0.25}$. If home

range scales with body mass as $H \propto M^b$, then population density can be directly related to the size of home range: $D \propto HM^{-(0.75+b)}$. The best known example of spatial allometry, however, may well be the *species-area relationship* (SAR). SAR is commonly described by a power-law function: $S = cA^z$, where c is a constant influenced by the effect of geographical variations on S, and z is the scaling exponent with a value close to 0.25. SAR has been regarded as "ecology's most general, yet protean pattern" (Lomolino 2000) and one of the few widely accepted laws in ecology (Schoener et al. 2001).

Some recent studies suggested that SAR is an example of scale invariance that reflects self-similarity in species abundance and distribution (e.g., Harte et al. 1999, Kunin 1999). However, many others have indicated that the value of the scaling exponent of SAR may vary widely and that the power-law scaling only holds over a finite range of spatial scales in real landscapes (Crawley and Harral 2001, Schoener et al. 2001). While scale invariant pattern is often believed to imply a single underlying process, SAR may have multiple scaling domains if examined over many orders of magnitude in space. This observation favors the explanation that different factors determine species diversity at different ranges of scales (Shmida and Wilson 1985, Crawley and Harral 2001, Whittaker et al. 2001). For example, Lomolino (2000) argued that, for isolated ecosystems, SAR has three fundamentally different realms: (1) erratic changes influenced by idiosyncratic differences among islands and random catastrophic disturbance events for small islands, (2) a monotonic deterministic pattern determined by island area and associated ecological factors for intermediate-sized islands, and (3) again a monotonically increasing pattern for islands large enough for *in situ* speciation. Nevertheless, as with the self-thinning law, the debate and controversies on the universality, scale invariance, and ecological interpretation of SAR do not necessarily invalidate the use of the allometric scaling approach; it actually demonstrates its usefulness as a research tool.

In landscape geomorphology, it has long been noted that landform attributes exhibit allometric relationships (Woldenberg 1969, Bull 1975, Church and Mark 1980). For example, Hood (2002) identified several allometric scaling relations between slough attributes (e.g., area, outlet width, perimeter, length) for rivers in the Pacific Northwest of the United States, and showed that detrital insect flotsam density was also allometrically related to slough perimeter. In a recent study of the landscape dynamics of over 640 peatland bog pools in northern Scotland, Belyea and Lancaster (2002) found that the pools became deeper and more convoluted in shape with increasing size, and that the relationships between the area, depth, width, and length of the bog pools showed allometric (rather than geometric) scaling. Schneider (2001a,b) provided a number of examples of spatial allometry for lake ecosystems and aquatic mesocosms in terms of the geometric attributes of the systems (e.g., volume, area, perimeter, and depth of lakes or mesocosms) and biological properties (e.g., fish catch, primary production). In recent decades, the allometric study of landform, or *landscape allometry*, has been elevated to a new level of enthusiasm and insight by applying the concepts of fractals and

self-organization (Mandelbrot 1982, Turcotte 1995, Rodriguez-Iturbe and Rinaldo 1997, Phillips 1999, Schneider 2001a, b).

In landscape ecology, there have been many examples of spatial patterns exhibiting allometric or fractal scaling relations (e.g., Milne 1991, Nikora et al. 1999, Wu 2004). Although some authors attempt to associate power scaling relations to underlying "universal" laws or scale invariance theories, such scaling relations usually only hold for limited ranges of scale (Milne 1991, Berntson and Stoll 1997, Wu 2004). Without resorting to any such grandiose assumptions, however, spatial allometry can still be used as a valuable empirical scaling method to summarize and extrapolate observed patterns over a range of scales, and to provide clues about the underlying processes, using a "scalogram approach" (Wu 2004).

2.3 DYNAMIC MODEL-BASED SCALING METHODS

2.3.1 Some Concepts of Scaling with Dynamic Models

In contrast with similarity-based scaling methods that deal with complex phenomena in a relatively simple manner, dynamic model-based scaling methods focus more on the processes and mechanisms of the phenomena under study. They may incorporate, but do not rely on, similarity concepts in theory and dimensional techniques in practice. *Dynamic models* are composed of state variables, rate variables, input variables, output variables, parameters, and constants. Parameters and constants help define rate variables and relate input and output variables to state variables. Because these terms are defined differently in the literature, some clarifications are needed here to avoid confusion. Following Bierkens et al. (2000), parameters may change in space, but not in time; constants are the only part of a model that does not change in space and time (i.e., scale-invariant); and all other model components may change in both space and time. Dynamic models can be implemented in mathematically explicit forms (e.g., differential or difference equations) or mathematically implicit forms (e.g., mathematical relation-based or rule-based simulation algorithms written in computer languages).

To illustrate different scaling methods clearly and precisely, let's assume that a dynamic model at a local scale s_1 is:

$$y(s_1) = f(v, \theta, i) \tag{2.5}$$

where $y(s_1)$, v, θ, and i are the output variables, state variables, parameters, and input variables at scale s_1, respectively.

Also, let's assume that a model can be developed at a broader scale s_2 as:

$$Y(s_2) = F(V, \Theta, I) \tag{2.6}$$

where $Y(s_2)$, V, Θ, and I are the output variables, state variables, parameters, and input variables at scale s_2, respectively.

Note that all the model arguments can be vectors. Then, transferring information from s_1 to s_2 usually involves one or more of the following transformations: $v \leftrightarrow V$, $\theta \leftrightarrow \Theta$, $i \leftrightarrow I$, and $f(v,\theta,i) \leftrightarrow F(V,\Theta,I)$, depending on how the model arguments and relationships at the two scales are linked (Blöschl and Sivapalan 1995, Wu 1999, Bierkens et al. 2000). Thus, the transfer of information between scales using dynamic models is done through rescaling or other kinds of alterations of inputs, parameters, state variables, and model conceptualizations.

Scaling with dynamic models typically consists of two major steps. For *upscaling*, the two steps are characterizing heterogeneity and aggregating information (Figure 2.1). First, characterizing spatial heterogeneity involves the classification and quantification of spatial patterns (e.g., the number, size, and spatial configuration of different types of patches in a landscape), which influence model inputs and parameters. Spatial heterogeneity can be characterized either in a spatially explicit way (i.e., in the form of maps) or in statistical terms (e.g., statistical moments, probability density functions, or pattern indices). In cases where data do not cover the entire study area, interpolation is often needed. Many spatial interpolation methods exist (Lam 1983, Goovaerts 1997), and geostatistical methods such as kriging are particularly useful. In all these cases, geographic information systems (GIS) and remote sensing have proven extremely useful (Quattrochi and Goodchild 1997, Marceau 1999, Hay et al. 2001).

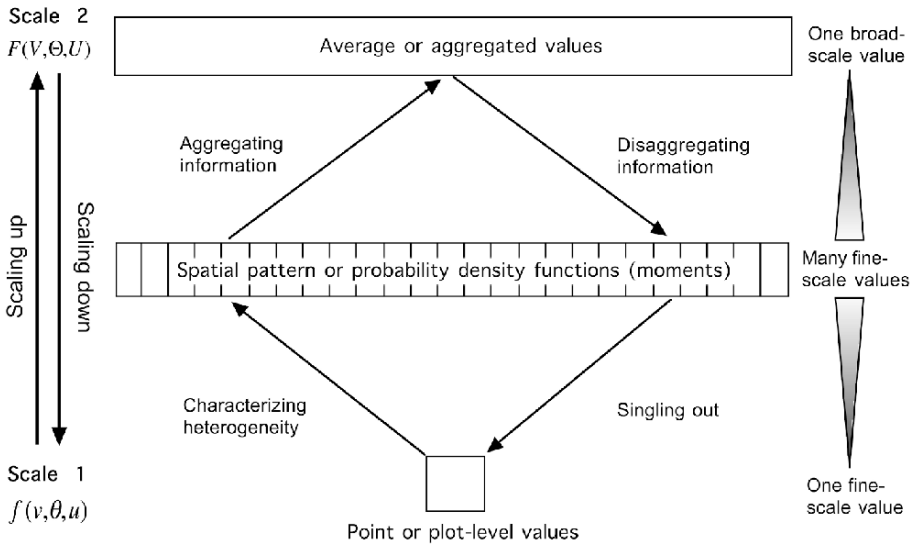


Figure 2.1. *Upscaling and downscaling as a two-step process when a dynamic model-based approach is used (modified from Blöschl and Sivapalan 1995).*

Second, aggregating information is to incorporate the quantitative description of spatial heterogeneity into local models to obtain predictions at a broader scale (or the target scale), be it a larger grain (coarse-graining) or a larger extent (extrapolation). Again, a variety of methods may be used in this step depending on what and how to aggregate, as discussed below. In general, aggregating state variables and inputs can be readily done following such first principles as the laws of mass and energy conservation, but aggregating model parameters can be quite challenging (Blöschl and Sivapalan 1995).

For *downscaling*, the two steps are disaggregating information and singling out (Figure 2.1). Disaggregating coarse-grained information is to derive the detailed pattern within a spatial domain (fine-graining) with auxiliary data. Because of the lack of within-grain (or within-pixel) information, this often requires stochastic or probabilistic methods. Singling out is simply to find the location of the disaggregated pattern that corresponds to the site of interest, which is usually a trivial matter. Note that scaling relations (power laws) derived from similarity-based methods are supposed to work for both scaling up and scaling down. In contrast, scaling with dynamic models employs both deterministic and stochastic formulations, which may differ significantly for upscaling versus downscaling. In the following, we discuss several dynamic model-based scaling methods in detail.

2.3.2 Upscaling Methods

The literature on upscaling methods is both abundant and confusing because of the diversity in disciplines and approaches as well as the idiosyncrasy in terminologies and traditions. From a landscape modeling perspective, King (1991) distinguished four extrapolation methods: extrapolation by lumping, direct extrapolation, extrapolation by expected value, and extrapolation by explicit integration. Blöschl and Sivapalan (1995), Becker and Braun (1999), and Bierkens et al. (2000) discussed a number of scaling methods in the context of hydrological modeling and soil physics. Based on these and other studies, we compare and contrast several model-oriented upscaling methods. We focus more on the second step of upscaling – aggregating information. It must be emphasized, however, that adequately characterizing spatial heterogeneity is a crucial and necessary first step for upscaling with dynamic models, because the accuracy in representing spatial pattern may not only affect the accuracy of scaling results (see Li and Wu, Chapter 3), but also the model conceptualization and simulation scheme.

2.3.2.1. Extrapolation by lumping

One of the simplest ways to transfer information between two scales is to obtain the target-scale estimate as the output of the local-scale model with the mean values of parameters and inputs averaged over the study area – a method called *extrapolation by lumping* (King 1991) or *simple averaging* (Bierkens et al. 2000). If the local model is deterministic, only one model run is needed. This method can be used for extrapolation with increasing extent as well as for coarse-graining with increasing

grain size (see Wu and Li, Chapter 1). In this case, the local model $f()$ is assumed to remain valid at the target scale. That is, $f() = F()$. The method also assumes that $f()$ is linear such that changes in output variables are proportional to changes in input variables and parameters. In addition, horizontal interactions and feedbacks are negligible or at steady state. Under these assumptions, the treatment of the spatial heterogeneity of the study system is extremely simplistic: spatial variability is all squeezed into the averages of model arguments. Mathematically, the lumping method can be expressed as:

$$Y(s_2) = f(< v >,< \theta >,< i >) \tag{2.7}$$

where <*variable*> denotes the average of a variable, and all other terms are the same as before.

Equation 2.7 holds when y and Y represent a density measure (e.g., population density, flux density, carbon storage per unit area). If y and Y represent some cumulative or additive variable (e.g., population size, total flux, biomass), then the target-scale estimate becomes the product of the local-scale average multiplied by the total area, i.e.:

$$Y(s_2) = A f(< v >,< \theta >,< i >) \tag{2.8}$$

where A is the size of the target grain size in the case of coarse-graining, or the spatial extent in the case of extrapolation.

Thus, extrapolation by lumping uses the same local model at the target scale, with highly aggregated values of parameters and inputs. Simply put, lumping is about averaging over space. If grain elements are of the same size, arithmetic averaging is usually used. But if grain elements are of different sizes, area-weighted averaging should be used, instead. The local model is defined at the scale of a grain element or patch in the case of increasing extent, and at the size of the smaller grain element whose aggregates form the larger grain element in the case of coarse-graining. As a consequence of the oversimplifying assumptions, extrapolation by lumping is expected to produce large scaling errors when the model is nonlinear, when the local model formulation is no longer applicable at the target scale, or when horizontal interactions between grain elements (or patches) are strong and asymmetric (King 1991, Bierkens et al. 2000). For example, if the target scale is a landscape consisting of a number of different interacting ecosystems, the simple lumping method is more than likely to fail.

2.3.2.2 Extrapolation by effective parameters

Similar to the simple lumping method, *extrapolation by effective parameters* assumes that the local model applies to the target scale such that upscaling can be done by manipulating its parameters and inputs. However, instead of simply averaging parameters and inputs over space, the method uses "effective" (also called "equivalent" or "representative") parameters and inputs to produce the target scale

estimates (L'homme et al. 1994, Blöschl and Sivapalan 1995, Bierkens et al. 2000). That is, the estimated value of a variable at the target scale, $Y(s_2)$, is the output of the local model with a set of effective parameters and inputs:

$$Y(s_2) = f(\Theta_r, I_r) \tag{2.9}$$

where Θ_r and I_r are the effective or representative parameters and inputs.

The effective parameter approach has been widely used in soil physics, hydrology, and micrometeorology (L'homme et al. 1996, Bierkens et al. 2000). A prototypical example of using this method is to find the effective hydraulic conductivity for models of groundwater or soil water dynamics (Blöschl and Sivapalan 1995). For uniform steady saturated flow through a soil block that is made up of smaller blocks of different hydraulic conductivities, the effective conductivity equals the arithmetic mean of the small-block conductivities when they are arranged in parallel, and the harmonic mean when the small blocks of soil are arranged in series. For unsaturated flow, infiltration, and overland flow, such general effective conductivity does not exists, and a number of factors other than the porous medium itself may affect hydraulic conductivity, even though the geometric mean is found to be "effective" in some situations. Micrometeorologists have long used the "flux matching" technique in modeling fluxes over heterogeneous landscapes (L'homme et al. 1996, Wood 1998). That is, to upscale a certain surface flux over a patchy geographic area, one assumes that the plot-scale model, $f(\theta(x), i(x,t))$, still holds at the landscape scale, and then seeks the representative parameters $\Theta(x)$ and inputs $I(x,t)$ that produce the same estimated landscape-scale flux as the summation of fluxes predicted by the local model using heterogeneous parameters and inputs, i.e.,

$$F(\Theta_r, I_r) = f(\Theta_r, I_r) = \int_A f[\theta(x), i(x,t)]dA(x) \tag{2.10}$$

where A is the area of a region over which the integration is performed.

Thus, extrapolation by effective parameters essentially is to run "micro-scale equations" using "macro-scale parameters" whose values are assigned to points within the study area to assure that the "uniform parameter field" produces the same model output as the "heterogeneous parameter field" (Blöschl and Sivapalan 1995). The key to this method is to successfully derive a set of representative parameters and inputs for scale s_2 from parameters and inputs at scale s_1 through synoptic descriptions of the fine-scale heterogeneous patterns. This is relatively easy for linear models where the representative parameters and inputs can be found by simply averaging over space. In this case, it becomes the simple lumping method again. For nonlinear models, however, finding effective parameters can be a difficult task. The values of representative parameters and inputs are determined by several factors: the values of parameters and inputs at scale s_1, the detailed formulation of the local model, and the difference between the two scales as measured by their ratio, s_2/s_1 (Bierkens et al. 2000). Effective parameters are rarely unique, and

fine-scale parameters used to estimate the target-scale effective parameters are themselves mostly approximations (Bierkens et al. 2000). A number of methods have been used for deriving representative parameters and inputs, including analytical approaches, Monte Carlo simulations, inverse modeling, and direct measurements (Blöschl and Sivapalan 1995, Dolman and Blyth 1997, Bierkens 2000). In general, the accuracy of extrapolation by effective parameters is dependent upon whether the local model is valid at the target scale and whether representative parameters and inputs can be found that adequately account for the spatial heterogeneity of the system under study.

2.3.2.3 Direct extrapolation

Different from the lumping and effective parameter methods, *direct extrapolation* (*sensu* King 1991) does not assume that the local model applies at the target scale. Instead of averaging parameters and inputs before running the model as in the lumping method, direct extrapolation obtains the target-scale estimates by first running the local model for each grain element (or patch) of the system with geospatially corresponding parameters and inputs, and then averaging (or summing up) the model outputs for all the grain elements. Mathematically, the relationship between the two scales can be described as:

$$Y(s_2) = \; < y(s_1) > \; = \; <f(v,\theta,i) > \tag{2.11}$$

$$\text{or,} \quad Y(s_2) = \sum_{k=1}^{N} y(s_1)_k = \sum_{k=1}^{N} f(v,\theta,i)_k \tag{2.12}$$

where N is the total number of areal units at scale s_1. Again, Equation 2.11 is applicable when y and Y represent a density measure, while Equation 2.12 is used if y and Y represent an additive variable.

In direct extrapolation, therefore, spatial heterogeneity is treated explicitly in terms of model parameters and inputs. This spatial explicitness can be retained in the target-scale estimates as well because the model outputs can also be presented in the form of maps in the case of coarse-graining (see Wu and Li, Chapter 1). More importantly, if the local model is nonlinear, running it first with spatial data and then averaging the outputs can reduce the errors due to model nonlinearity that are associated with the lumping method. If the local model is linear (rarely true in reality), however, direct extrapolation should produce the same results as the simple lumping method, but with higher demands for data preparation and computation.

Direct extrapolation is conceptually intuitive and technically straightforward, and it has been widely used in ecology, hydrology, and soil science. Examples include numerous spatially explicit or spatially distributed ecosystem and landscape models. Nevertheless, because it requires running the local model at all grain elements, direct extrapolation may suffer from excessive computational demand and redundancy when the total number of grain elements is great. This can be a real problem even with the most advanced computing facility because the computational demand for

such simulations increases exponentially with the number of grain elements. This problem can be alleviated by running the local model for each patch type (the aggregate of grain elements of the same type) instead of each grain element or individual patch because direct extrapolation is spatially "*explicit*," but not spatially "*interactive*." In other words, grain elements or patches are treated individually without considering the influences among them (see Peters et al., Chapter 7). The scaling results in the two cases should be similar as long as the variability of inputs and parameters within each patch type is insignificant.

Moreover, it is important to note that the assumption behind direct extrapolation that horizontal interactions and feedbacks are negligible or at steady state may not be valid in many situations, especially when lateral hydrological flows and exchanges of energy, materials and biological organisms in landscapes are significant and asymmetric. In other words, only one of the two components of spatial pattern, compositional heterogeneity (the diversity and relative abundance of patches), is recognized in this approach, leaving configurational heterogeneity (the geometric and spatial arrangement of patches) and its functional consequences unaccounted. Another implicit assumption is that there are no new patterns and processes emerging as the spatial extent increases. This can be a severe problem when direct extrapolation is practiced over a broad range of scale. Finally, spatial data on many processes often are not available for all the grain elements, resulting in obstacles for the effective application of this method.

2.3.2.4 Extrapolation by expected value

Similar to direct extrapolation, *extrapolation by expected value* (*sensu* King 1991) does not require the local-scale model to apply at the target scale. The expected value method treats model arguments as random variables, quantifies spatial heterogeneity by the joint probability distribution of model arguments, and uses the expected value of the local model outputs (also random variables) as the target-scale estimate (King 1991), i.e.,

$$Y(s_2) = E[y(s_1)] = E[f(v,\theta,i)] \tag{2.13}$$

$$\text{or,} \quad Y(s_2) = A\,E[y(s_1)] = A\,E[f(v,\theta,i)] \tag{2.14}$$

where $E[]$ is the mathematical expectation operator.

For discrete variables, the mathematical expectation of model outputs can be expressed as summations, i.e.,

$$E[y(s_1)] = \sum_{j=1}^{q} \sum_{k=1}^{r} \sum_{l=1}^{s} f(v_j,\theta_k,i_l)\ \phi(v_j,\theta_k,i_l) \tag{2.15}$$

where $\phi()$ is the joint probability distribution functions of the model arguments that describe the spatial heterogeneity of the model domain. For continuous variables, integrals are used in the places of summations.

Apparently, the most critical step in upscaling with the expected value method is to compute the mathematical expectation of the outputs of the local model, which requires an accurate description of the fine-scale spatial heterogeneity or variability. If the joint probability density functions, $\phi()$, can be estimated explicitly, the expected value of model outputs can be derived directly using analytical or numerical techniques (King 1991). However, in many cases the explicit forms of the joint probability density functions cannot be found, and thus a sampling approach (e.g., Monte Carlo simulation) may be used to compute the expected value of the outputs of the local model. While direct extrapolation requires a full representation of the heterogeneity in a spatially explicit manner, the expected value method defines the fine-scale spatial heterogeneity only in statistical terms. However, both methods share two key issues: accurately describing spatial heterogeneity in terms of model arguments at the fine scale, and properly aggregating outputs from the local model to derive the estimate at the broader scale.

Because the local model does not have to be applied to the target scale, extrapolation by expected value can be a rather general upscaling approach. It also overcomes some of the problems encountered in the previous methods: for example, the problem of oversimplifying spatial heterogeneity in the lumping method, difficulties in deriving effective parameters, and excessive data and computational demands with direct extrapolation when the target scale is too broad in extent. The method of expected value is quite amenable to uncertainty analysis (Li and Wu, Chapter 3) and, particularly, the results of Monte Carlo simulation may be used to estimate an unbiased confidence interval for the extrapolated value at the target scale (King 1991). This approach has been applied to a number of important ecological phenomena, including extrapolating information on primary production, trace gas fluxes, and other biophysical properties from local plots to landscapes or regions (e.g., King 1991, King et al. 1991).

A major potential problem with the expected value method is that, as in direct extrapolation, neither the geometric attributes and spatial arrangement of patches nor the interactions and feedbacks among grain elements (or patches) are explicitly considered. Therefore, scaling errors with this method are expected to increase when lateral flows and feedbacks are strong and significantly asymmetric or far from a steady state. Also, as with direct extrapolation, this method does not account for any new patterns and processes that emerge with increasing extent. In real landscapes, however, new processes and controls do emerge at progressively broader scales. If these "new" attributes have significant nonlinear effects, a method completely ignoring them is not only theoretically improper but practically inaccurate as well.

2.3.2.5 Explicit integration

If the spatial variations of all the arguments of the local-scale model can be adequately and explicitly represented as functions of space (x, y) in closed form, and

if the indefinite integral of the local model with respect to space exists (and can be obtained), upscaling between two scales can be accomplished by directly integrating the local-scale model. This method was termed *explicit integration* (King 1991), i.e.,

$$Y(s_2) = \iint_R f\left(v(x, y), \theta(x, y), i(x, y)\right) dA \tag{2.16}$$

where x and y are spatial coordinates, and A is the area of the region R over which the integration is performed.

The local model only needs to be evaluated once, and the prediction can be made precisely at any spatial scale within the defined region, R. In contrast with the lumping, effective parameters, direct, and expected value methods discussed earlier, the structure of the local-scale model now changes as a function of space during extrapolation by explicit integration. Explicit integration is elegant, efficient, and accurate when all of its requirements are met. However, because of the prevalence of nonlinear relationships in ecological models and complex spatial structures of model arguments, the applicability of explicit integration as an upscaling method is rather limited in practice. First of all, it is difficult or even impossible to represent spatial heterogeneity with closed-form mathematical functions of model arguments with acceptable accuracy. Second, even if this can be done, finding the indefinite integrals of nonlinear models, in general, is a formidable task. If the closed-form indefinite integrals cannot be found, approximating the double definite integral by numerical methods is equivalent to either direct extrapolation or extrapolation by expected value (King 1991).

2.3.2.6 Spatially interactive modeling

Not only do many processes in landscapes vary in their characteristics from one place to another, but also they often interact in space to generate feedbacks and emergent properties (Raupach et al. 1999, Wu 1999, Peterson 2000). Such examples include population dynamics in patchy environment, hydrological and biogeochemical dynamics in complex landscapes, and land-water and land-atmosphere interactions on various scales. To adequately understand and predict such phenomena across scales, models must explicitly consider the horizontal interactions of the processes under study.

Spatially interactive modeling integrates the two aspects of spatial scaling – characterizing heterogeneity and aggregating information between scales – into the dynamic models themselves. In this case, local-scale models or submodels are usually embedded in the larger-scale model (e.g., a metapopulation model consisting of many interacting local population models or a landscape model composed of multiple ecosystem models). Spatially interactive modeling deals explicitly with not only spatial variations in model arguments, but also the interactions among grain elements or patches. In other words, the values of the arguments of a local-scale model in one grain element or patch not only differ from, but also are functions of, the attributes of other (often neighboring) patches and the landscape matrix. Spatially interactive

modeling is able to incorporate feedbacks, time delays, and new features on larger scales. Different types of models emphasize different aspects of between-patch interactions (e.g., dispersal and species interactions in population and community models; hydrological and material exchanges in ecosystem models; spreading disturbances like fires and pests in many landscape models). In this case, the model conceptualization at the local scale (s_1) may be described as follows:

$$y(s_1) = f\left(v_k, \theta_k, i_k \right) + \delta p_k + \delta m_k \tag{2.17}$$

where $y(s_1)$ is the model output at the local scale s_1, v_k is a state variable for patch k, θ_k and i_k are parameters and input variables for the same patch k, δp_k is the net exchange between patch k and all other patches, δm_k is the net exchange between patch k and the landscape matrix, and $f()$ defines the local model in terms of within-patch dynamics. Then, the prediction at the target scale at any point of time t, $Y(s_2,t)$, can be described as:

$$Y(s_2,t) = \sum_{k=1}^{N} y_k(s_1,t) = \sum_{k=1}^{N} \left(f\left(v_k, \theta_k, i_k \right) + \delta p_k + \delta m_k \right) \tag{2.18}$$

where N is the total number of grain elements or patches at scale s_1.

For ecological processes in a shifting landscape mosaic, the dynamics of both landscape pattern and ecological processes, as well as their interactions, have to be modeled explicitly. For example, the overall population density in a dynamic landscape may be predicted by coupling the patch dynamics of the biological population of interest and the changing habitat patches (Levin and Paine 1974, Levin 1976):

$$\begin{aligned} \bar{n}_j(t) = \frac{1}{A} \{ [A - \int_0^\infty \int_0^\infty \xi \rho(t,\alpha,\xi)\, d\alpha d\xi]\, \varphi_j^O(t) \\ + \int_0^\infty \int_0^\infty \xi \rho(t,\alpha,\xi)\, \varphi_j(t,\alpha,\xi)\, d\alpha d\xi \} \end{aligned} \tag{2.19}$$

where $\bar{n}_j(t)$ is the overall population density of species j in the landscape, $\rho(t,\alpha,\xi)$ is the probability density function describing the frequency distribution of patches of age α and size ξ at time t, $\varphi_j(t,\alpha,\xi)$ is the population density of species j within a patch of age α and size ξ at time t, $\varphi_j^0(t)$ is the population density of the same species in the non-patch area, and A is the total area of the landscape.

Specifying such models in closed forms and solving them analytically are difficult or impossible for real landscapes. Spatially explicit simulation modeling with Monte Carlo integration provides a general approach to dealing with such complex patch dynamics problems (Wu and Levin 1997).

In the past two decades, a great number of spatially interactive models of different kinds have been developed in ecology and earth sciences, examining such topics as metapopulation dynamics (Hanski 1999), landscape dynamics (e.g., Li et al. 1993, Wu and Levin 1997), hydrological and biogeochemical processes (e.g., Tenhunen and Kabat 1999, Beven 2000), and socioeconomic processes (Schweitzer 1997). Many modeling approaches have been used, including various grid-based models, cellular automata, and individual-based and agent-based models. Most, if not all, spatially interactive models are multiscaled or hierarchical. For scaling across broader spatial scales such as landscapes and regions, GIS and remote sensing techniques have increasingly been used in such models. The main sources of error in upscaling with spatially interactive modeling vary from case to case, but are due primarily to the characterization of spatial heterogeneity, formulation of spatial interactions, interface of multiple scales, and computational algorithms. In addition, such models can be quite demanding in data and computational requirements.

2.3.2.7 Extrapolation along a scaling ladder

The upscaling methods discussed above are in principle only applicable to situations where there are neither significant asymmetric between-patch interactions nor scale-dependent or emergent patterns and processes. These are typically "short-range" scaling methods because the assumptions behind them are less likely to be satisfied over a broad range of scale. When scaling involves multiple scale domains or levels of organization, new patterns and processes at different scales as well as vertical linkages need to be taken into account. How can these "short-range" upscaling methods be used for transferring information over a "long range" of scales that have multiple scaling domains or organizational levels?

To address this question, Wu (1999) proposed a "scaling ladder" approach based on the *hierarchical patch dynamics* (HPD) paradigm, which integrates hierarchy theory and patch dynamics (Wu and Loucks 1995). The *scaling ladder approach* facilitates the understanding and scaling of patterns and processes in different kinds of heterogeneous landscapes (Hay et al. 2001, Poole 2002, Wu and David 2002, Burnett and Blaschke 2003). The first step in this approach is to construct a spatially nested hierarchical system with distinctive scaling domains or levels of organization. Top-down (partitioning) or bottom-up (aggregation) schemes can be used in this step (Wu 1999). A top-down approach identifies the levels of a hierarchy by progressively partitioning the entire system downscale, whereas a bottom-up scheme involves successively aggregating or grouping similar entities upscale. *A priori* spatial hierarchies based on empirical observations and natural biophysical boundaries may be used as long as they are spatially nested and relevant to processes of interest. Such empirical hierarchies are commonly found in all fields of study, including the individual-population-community-biome or plot-ecosystem-landscape-region hierarchies (e.g., Urban et al. 1987, Jarvis 1995), soil-type hierarchies (Woodmansee 1990), hydrologic unit hierarchies (e.g., site-drainage-subwatershed-watershed-subbasin-basin-subregion-region; Griffith et al. 1999), and landscape and geomorphological hierarchies (e.g., Reynolds and Wu 1999).

It is important to realize that the appropriateness of a given hierarchy varies with the process under study and research questions to be addressed (Wu and Loucks 1995, Wu 1999, Omernik 2003). Quantitative methods, including landscape metrics, spatial statistics, and object-oriented approaches can be used for identifying patch hierarchies (e.g., O'Neill et al. 1991, Li and Wu 2004, Hay et al. 2001, 2003, Hall et al. 2004, Wu 2004). As emphasized in Wu and Loucks (1995) and Wu (1999), these spatial hierarchies should be constructed with consideration of both pattern and process, such that they are neither completely arbitrary, nor simply products of spatial analyses. Also, they are dynamic, not static, as indicated by the phrase "hierarchical patch dynamics." Once an appropriate patch hierarchy, or a scaling ladder, is established, the short-range scaling methods discussed above can be applied within each scale domain or between two adjacent hierarchical levels (Figure 2.2). This can be accomplished by changing grain, extent, or both.

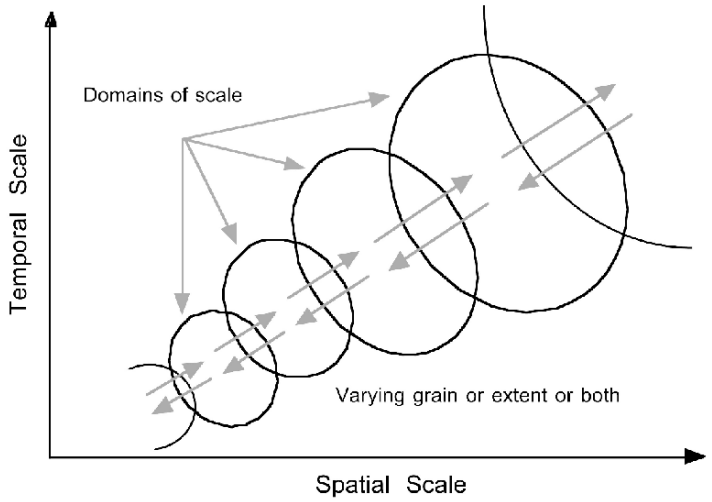


Figure 2.2. *Illustration of the scaling ladder approach in which scaling up (or down) is implemented by changing model grain size, extent, or both across successive domains of scale (redrawn from Wu 1999).*

2.3.3 Downscaling Methods

The central question of *downscaling* is: given the aggregated values, the probability distributions, or the functional relationships of variables at a particular scale, how can they be derived at a smaller scale (Bierkens et al. 2000)? In many applications, the objective of downscaling is, "given the average value over a certain domain, to derive the detailed pattern within that domain" (Blöschl and Sivapalan 1995). Downscaling methods usually share the assumption that the variable to be downscaled varies according to some function of space within the support units, and thus downscaling the variable is essentially to seek the parameters of this function such that the average value of the support unit or its probability distribution is

maintained (Bierkens et al. 2000). In principle, downscaling does not produce unique solutions because the values of a variable at scale s_1 may vary within a grain unit in an infinite number of ways without changing the average value at s_2. For example, for a grain unit at scale s_2 composed of 4 smaller grain units at scale s_1, there are an infinite number of combinations of the values that the four s_1 grain units may take to produce the average value of 1.

Bierkens et al. (2000) discussed three kinds of downscaling methods (1) deterministic downscaling, (2) conditional stochastic downscaling, and (3) unconditional stochastic downscaling. In deterministic downscaling, the average property at scale s_2 is known exactly, and the objective is to find a single deter-ministic function to describe the spatial or temporal variation of values at scale s_1, such that the average of disaggregated values matches the known average value at scale s_2. In conditional stochastic downscaling, the average property at scale s_2 is also known exactly, but the objective is to find a set of equally probable functions that can predict the disaggregated values at s_1 while maintaining the same known average value at s_2. Monte Carlo analysis can be used in this case. Unconditional stochastic downscaling occurs when only the probability density function of the average property at scale s_2 is known. The objective is to find a family of equally probable functions of the spatial or temporal variation at s_1 that produce the known probability density function of the average property at s_2. Blöschl and Sivapalan (1995) and Bierkens et al. (2000) provided several examples in the context of soil and hydrological sciences (e.g., disaggregating information on soil properties in a soil profile or over a geographic area, and downscaling hydrological time series or spatial pattern of rainfall). These three kinds of downscaling methods can employ either empirical functions or mechanistic models (Bierkens et al. 2000).

While there are various needs for downscaling over relatively fine scales, the current literature on downscaling is clearly dominated by climate studies on much broader scales. Because General Circulation Models (GCMs) operate at spatial resolutions (usually $>2^{\circ}$ in both latitude and longitude) that are too coarse for understanding the regional and local impacts of global climate change, there has been a great deal of research in climate downscaling in the past few decades. Specifically, climate downscaling refers to "a set of procedures by which we attempt to take information available at the relatively poor spatial resolution of the GCMs, and derive regional-scale data that can be used for ecosystems modeling, climate impact assessment, and other tasks that require higher resolution climate data" (Crane et al. 2002). Thus, the primary objective of climate downscaling is disaggregating (i.e., fine-graining) GCM outputs to produce patterns of surface climatic conditions (e.g., temperature, precipitation, wind velocity) at regional (and eventually local) scales on which most ecological and socioeconomic processes operate. Parallel to the two general approaches to scaling, the methods of downscaling are also commonly classified into two main approaches: empirically-based statistical downscaling and process model-based downscaling (Hewitson and Crane 1996, Crane et al. 2002).

The *empirically-based statistical downscaling approach* derives quantitative relationships between large-scale synoptic circulation features (e.g., upper level winds, geopotential heights, and sea level pressure) predicted by GCMs and regional climate conditions (e.g., temperature, precipitation, wind velocity) based on observations. These relations, in some form of $y = f(x)$, are often referred to as "transfer functions," which are obtained through multiple linear regression, artificial neural networks (ANNs), classification and regression trees (CART), and other statistical methods (Hewitson and Crane 1996, Wilby et al. 1998, Sailor et al. 2000, Crane et al. 2002). Note that, in contrast with the empirical methods in allometric and similarity analysis discussed earlier, empirical downscaling in climate studies rarely invokes similarity assumptions of any kind. The feasibility and the validity of the empirical approach hinge on the fundamental assumption that "stable empirical relationships can be established between atmospheric processes occurring at disparate temporal and/or spatial scales" (Wilby et al. 1998). Thus, such empirical or statistical relationships do not have the capacity to explain how circulation-related processes affect regional/local climate, and their predictive ability is undermined by the fact that transfer functions are often varying in time. Empirical techniques in climate downscaling have been widely used because they are operationally simpler and computationally much less demanding than the process modeling approach.

The *process-based downscaling approach*, also called *nested modeling*, embeds a higher-resolution regional climate model within a global GCM so that the coarse-grained predictions from GCM are dynamically translated into fine-grained outputs of the regional model. In most cases, the output of GCM from a large grid cell is used to provide boundary conditions for the regional climate model inside the grid cell in one-way nesting schemes. There are also two-way nesting schemes in which GCM and the embedded regional climate model run simultaneously and interact with each other across scales. Because the land surface characteristics may significantly affect local, regional, and even global climatic processes (Pielke and Avissar 1990, Raupach et al. 1999), nested modeling provides a necessary and promising approach to climate downscaling and to coupling geophysical and ecological processes across scales in general. However, because of the excessive computational demands and the lack of understanding of interface mechanisms of climatic processes at different scales, two-way nesting has not yet been commonly used in climate downscaling with GCMs.

2.4 DISCUSSION AND CONCLUSIONS

We have classified numerous scaling methods into two general approaches: similarity-based versus dynamic model-based. The first approach itself has a diversity of methods, including those relying on empirical, statistical methods and those based on first principles and analytical methods such as dimensional and similarity techniques. The dynamic model-based scaling approach, in contrast, puts more emphasis on the processes and mechanisms of interest, and employs a variety of methods for modeling (deterministic and stochastic), quantifying heterogeneity, and aggregating information across scales. Similarity-based scaling techniques also

have been used in dynamic model-based scaling methods. Thus, the two general approaches are not mutually exclusive; on the contrary, they should be considered complementary.

The similarity-based scaling approach is elegant and powerful when it is found applicable to the problem at hand. If the scaling relations are valid, using them for extrapolation is simple and bi-directional (both up and down). However, caution must be taken when underlying processes and mechanisms are inferred from such scaling relations. Brown et al. (2000) asserted that "most biological scaling relationships are manifestations of a single underlying scaling process, which appears to be based on quarter powers and to be unique to living things." However, because ecological systems are mixtures of physical, chemical, biological, and socioeconomic processes, it is unlikely that a single similarity criterion applies to all or most of them (Prothero 1986). The empirical determination of a power scaling relationship cannot be simply taken as *prima facie* evidence for similarity (i.e., objects of different sizes are similar in geometry and shape) because "pseudo-similarity" may also result in straight lines in a log-log regression plot (Prothero 1986). Thus, geometric similarity will for sure result in power laws of simple multiples of 1/3, but the reverse is not guaranteed (Niklas 1994). The same can be said about the relationship between power laws and other kinds of similarity. One needs to bear in mind that statistical equations merely express correlations, any of which can be spurious (Prothero 1986).

There are several reasons why dynamic-model scaling methods ought to be used in many situations. First of all, not all ecological patterns and processes adequately meet the criteria of complete similarity (e.g., geometric similarity) or incomplete similarity (e.g., self-similarity or self-affinity). Therefore, alternative methods are needed for translating information from one scale to another for those patterns and processes. Second, the goals of scaling may be not only to describe and predict across scales, but also to understand patterns and processes at multiple scales. However, most similarity-based methods are empirical, relying primarily on statistical techniques, and do not deal directly with dynamic processes. Empirical equations provide useful information on quantitative relations among variables, but tell us little beyond the quantitative relations themselves (Prothero 1986). On the other hand, analytical similarity-based methods either demand explicit mathematical expressions of processes under study or start with well-established first principles. These requirements can rarely be met for most complex ecological problems. Third, ecologists are often interested in quantifying how spatial heterogeneity interacts with ecological processes in their efforts to scale across space. This requires an approach that deals with space explicitly and processes directly. Apparently, similarity-based methods that rely on regression or differential equations are inadequate to achieve this objective.

Dynamic model-based scaling methods can be used to overcome these shortcomings because (1) they are not constrained by similarity assumptions, (2) they can explicitly consider dynamic processes and their interactions, and (3) they can couple patterns and processes in spatially explicit fashion. Spatial scaling with dynamic models has two critical steps: accurately defining and quantifying the fine-scale heterogeneity and properly aggregating or integrating this heterogeneity

through model arguments or outputs to derive broad-scale projections (King 1991, Blöschl and Sivapalan 1995, Wu 1999). Besides the methods discussed here, there are other dynamic modeling techniques also relevant to spatial upscaling. For example, the methods of model simplification or model aggregation (e.g., O'Neill and Rust 1979, Iwasa et al. 1989, Cale 1995) are useful for upscaling especially when they directly address the problem of spatial aggregation. Metamodeling – developing coarse-scale models based on fine-scale models – is directly relevant to upscaling (e.g., de Vries et al. 1998, Urban et al. 1999, Bierkens et al. 2000).

To develop a science of scaling, a pluralistic strategy is necessary. The pluralism should not only be reflected in the views and theories of scaling, but also need to be implemented in the methods and applications of transferring information across scales. Pluralism does not mean arbitrary division and diversification; rather, it provides a realistic basis for enlightened scaling. It would be nice if the systems of all kinds in the universe behaved like a sandpile, so that simple power laws could adequately describe "how nature works." Although some physical, ecological, and socioeconomic systems may indeed exhibit scale-invariant behavior within certain temporal and spatial scale domains (e.g., Chave and Levin 2003, Wu 2004), scaling is certainly more than just deriving power laws. In particular, the progress in ecological scaling depends on how well we can integrate the different scaling approaches and methods, and use them appropriately for the intended problems.

ACKNOWLEDGEMENTS

We would like to thank Geoffrey Hay, Fangliang He, and Simon Levin for their comments on an earlier version of the chapter. JW's research on scaling has been supported in part by grants from US EPA's Science to Achieve Results (STAR) Program (R827676-01-0) and US NSF (DEB 9714833, CAPLTER).

REFERENCES

Barenblatt, G. I. 1996. Scaling, Self-Similarity, and Intermediate Asymptotics. Cambridge University Press, Cambridge.

Becker, A., and P. Braun. 1999. Disaggregation, aggregation and spatial scaling in hydrological modelling. Journal of Hydrology 217:239-252.

Belyea, L. R., and J. Lancaster. 2002. Inferring landscape dynamics of bog pools from scaling relationships and spatial patterns. Journal of Ecology 90:223-234.

Berntson, G. M., and P. Stoll. 1997. Correcting for finite spatial scales of self-similarity when calculating the fractal dimensions of real-world structures. Proceedings of the Royal Society of London Series B-Biological Sciences 264:1531-1537.

Beven, K. 2000. On the future of distributed modelling in hydrology. Hydrological Processes 14: 3183-3184.

Bierkens, M. F. P., P. A. Finke, and P. de Willigen. 2000. Upscaling and Downscaling Methods for Environmental Research. Kluwer Academic Publishers, Dordrecht.

Binney, J. J., N. J. Dowrick, A. J. Fisher, and M. E. Newman. 1993. The Theory of Critical Phenomena: An Introduction to the Renormalization Group. Oxford Science, Oxford.

Blöschl, G., and M. Sivapalan. 1995. Scale issues in hydrological modelling: A review. Hydrological Processes 9:251-290.

Bokma, F. 2004. Evidence against universal metabolic allometry. Functional Ecology 18:184-187.

Brown, J. H., V. K. Gupta, B.-L. Li, B. T. Milne, C. Restrepo, and G. B. West. 2002. The fractal nature of nature: power laws, ecological complexity and biodiversity. Philosophical Transactions of the Royal Society (London B) 357:619-626.

Brown, J. H., G. B. West, and B. J. Enquist. 2000. Patterns and processes, causes and consequences. Pages 1-24 *in* J. H. Brown and G. B. West, editors. Scaling in Biology. Oxford University Press, New York.

Bull, W. B. 1975. Allometric change of landforms. Geological Society of America Bulletin 86:1489-1498.

Burnett, C., and T. Blaschke. 2003. A multi-scale segmentation/object relationship modeling methodology for landscape analysis. Ecological Modelling 168:233-249.

Calder, W. A. 1983. Ecological scaling: mammals and birds. Annual Review of Ecology and Systematics 14:213-230.

Cale, W. G. 1995. Model Aggregation – Ecological perspectives. Pages 230-241 *in* B. C. Patten, S. E. Jorgensen, and S. I. Auerbach, editors. Complex Ecology. Prentice-Hall, Englewood Cliffs.

Carbone, C., and J. L. Gittleman. 2002. A common rule for the scaling of carnivore density. Science 295:2273-2276.

Chave, J., and S. A. Levin. 2003. Scale and scaling in ecological and economic systems. Environmental and Resource Economics 26:527-557.

Church, M., and D. M. Mark. 1980. On size and scale in geomorphology. Progress in Physical Geography 4:342-390.

Crane, R. G., B. Yarnal, E. J. Barron, and B. Hewitson. 2002. Scale interactions and regional climate: examples from the Susquehanna River Basin. Human and Ecological Risk Assessment 8:147-158.

Crawley, M. J., and J. E. Harral. 2001. Scale dependence in plant biodiversity. Science 291:864-868.

Cyr, H., and S. C. Walker. 2004. An illusion of mechanistic understanding. Ecology 85:1802-1804.

de Vries, W., J. Kros, C. van der Salm, J. E. Groenenberg, and G. J. Reinds. 1998. The use of upscaling procedures in the application of soil acidification models at different spatial scales. Nutrient Cycling in Agroecosystems 50:223-236.

Dolman, A. J., and E. A. Blyth. 1997. Patch scale aggregation of heterogeneous land surface cover for mesoscale meteorological model. Journal of Hydrology 190:252-268.

Enquist, B. J., J. H. Brown, and G. B. West. 1998. Allometric scaling of plant energetics and population density. Nature 395:163-165.

Gardner, R. H., B. T. Milne, M. G. Turner, and R. V. O'Neill. 1987. Neutral models for the analysis of broad-scale landscape pattern. Landscape Ecology 1:19-28.

Goovaerts, P. 1997. Geostatistics for Natural Resources Evaluation. Oxford University Press, New York.

Gould, S. J. 1966. Allometry and size in ontogeny and phylogeny. Biological Review 41:587-640.

Griffith, G. E., J. M. Omernik, and A. J. Woods. 1999. Ecoregions, watersheds, basins, and HUCs: how state and federal agencies frame water quality. Journal of Soil and Water Conservation 54:666-677.

Gunther, B. 1975. Dimensional analysis and theory of biological similarity. Physiological Reviews 55:659-699.

Hall, O., G. J. Hay, A. Bouchard, and D. J. Marceau. 2004. Detecting dominant landscape objects through multiple scales: an integration of object-specific methods and watershed segmentation. Landscape Ecology 19:59-76.

Hanski, I., editor. 1999. Metapopulation Ecology. Oxford University Press, New York.

Harte, J., A. Kinzig, and J. Green. 1999. Self-similarity in the distribution and abundance of species. Science 284:334-336.

Hastings, H. M., and G. Sugihara. 1993. Fractals: A User's Guide for the Natural Sciences. Oxford University Press, Oxford.

Haverkamp, R., J.-Y. Parlange, R. Cuenca, P. J. Ross, and T. S. Steenhuis. 1998. Scale analyses for land-surface hydrology. Pages 190-223 *in* G. Sposito, editor. Scale Dependence and Scale Invariance in Hydrology. Cambridge University Press, Cambridge.

Hay, G., D. J. Marceau, P. Dubé, and A. Bouchard. 2001. A multiscale framework for landscape analysis: object-specific analysis and upscaling. Landscape Ecology 16:471-490.

Hay, G. J., T. Blaschke, D. J. Marceau, and A. Bouchard. 2003. A comparison of three image-object methods for the multiscale analysis of landscape structure. ISPRS Journal of Photogrammetry and Remote Sensing 57:327-345.

Hewitson, B. C., and R. G. Crane. 1996. Climate downscaling: techniques and application. Climate Research 7:85-95.

Hood, W. G. 2002. Application of landscape allometry to restoration of tidal channels. Restoration Ecology 10:213-222.
Iwasa, Y., S. A. Levin, and V. Andreasen. 1989. Aggregation in model ecosystems: II. approximate aggregation. IMA Journal of Mathematics Applied in Medicine & Biology 6:1-23.
Jarvis, P. G. 1995. Scaling processes and problems. Plant, Cell and Environment 18:1079-1089.
King, A. W. 1991. Translating models across scales in the landscape. Pages 479-517 *in* M. G. Turner and R. H. Gardner, editors. Quantitative Methods in Landscape Ecology. Springer-Verlag, New York.
King, A. W., A. R. Johnson, and R. V. O'Neill. 1991. Transmutation and functional representation of heterogeneous landscapes. Landscape Ecology 5:239-253.
Kozlowski, J., and M. Konarzewski. 2004. Is West, Brown and Enquist's model of allometric scaling mathematically correct and biologically relevant? Functional Ecology 18:283-289.
Kunin, W. E. 1998. Extrapolating species abundance across spatial scales. Science 281:1513-1515.
L'homme, J. P., A. Chehbouni, and B. A. Monteny. 1996. Canopy to region scale translation of surface fluxes. Pages 161-182 *in* J. B. Stewart, E. T. Engman, R. A. Feddes, and Y. Kerr, editors. Boundary-Layer Meteorology. Wiley, Chichester.
LaBarbera, M. 1989. Analyzing body size as a factor in ecology and evolution. Annual Review of Ecology and Systematics 20:97-117.
Lam, N. S. 1983. Spatial interpolation methods: a review. The American Cartographer 10:129-149.
Levin, S. A. 1976. Population dynamic models in heterogeneous environments. Annual Review of Ecology and Systematics 7:287-310.
Levin, S. A. 1992. The problem of pattern and scale in ecology. Ecology 73:1943-1967.
Levin, S. A., and S. W. Pacala. 1997. Theories of simplification and sclaing of spatially distributed processes. Pages 271-295 *in* D. Tilman and P. Kareiva, editors. Spatial Ecology. Princeton University Press, Princeton.
Levin, S. A., and R. T. Paine. 1974. Disturbance, patch formation and community structure. Proceedings of the National Academy of Sciences 71:2744-2747.
Li, H., and J. Wu. 2004. Use and misuse of landscape indices. Landscape Ecology 19:389-399.
Li, H., J. F. Franklin, F. J. Swanson, and T. A. Spies. 1993. Developing alternative forest cutting patterns: a simulation approach. Landscape Ecology 8:63-75.
Lomolino, M. V. 2000. Ecology's most general, yet protean pattern: the species-area relationship. Journal of Biogeography 27:17-26.
Lonsdale, W. M. 1990. The self-thinning rule: dead or alive? Ecology 71:1373-1388.
Magnani, F. 1999. Plant energetics and population density. Nature 398:572.
Mandelbrot, B. B. 1982. The Fractal Geometry of Nature. W. H. Freeman and Company, New York.
Mandelbrot, B. B. 1999. Multifractals and 1/f Noise: Wild Self-Affinity in Physics (1963-1976). Springer-Verlag, New York.
Marceau, D. J. 1999. The scale issue in social and natural sciences. Canadian Journal of Remote Sensing 25:347-356.
Miller, E. E., and R. D. Miller. 1956. Physical theory for capillary flow phenomena. Journal of Applied Physics 27:324-332.
Milne, B. T. 1991. Heterogeneity as a multiscale characteristic of landscapes. Pages 69-84 *in* J. Kolasa and S. T. A. Pickett, editors. Ecological Heterogeneity. Springer-Verlag, New York.
Milne, B. T. 1998. Motivation and benefits of complex systems approaches in ecology. Ecosystems 1:449-456.
Niklas, K. J. 1994. Plant Allometry: The Scaling of Form and Process. University of Chicago Press, Chicago.
Nikora, V. I., C. P. Pearson, and U. Shankar. 1999. Scaling properties in landscape patterns: New Zealand experience. Landscape Ecology 14:17-33.
O'Neill, R. V. 1979. Natural variability as a source of error in model predictions. Pages 23-32 *in* G. S. Innis and R. V. O'Neill, editors. Systems Analysis of Ecosystems. International Co-operative Publishing House, Fairland, MD.
O'Neill, R. V., and B. Rust. 1979. Aggregation error in ecological models. Ecological Modelling 7: 91-105.
O'Neill, R. V., R. H. Gardner, B. T. Milne, M. G. Turner, and B. Jackson. 1991. Heterogeneity and Spatial Hierarchies. Pages 85-96 *in* J. Kolasa and S. T. A. Pickett, editors. Ecological Heterogeneity. Springer-Verlag, New York.

Omernik, J. M. 2003. The misuse of hydrologic unit maps for extrapolation, reporting, and ecosystem management. Journal of American Water Resources Association 39:563-573.

Peters, R. H. 1983. The Ecological Implications of Body Size. Cambridge University Press, Cambridge.

Peterson, G. D. 2000. Scaling ecological dynamics: Self-organization, hierarchical structure, and ecological resilience. Climatic Change 44:291-309.

Phillips, J. D. 1999. Earth Surface Systems: Complexity, Order and Scale. Blackwell Publishers Ltd., Malden, Massachusetts.

Pielke, R. A., and R. Avissar. 1990. Influence of landscape structure on local and regional climate. Landscape Ecology 4:133-155.

Poole, G. C. 2002. Fluvial landscape ecology: addressing uniqueness within the river discontinuum. Freshwater Biology 47:641-660.

Prothero, J. 1986. Methodological aspects of scaling in biology. Journal of Theoretical Biology 118:259-286.

Quattrochi, D. A., and M. F. Goodchild, editors. 1997. Scale in Remote Sensing and GIS. CRC Lewis Publishers, Boca Raton.

Raupach, M. R., D. D. Baldocchi, H.-J. Bolle, L. Dümenil, W. Eugster, F. X. Meixner, J. A. Olejnik, R. A. Pielke, J. D. Tenhunen, and R. Valentin. 1999. How is the atmospheric coupling of land surfaces affected by topography, complexity in landscape patterning, and the vegetation mosaic? Pages 177-196 *in* J. D. Tenhunen and P. Kabat, editors. Integrating Hydrology, Ecosystem Dynamics, and Biogeochemistry in Complex Landscapes. Wiley, Chichester.

Reynolds, J. F., and J. Wu. 1999. Do landscape structural and functional units exist? Pages 273-296 *in* J. D. Tenhunen and P. Kabat, editors. Integrating Hydrology, Ecosystem Dynamics, and Biogeochemistry in Complex Landscapes. Wiley, Chichester.

Rodriguez-Iturbe, I., and A. Rinaldo. 1997. Fractal River Networks. Cambridge University Press, New York.

Sailor, D., T. Hu, X. Li, and J. N. Rosen. 2000. A neural network approach to local downscaling of GCM output for assessing wind power implications of climate change. Renewable Energy 19:359-378.

Scheurer, D. L., D. C. Schneider, and L. P. Sanford. 2001. Scaling issues in marine experimental ecosystems. Pages 330-360 *in* R. H. Gardner, W. M. Kemp, V. S. Kennedy, and J. E. Petersen, editors. Scaling Relations in Experimental Ecology. Columbia University Press, New York.

Schmid, P. E., M. Tokeshi, and J. M. Schmid-Araya. 2000. Relation between population density and body size in stream communities. Science 289:1557-1560.

Schmidt-Nielsen, K. 1984. Scaling: Why Is Animal Size So Important? Cambridge University Press, Cambridge.

Schneider, D. C. 2001a. Spatial allometry. Pages 113-153 *in* R. H. Gardner, W. M. Kemp, V. S. Kennedy, and J. E. Petersen, editors. Scaling Relations in Experimental Ecology. Columbia University Press, New York.

Schneider, D. C. 2001b. The rise of the concept of scale in ecology. Bioscience 51:545-553.

Schoener, T. W., D. A. Spiller, and J. B. Losos. 2001. Natural restoration of the species-area relation for a lizard after a hurricane. Science 294:1525-1528.

Schroeder, M. 1991. Fractals, Chaos, Power Laws: Minutes from an Infinite Paradise. W.H. Freeman and Company, New York.

Schweitzer, F., editor. 1997. Self-Organization of Complex Structures: From Individual to Collective Dynamics. Gordon and Breach Science Publishers, Amsterdam.

Shmida, A., and M. Wilson. 1985. Biological determinants of species diversity. Journal of Biogeography 12:1-20.

Sposito, G., editor. 1998. Scale Dependence and Scale Invariance in Hydrology. Cambridge University Press, Cambridge.

Stanley, and H. E. P. Meakin. 1988. Multifractal phenomena in physics and chemistry. Nature 335: 405-409.

Tenhunen, J. D., and P. Kabat, editors. 1999. Integrating Hydrology, Ecosystem Dynamics, and Biogeochemistry in Complex Landscapes. Wiley, Chichester.

Turcotte, D. L. 1995. Scaling in geology: landforms and earthquakes. Proceedings of the National Academy of Science 92:6697-6704.

Turner, M. G., V. H. Dale, and R. H. Gardner. 1989. Predicting across scales: theory development and testing. Landscape Ecology 3:245-252.

Urban, D. L., M. F. Acevedo, and S. L. Garman. 1999. Scaling fine-scale processes to large-scale patterns using models derived from models: meta-models. Pages 70-98 *in* D. J. Mladenoff and W. L. Baker, editors. Spatial Modeling of Forest Landscape Change: Approaches and Applications. Cambridge University Press, Cambridge, UK.

Urban, D. L., R. V. O'Neill, and H. H. Shugart. 1987. Landscape ecology: a hierarchical perspective can help scientists understand spatial patterns. Bioscience 37:119-127.

Weller, D. E. 1987. A reevaluation of the -3/2 power rule of plant self-thinning. Ecological Monographs 57:23-43.

West, G. B., J. H. Brown, and B. J. Enquist. 1997. A general model for the origin of allometric scaling laws in biology. Science 276:122-126.

Whittaker, R. J., K. J. Willis, and R. Field. 2001. Scale and species richness: towards a general, hierarchical theory of species diversity. Journal of Biogeography 28:453-470.

Wilby, R. L., T. M. L. Wigley, D. Conway, P. D. Jones, B. C. Hewitson, J. Main, and D. S. Wilks. 1998. Statistical downscaling of general circulation model output: a comparison of methods. Water Resources Research 34:2995-3008.

Wilson, K. G. 1975. Renormalization group methods. Advances in Mathematics 16:170-186.

Woldenberg, M. J. 1969. Spatial order in fluvial systems: Horton's laws derived from mixed hexagonal hierarchies of drainage basin areas. Geological Society of America Bulletin 80:97-112.

Wood, E. F. 1998. Scale analyses for land-surface hydrology. Pages 1-29 *in* G. Sposito, editor. Scale Dependence and Scale Invariance in Hydrology. Cambridge University Press, Cambridge.

Woodmansee, R. G. 1990. Biogeochemical cycles and ecological hierarchies. Pages 57-71 *in* I. S. Zonneveld and R. T. T. Forman, editors. Changing Landscapes: An Ecological Perspective. Springer-Verlag, New York.

Wu, J. 1990. Modelling the energy exchange processes between plant communities and environment. Ecological Modelling 51:233-250.

Wu, J. 1999. Hierarchy and scaling: extrapolating information along a scaling ladder. Canadian Journal of Remote Sensing 25:367-380.

Wu, J. 2004. Effects of changing scale on landscape pattern analysis: scaling relations. Landscape Ecology 19:125-138.

Wu, J., and J. L. David. 2002. A spatially explicit hierarchical approach to modeling complex ecological systems: theory and applications. Ecological Modelling 153:7-26.

Wu, J., and R. Hobbs. 2002. Key issues and research priorities in landscape ecology: an idiosyncratic synthesis. Landscape Ecology 17:355-365.

Wu, J., and S. A. Levin. 1997. A patch-based spatial modeling approach: conceptual framework and simulation scheme. Ecological Modelling 101:325-346.

Wu, J., and O. L. Loucks. 1995. From balance-of-nature to hierarchical patch dynamics: a paradigm shift in ecology. Quarterly Review of Biology 70:439-466.

Wu, J., D. Jelinski, M. Luck, and P. Tueller. 2000. Multiscale analysis of landscape heterogeneity. Geographic Information Sciences 6:6-19.

Zeide, B. 1987. Analysis of the 3/2 power law of self-thinning. Forest Science 33:517-537.

CHAPTER 3

UNCERTAINTY ANALYSIS IN ECOLOGICAL STUDIES:

An Overview

HARBIN LI AND JIANGUO WU

3.1 INTRODUCTION

Large-scale simulation models are essential tools for scientific research and environmental decision-making because they can be used to synthesize knowledge, predict consequences of potential scenarios, and develop optimal solutions (Clark et al. 2001, Berk et al. 2002, Katz 2002). Modeling is often the only means of addressing complex environmental problems that occur at large scales (Klepper 1997, Petersen 2000). For example, investigations of global climate change (Wobbles et al. 1999), regional assessments of net primary productivity and carbon dynamics (Jenkins 1999, Peters et al., Chapter 7, Law et al., Chapter 9), and landscape analysis of fire spread (Hargrove et al. 2000) rely heavily on simulation modeling at various scales. However, uncertainty in simulation modeling is often overlooked even though it is a fundamental characteristic of modeling that can be caused by incomplete data, limitations of models, and lack of understanding of underlying processes (Beck 1987, Reckhow 1994, Clark et al. 2001, Berk et al. 2002, Katz 2002, Stott and Kettleborough 2002, Urban et al., Chapter 13). If simulation results are to be useful, researchers must show the reliability of the model output by providing information about model adequacy and limitations, prediction accuracy, and the likelihood of scenarios (Clark et al. 2001, Katz 2002).

Uncertainty affects every aspect of modeling (Reckhow 1994, Klepper 1997, Jansen 1998, Katz 2002, Stott and Kettleborough 2002, Urban et al., Chapter 13). Data may contain errors that result from problems with sampling, measurement, or estimation procedures (O'Neill and Gardner 1979, Regan et al. 2002). Incomplete data are a common problem, especially in spatial modeling at broad scales. Models are imperfect because they are simplifications of real systems and always have errors in their assumptions, formulation, and parameterization. Moreover, effects of these errors on model adequacy are often insufficiently evaluated (Beck 1987, Reckhow 1994). In fact, most large-scale models are not fully validated, partly

J. Wu, K.B. Jones, H. Li, and O.L. Loucks (eds.),
Scaling and Uncertainty Analysis in Ecology: Methods and Applications, 45–66.

because validation data are not available (sometimes no data can be collected under the existing technological and logistical constraints) and partly because techniques for validating spatial models have not been perfected. Although the importance of uncertainty in modeling is well recognized, few studies of ecological modeling provide critical information about uncertainty, confidence levels or likelihood associated with simulation results (Reckhow 1994, Clark et al. 2001, Rypdal and Winiwarter 2001, Katz 2002). This lack of discussion and reporting is unfortunate because predictions that are not accompanied by information about uncertainty are of limited value in policy- or decision-making. Researchers must adopt a new modeling philosophy that requires that uncertainty in models and modeling be understood, quantified when possible, and reduced to an acceptable level when feasible.

Scaling is the translation or extrapolation of information from one scale to another in time or space or both (Blöschl and Sivapalan 1995, Wu 1999, Wu and Li, Chapters 1 and 2). For example, scaling is needed to resolve most of the large-scale management problems because most of our knowledge and data is obtained by means of small-scale research. In the process of scaling, errors in data and models may be propagated into results. It is not adequate simply to ask how to scale: one must ask how to scale with known reliability and uncertainty even when ecological systems and models involved are often complex. Thus, uncertainty analysis is an essential part of scaling because it provides critical information about the adequacy of models or algorithms used in the scaling process and about the accuracy of scaling results (Katz 2002).

In this overview, we will focus on the major concepts and techniques of uncertainty analysis associated with up-scaling methods (i.e., those that extrapolate information from fine scales to coarse scales; Wu and Li, Chapters 1 and 2). Specifically, we will identify sources of uncertainty in the scaling process and illustrate approaches to and techniques of uncertainty analysis. Because translating or extrapolating is usually done with the help of models (Wu and Li, Chapter 2), scaling can be regarded as a special case of modeling (i.e., modeling with changing scales). Therefore, most discussion of uncertainty in modeling is directly applicable to uncertainty in scaling. Also, it should be noted that most of the techniques of uncertainty analysis discussed below are more suitable for ecological models with low to intermediate complexity than for highly complex models like the general circulation models employed in climate change research (Allen et al. 2000, Forest et al. 2002, Stott and Kettleborough 2002).

3.2 UNCERTAINTY AND RELATED CONCEPTS

The term uncertainty implies some kind of error, inexactness, unreliability, and imperfection in our knowledge and understanding of the systems under investigation (Funtowwicz and Ravetz 1990, Petersen 2000, Regan et al. 2002, Katz 2002). Some degree of uncertainty is unavoidable in modeling and scaling because there are always errors associated with the stochastic nature of ecological processes, system complexity caused by spatial heterogeneity and nonlinear relationships, unreliability and unavailability of data, and/or imperfections of models (Mitchell and Hulme

1999, Katz 2002, Regan et al. 2002, Stott and Kettleborough 2002, Groffman et al., Chapter 10, Urban et al., Chapter 13). However, many of the uncertainties in simulation modeling should be quantified and reduced, and different sources of uncertainty be ranked with respect to their relative contributions to errors in model output. The term uncertainty is sometimes used to mean levels of risk involved in a scenario, action, or inaction defined by policy or management decisions, but such usage may require caution because risk implies probability and consequence that can be themselves uncertain (Pate-Cornell 1996, Katz 2002). All of these uncertainties should be studied thoroughly and systematically (Reckhow 1994), but they may require different treatments. Some uncertainties can be quantified and reduced (e.g., input uncertainty; Katz 2002), some can be quantifiable but hard to reduce (e.g., natural variability of data; Nauta 2000, Regan et al. 2002), some may be unquantifiable (e.g., model uncertainty; Beck 1987, Klepper 1997, Regan et al. 2002, Stott and Kettleborough 2002), and some may have only insignificant effects on model output in a particular modeling exercise (e.g., omission of minor processes or variables; Katz 2002).

Uncertainty analysis is the process of assessing uncertainty in modeling or scaling to identify major uncertainty sources, quantify their degree and relative importance, examine their effects on model output under different scenarios, and determine prediction accuracy (Jansen 1998, Katz 2002). Uncertainty analysis is employed primarily to determine whether the estimated prediction uncertainty is acceptable for a particular model application and, if it is not, to highlight factors whose uncertainty is pivotal in policy considerations and to recommend ways of reducing prediction error (Jansen 1998, Katz 2002). Specifically, uncertainty analysis addresses questions like: What is the magnitude of error in large-scale estimates? How is error propagated in the scaling process? Which factors are most critical, most poorly understood, or least predictable? How can errors be reduced? What is the probability that an event or scenario will occur? Uncertainty analysis can increase the credibility of modeling even when much of uncertainty may not be reduced (Rykiel 1996, Rypdal and Winiwarter 2001). Thus, understanding, quantifying, reporting, and ultimately reducing uncertainty in large-scale assessment is of great interest to policy makers.

Sensitivity analysis and scenario analysis are closely related to uncertainty analysis (Saltelli et al. 2000, Melching and Bauwens 2001, Katz 2002). The similarity is that they all involve in running models under perturbations (e.g., changes in model structure, parameters, or input data) and may use similar techniques (e.g., Monte Carlo simulation). Sensitivity analysis quantifies the rate of change in model output when one or more input variables and parameters are varied by a fixed amount or proportion while the others are held constant (Klepper 1997, Katz 2002). A more formal approach to sensitivity analysis with various statistical sampling methods may also be used (Saltelli et al. 2000). Sometimes, the absolute rate of change is converted into a relative measure to make comparisons among different parameters more meaningful (i.e., absolute vs. relative sensitivity). Sensitivity analysis is often used as a model-testing tool to examine model behavior in terms of the most sensitive parameters. A scenario is a possible future boundary condition (represented by a set of key input values or sometimes by trajectories of

key input values) on the basis of which simulations are run (Clark et al. 2001). In other words, a scenario is the IF in the WHAT-IF questions of simulation modeling. In scenario analysis, all of the inputs are changed simultaneously. Scenarios are model input rather than predictions, but they may be defined by results of previous simulation studies. Scenario analysis usually focuses on policy-relevant possibilities of future conditions (e.g., best case vs. most likely case vs. worst case), and is an effective means of communicating a large amount of technical information obtained by simulations with large, complex models (Wobbles et al. 1999, Clark et al. 2001, Katz 2002). Both sensitivity and scenario analyses may best be regarded as necessary precursors to full-fledged, probability-based uncertainty analysis (Katz 2002).

3.3 SOURCES OF UNCERTAINTY

Uncertainty in modeling may come from many different sources, which in turn may be classified into many different categories (O'Neill and Gardner 1979, Funtowwicz and Ravetz 1990, Jansen 1998, Petersen 2000, Katz 2002, Stott and Kettleborough 2002, Urban et al., Chapter 13). However, most classifications of uncertainty sources consider similar sets of factors when viewed as a whole. In this overview, we discuss three main sources of uncertainty: models themselves, input data, and scaling algorithms (O'Neill and Gardner 1979, Jansen 1998, Katz 2002, Urban et al., Chapter 13). Note that uncertainty in scaling algorithms may be regarded as part of model uncertainty, but we separate them here because it presents a unique problem and is a focus of this book.

3.3.1 Model

Model uncertainty has two basic components, model structure and model parameters (Figure 3.1; Morgan and Henrion 1990, Klepper 1997, Katz 2002). Model structure uncertainty is caused by the modeling processes of simplification and formulation. Model simplification is essential to modeling and is the identification and selection of the processes, relationships, and variables that are the most important to the system of interest and the modeling objectives. Simplification is done by assuming that some processes may be ignored because they explain an insignificant amount of variability in model predictions. Model simplification may also reflect a failure to understand certain processes. Model formulation focuses on the mathematical translations of relationships and the designs of algorithms and computer codes. Because many of the assumptions and subjective judgments must be made during model construction but are not often reported, most hidden error is created in the process of model formulation. Therefore, model structure uncertainty is the failure to include relevant processes and the unreliability caused by deficiencies in confidence, quality, and scientific basis of the equations and algorithms that represent the selected processes and their interactions.

Model parameter uncertainty is introduced by the modeling process of parameterization of models (Morgan and Henrion 1990, Klepper 1997, Katz 2002).

Model parameterization is the estimation and calibration of parameters. Calibration produces a set of optimal parameter values by forcing selected model outputs to agree with testing data. Model parameter values are built into models and may be fixed or change in space and time. Model parameter uncertainty results from imperfect knowledge about the parameters, lack of data or understanding, and errors in the estimation and calibration processes. Given the large number of parameters involved in ecological models, the first step in uncertainty analysis is often to perform sensitivity analysis to identify the parameters that may have significant effects on model output. Uncertainty analysis requires that statistical distributions (or ranges, means, variances) of parameters be known. However, a common problem in uncertainty analysis is that the accuracy of measurements and/or estimates of parameters are unknown. Modelers often have no information about the variability of parameters and have to make assumptions about parameter distributions (Urban et al., Chapter 13).

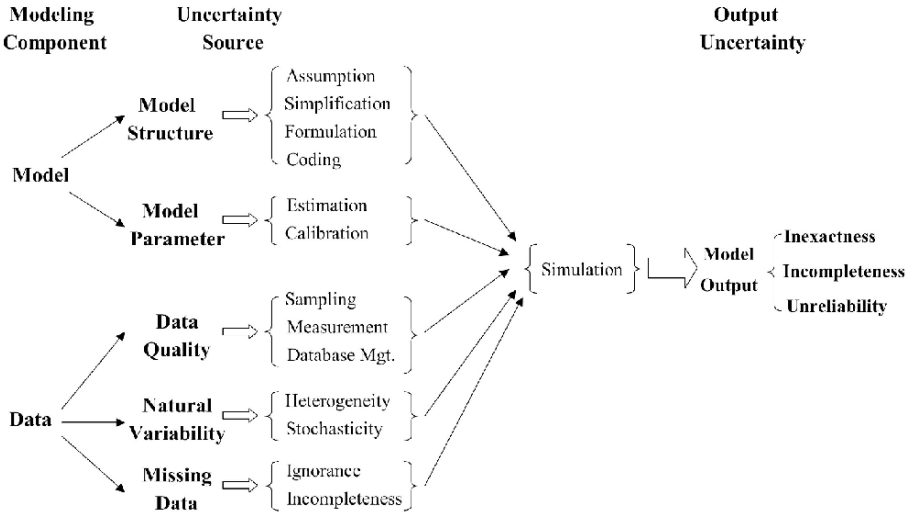


Figure 3.1. *Sources of uncertainty in scaling and modeling.*

A lesser known but perhaps more critical problem in model parameterization is the use of free (or fuzzy) parameters in model functions (Petersen 2000, Berk et al. 2002). Free parameters are those whose values are arbitrarily assigned or tuned in model calibration to make the model output fit the testing data. Problems arise when a model contains many free parameters that have no physical meanings and thus are not subject to evaluations by observation or measurement. Although the use of free parameters is often unavoidable when models are complex, their overuse can introduce large uncertainty into model output, and their uncertainty can severely diminish the value of the calibrated model (Petersen 2000). Thus, if models require intensive calibration of key parameters before they can be applied to new sites, they are of limited use in large-scale spatial simulation or scaling because data needed for

such calibration may be unavailable and because model predictions may have high uncertainty due to the calibration. Reducing the number of free parameters in models should alleviate these problems and improve models (Petersen 2000) but poses a daunting challenge to model developers.

3.3.2 Input Data

Input data are those required to run models for specific applications. Input data are usually from measurements or observations and are composed of systems attributes and driving variables. Systems attributes define the simulation settings about characteristics of the modeled system and its environment (e.g., evapotranspiration rate, leaf area index, spatial distribution of vegetation). Systems attributes are often fixed as constants even though they may have a range of values and change over time. Uncertainty in systems attributes is a major focus and is relatively easy to handle in uncertainty analysis. Driving variables define the environmental conditions (e.g., climate variables) that change in space and time, but are not affected by the model. Driving variables are often not examined in uncertainty analysis, but treated as part of the simulation conditions. This is primarily due to technical difficulties involved. However, if variability of a driving variable needs to be considered, it can be done via scenario analysis.

Data uncertainty can be caused by unknown data quality, high natural variability of the system, or lack of information (Figure 3.1; O'Neill and Gardner 1979, Jansen 1998, Regan et al. 2002). Data quality is affected by instrument or measurement errors, sampling errors, and database management errors. Data quality is always a concern in modeling and it becomes a major problem in uncertainty analysis when errors in sampling or measurement for observed data and errors in interpolation or aggregation for estimated data are not reported (Berk et al. 2002, Regan et al. 2002). The data employed in modeling exercises are usually collected for other purposes, and this often causes difficulties in model construction and testing because they may not adequately represent key characteristics of the modeled system. In such cases, critical reviews of data quality should be required before model simulations. A related problem is inappropriate use of data outside their intended purpose or domain. For example, whether model testing data are obtained independently and at the appropriate scale is a question of data quality.

Variability in ecological systems may result from spatial heterogeneity of environmental conditions or from randomness in interactions of different processes. Natural variability of data is a critical factor in uncertainty analysis that must be considered because modeling only with average values can produce severe bias in predictions, especially for nonlinear models (O'Neill 1979, Scherm and van Bruggen 1994). Spatial variability in systems attributes and driving variables need to be effectively incorporated into simulation modeling. We will discuss this point in the next section. Natural variability in data is the most studied in uncertainty analysis.

Lack of information is a pervasive problem in ecological research, especially in large-scale modeling where the emphasis is on synthesis. Unavailability of large-scale and long-term data greatly hinders uncertainty analysis and model evaluation

because large-scale processes often cannot be predicted directly from fine-scale data (Clark et al. 2001). Thus, it is imperative to obtain experimental and observational data at landscape or regional scales. Other data availability concerns include inadequacy of resolution and duration of observational studies, and gaps in temporal and spatial coverage (Clark et al. 2001). Techniques for dealing with missing data must be developed for scaling; at the present state of knowledge, data requirements often cannot be met and uncertainty analysis must be conducted with key data missing (Berk et al. 2002). Missing data is the uncertainty source that cannot be quantified (Funtowwicz and Ravetz 1990).

3.3.3 Scaling Algorithm

Scaling algorithms are a new source of uncertainty. One perspective of scaling is that it is an uncertainty problem of error propagation. This is partly because the factors that cause problems in scaling (e.g., spatial heterogeneity, nonlinearity) are also those that contribute greatly to uncertainty (Schulze 2000). Scaling in space brings forward two causes of uncertainty: mismatch of scales in model or data and spatial heterogeneity of system variables and parameters. Mismatch of scales is an issue of model adequacy and occurs when models are applied at scales different from those for which they have been developed or when the support of a model (or data) changes with changing scales. Support refers to the nature of the modeled entities, such as size, shape, orientation, and heterogeneity (Heuvelink 1998a, Wu and Li, Chapter 1), and changes of the support may cause changes in parameter values and even in functional forms of the model (Heuvelink 1998a, Katz 2002). Models are often developed for application at a specific scale or domain of scales (Reynolds et al. 1993, Heuvelink 1998a, Katz 2002). Thus, when a model is applied outside its designed domain of scales, the uncertainty of model structure may increase as a consequence of loss of model adequacy (Rykiel 1996). Spatial heterogeneity and its representation in scaling algorithms is a major source of scaling uncertainty. Ecological processes and phenomena exhibit both stochastic and patterned variations over a wide range of spatial scales. Such spatial heterogeneity increases system complexity and raises questions about the adequacy or representativeness of sampling methods and data. The problem of accounting for spatial heterogeneity explicitly in scaling (or modeling in general) is a critical challenge and remains to be resolved (Hunsaker et al. 2001, Lowell and Jaton 1999, Groffman et al., Chapter 10, Urban et al., Chapter 13). The scale-specific nature of models and data and the heterogeneous characteristics of the system must be considered in scaling.

3.4 METHODS OF UNCERTAINTY ANALYSIS

Uncertainty analysis focuses on effects of uncertainty from different sources on model output under multiple scenarios. Uncertainty analysis is not employed routinely in ecological studies mainly because the existing techniques are neither widely known nor universally applicable and effective. This section reviews the

existing techniques of uncertainty analysis and points out their key characteristics and deficiencies. We discuss: (1) model evaluation to examine model structure uncertainty, (2) examination of error propagation to quantify uncertainties in model parameters and input data, and (3) prediction accuracy and error partitioning to present scaling (or model output) uncertainty.

3.4.1 Model Evaluation: Model Structure Uncertainty

Model structure uncertainty can affect model output significantly, but it is often not addressed in traditional uncertainty analysis (Morgan and Henrion 1990, Klepper 1997, Katz 2002, Regan et al. 2002). Models that are to be employed to solve real-world problems should first be subjected to model testing. Uncertainty analysis remains important even if model uncertainties have been dealt with by means of sensitivity analysis during model testing. It is critical to establish prediction confidence, especially when models must be applied to new sites or systems. A complete analysis of uncertainty that deals with all major sources of uncertainty should be pursued whenever possible (Reckhow 1994). Thus, it is a good practice to treat evaluation of model adequacy as part of uncertainty analysis (Heuvelink 1998a).

We use the term model evaluation in preference to the controversial term model validation (Oreskes et al. 1994, Rykiel 1996). Because models are always imperfect, it is the adequacy, not the validity, of models that is to be determined. Beldring (2002) defined two important aspects of model evaluation: scientific evaluation and performance evaluation. Scientific evaluation examines the extent to which the model's behavior is consistent with prevailing scientific theory and determines whether the model can describe the physical processes of interest. Performance evaluation determines the degree to which model-predicted values agree with a corresponding set of reliable and independently obtained observations. Model evaluation is usually done in the process of model construction, but a more detailed, systematic analysis should also be performed in model applications to ensure a model's practical value.

In practice, model evaluation examines the degree of adequacy in a model's assumptions, simplifications, formulations, and predictions (Rykiel 1996). Model adequacy is defined operationally by the following criteria of model behavior and prediction accuracy (Cale et al. 1983, Oreskes et al. 1994, Rykiel 1996, Beldring 2002): (1) Models should be consistent with prevailing scientific theory and concepts; (2) Models should have no detectable flaws in internal structure and logic chain; (3) Models should contain all necessary components, critical variables and processes to achieve the objectives; (4) Models should yield predictions that agree with observations; (5) Models should incorporate well-tested submodels with sound (acceptable) algorithms or formulations; (6) Models should be used within the domain of designed applicability and scales; (7) Models should be tested for multiple state variables, in multiple years, and at multiple locations (systems); and (8) Models should produce results with acceptable confidence levels. This last criterion is added because of the recognition of the important roles that uncertainty analysis plays in establishing credibility of models. However, specifics about the

implementation of this criterion still need to be developed because what constitutes an acceptable confidence level will depend on the objectives of a particular application. Although any deviation from these criteria can cause serious problems, a determination of model adequacy is in essence a judgment that takes into account the objectives of the study and the characteristics of the system of interest (Rykiel 1996).

Some of the model structural errors identified in model evaluation can be eliminated, as when modifications of the model can successfully remove the inconsistencies with theory and the logic flaws in model formulation and algorithms. Some can be quantified, as when outputs from alternative models can be contrasted to determine differences in their agreements to observations (Jansen 1998). Some can be reduced, as when mismatch of scales in models and data can be avoided or corrected and when missing critical variables and processes can be included. Unfortunately, some model structural uncertainty is not quantifiable and cannot be reduced or eliminated because it reflects imperfections that are inherent in all models (Beck 1987, Klepper 1997, Katz 2002, Stott and Kettleborough 2002). However, there are ways to cope with this unquantifiable uncertainty. Some techniques, like the Bayesian Forecasting System (Krzysztofowicz 1999a), combine all of the untreatable uncertainty and provide some measure of it as a whole. Also, good model testing can go a long way in reducing uncertainty. The key is to perform model evaluation thoroughly and systematically.

Model comparison can provide a benchmark for and insight into model uncertainty, especially when no data are available for model testing (Klepper 1997, Jain et al. 1997, Berk et al. 2002). One model may be compared with another to identify possible problems. Large differences in key model behavior indicate a need for detailed analysis and evaluation of the model under study. Similar to the hierarchical modeling approach to scaling (Reynolds et al. 1993), a fine-scale mechanistic model can be used as a surrogate for reality in testing models of lesser complexity (Jansen 1998, Urban et al., Chapter 13). However, model uncertainty cannot be quantified through such inter-model comparison alone because the true system values and formulations are still unknown. Moreover, similarity of model predictions is not a sufficient indicator of the new model's predictive quality (Jansen 1998). Thus, model comparison is useful but limited in its capability to determine model uncertainty.

3.4.2 Error Propagation: Uncertainties in Model Parameters and Input Data

Errors propagate from model parameters and input data to model outputs in the process of modeling or scaling. In essence, to quantify errors and their propagation is to determine how variances or standard deviations of random variables get combined and manifested in the model predictions or large-scale estimates of some state variables. The variability in the state variables (i.e., model output) is then used as a measure of the output uncertainty. Many techniques can be used to analyze uncertainties in model parameters and input data. These include applications of probability theory, Taylor series expansion, Monte Carlo simulation, generalized likelihood uncertainty estimation, Bayesian statistics, and sequential partitioning

(Gardner and O'Neill 1983, Gardner et al. 1990, Rastetter et al. 1992, Heuvelink 1998b, Jansen 1998, Wiwatenadate and Claycamp 2000, Katz 2002). In this section, we will describe these major techniques of uncertainty analysis. Other techniques (e.g., the fuzzy set method, Franks and Beven 1997, Scherm 2000; the Rosenblueth's method of approximation, Heuvelink 1998b; the fingerprinting techniques of climate variability, Allen et al. 2000, Forest et al. 2002, Stott and Kettleborough 2002) also exist. We refer the reader to the cited literature for methods not covered here.

To demonstrate these techniques, we suppose a simple system with three random variables, *X*, *Y*, and *Z*, in which *Z* is a monotonous function of *X* and *Y*:

$$Z = \Phi(X, Y). \tag{3.1}$$

We use μ_Z, σ_Z^2, μ_X, σ_X^2, μ_Y and σ_Y^2 to denote the means and variances of *Z*, *X,* and *Y*. Note that *X* and *Y* can be either predictive variables or model parameters, and that Φ is a model to estimate output *Z* from inputs *X* and *Y*.

3.4.2.1 Probability theory

Probability theory method employs probability theory of moments of linear combinations of random variables to define means and variances of random functions (Wiwatenadate and Claycamp 2000). The essence of this approach is to calculate analytically the mean and variance of the output as a function of random variables of input, using the basic statistics of the random variables as follows. For the case of *Z* as an addition or subtraction function of *X* and *Y*, i.e.,

$$Z = \Phi(X, Y) = X \pm Y, \tag{3.2}$$

the mean and variance of *Z* can be calculated by the equations

$$\mu_Z = \mu_X \pm \mu_Y \tag{3.3}$$

$$\sigma_Z^2 = \sigma_X^2 + \sigma_Y^2 \pm 2 \cdot \sigma_{XY}. \tag{3.4}$$

For the case of multiplication, i.e.,

$$Z = \Phi(X, Y) = X \cdot Y, \tag{3.5}$$

similar equations can be used to calculate the mean and variance of *Z*, i.e.,

$$\mu_Z = \mu_X \cdot \mu_Y, \tag{3.6}$$

$$\sigma_Z^2 = \sigma_X^2 \cdot \sigma_Y^2 + (\mu_Y)^2 \cdot \sigma_X^2 + (\mu_X)^2 \cdot \sigma_Y^2. \tag{3.7}$$

Here, X and Y are assumed to be independent. If X and Y are correlated, the variance equation can get complicated. Similar equations can also be derived for the case of division (or ratio), but the variance equation may not always exist.

The probability theory method is powerful. Its main advantage is that it is analytical and provides exact solutions, i.e., it has neither estimation error nor approximation error. For example, it works when one knows the functional form and the basic statistics of the random variables (e.g., means, variances). For example, when dealing with relationships from literature, one does not have to have the raw data for X and Y. Thus, use of the probability theory method to study error propagation is straightforward for simple linear models. The disadvantage of the method is that it does not apply to nonlinear functions. However, it may be used to deal with uncertainties in complex models when combined with other techniques.

3.4.2.2 Taylor series expansion

The Taylor series method uses the Taylor series expansion at the point of (μ_X, μ_Y) to estimate the mean and variance of a simple function of random variables (Rastetter et al. 1992, Heuvelink 1998b). The idea of the Taylor series method is to first approximate the model by a linear function and then solve analytically for the combined error from the error propagation. In most situations, the first or the second order Taylor approximation is sufficient. Higher order Taylor methods are seldom used because the gain in reduced approximation errors may be greatly outweighed by the increased complexity.

Where the second order Taylor approximation is employed, the method can be presented as follows. For the same system defined by Equation 3.1, the mean of Z is defined by applying the original function with the means of the component random variables, and the variance of Z is estimated by the Taylor series expansion, i.e.,

$$\mu_Z \approx \Phi(\mu_X, \mu_Y), \tag{3.8}$$

$$\sigma_Z^2 \approx \left[\partial\Phi/\partial X\right]^2 \cdot \sigma_X^2 + \left[\partial\Phi/\partial Y\right]^2 \cdot \sigma_Y^2 + 2 \cdot \left[\partial\Phi/\partial X\right] \cdot \left[\partial\Phi/\partial Y\right] \cdot \sigma_{XY}, \tag{3.9}$$

where $\partial\Phi/\partial X$ and $\partial\Phi/\partial Y$ are partial derivatives of Z with respect to X and Y, and σ_{XY} is the covariance. If X and Y are independent (i.e., $\sigma_{XY} = 0$), then a simplified equation can be obtained by eliminating the covariance term in Equation 3.9.

The Taylor series method is a useful alternative to the probability theory method, which cannot be employed when the random function Φ is complex. One important characteristic of the method is that the function can be of any kind; but obviously, different types of equations will have different approximation errors. Thus, the main advantages of the method are that it is analytical, is flexible in terms of functional forms, and can be applied to models of moderate complexity. The main disadvantages are that it requires that models be presented as differentiable functions

and that the approximation errors are usually unknown. The Taylor series method has been used in ecological modeling for carbon dynamics in peatland (Bubier et al. 1999) and for water quality in a Florida watershed (Zhang and Haan 1996).

3.4.2.3 Monte Carlo simulation

The Monte Carlo method computes output statistics (means, variances) by repeating simulations with random sampling of input variables and model parameters (Gardner and O'Neill 1983, Gardner et al. 1990, Rastetter et al. 1992, Heuvelink 1998b, Jansen 1998, Katz 2002). The basic procedures are to define input distributions, sample randomly from the input distributions, run simulations with repeated samplings, and determine probability distribution for the output.

The method can be presented as follows. For a model of any complexity (e.g., Equation 3.1), a realization of Z is obtained by

$$Z_i = \Phi\left(X_i,\ Y_i\right) \tag{3.10}$$

with X and Y being defined by

$$X_i = \{\, X \mid X \sim N\left(\mu_x,\ \sigma_x\right)\} \tag{3.11}$$

$$Y_i = \{\, Y \mid Y \sim N\left(\mu_y,\ \sigma_y\right)\} \tag{3.12}$$

where X_i and Y_i are values obtained from samples of normal distributions, and all other terms are the same as defined previously. With repeated sampling of size N, the statistics of the output can be calculated by

$$\mu_Z = \sum_i Z_i \Big/ N \tag{3.13}$$

$$\sigma_Z^2 = \sum_i \left(Z_i - \mu_Z\right)^2 \Big/ N - 1. \tag{3.14}$$

Notice that both X and Y are assumed to be independent and have normal distributions. Joint distributions are required if the independence assumption does not hold. Other theoretical distributions (e.g., lognormal, uniform) may also be used.

The Monte Carlo method is the most commonly used technique for uncertainty analysis because it has no strict requirement about the exact formulation of the function and is therefore easily implemented and generally applicable. The function can be either a simple empirical model or a complicated dynamic model; the steps of Monte Carlo error analysis effectively remain the same because the method treats the function as a black box (i.e., only input and output are considered) and studies the resulting outputs by systematically sampling from the input space. The main disadvantages of the Monte Carlo method are that the results are not in an analytical

form and that joint distributions for correlated variables are often unknown or difficult to derive. The Monte Carlo method is computationally intensive, but one can employ sampling schemes to reduce the computation burden. A common sampling scheme is the Latin hypercube sampling in which the range of a variable is stratified and each stratum is sampled once with an equal probability (McKay et al. 1979). The Monte Carlo method has been used to study uncertainty related to forest dynamics (Gardner et al. 1990), water quality (Gardner and O'Neill 1983, Zhang and Haan 1996), soil acidification at the European scale (Kros et al. 1999), and nitrate leaching at a regional scale (Hansen et al. 1999).

3.4.2.4 Generalized likelihood method

Generalized likelihood uncertainty estimation (GLUE) is a statistical technique for simultaneously calibrating the parameter and estimating the uncertainty of predictive models (Beven and Binley 1992, Zak and Beven 1999). It operates on the assumption that many parameter sets may be equally acceptable in producing reasonable simulations of the observed data. The method can be presented as follows (Zak and Beven 1999, Brazier et al. 2000). First, a likelihood measure is selected to determine the goodness of fit in comparing observations with model predictions. This measure is chosen on basis of its appropriateness in relation to the model, the observed data, and the objectives of the study. For example, if the absolute error is chosen as the likelihood measure, the likelihood function for a single observation is given by

$$L\left(Z \mid Z_{Obs}\right) = \left[1 - \left(\left|Z_{Obs} - Z\right| \middle/ Z_{Obs}\right)\right]^{N} \tag{3.15}$$

where Z_{Obs} is the observed value of the predicted variable Z, and N is the shaping factor of the likelihood function (Zak and Beven 1999). Second, Monte Carlo simulations are run with all parameter sets, using the following equations,

$$X_i = \left\{X \mid X \sim U(a,b)\right\} \tag{3.16}$$

$$Y_i = \left\{Y \mid Y \sim U(c,d)\right\} \tag{3.17}$$

where X and Y are uniformly distributed parameters with ranges of (a,b) and (c,d), respectively. Third, a predetermined threshold, L^*, is used with the likelihood measure to identify and exclude those parameter sets that perform poorly (i.e., having a likelihood of zero or below the threshold). This process defines an acceptable parameter space, R_a, which is composed of values of all physically reasonable parameter sets from the potential parameter space in Monte Carlo simulations, i.e.,

$$X_i^*, Y_i^* \in R_a \tag{3.18}$$

given that

$$L\left(Z_i^* \mid Z_{Obs}\right) > L^*. \tag{3.19}$$

Finally, the simulation results with acceptable parameter sets are then to define likelihood weighted distributions of predicted results from which uncertainty bounds are derived.

GLUE is a hybrid technique for parameter calibration and output uncertainty assessment that combines Monte Carlo simulations and likelihood analysis. The advantage of GLUE is that it provides a probabilistic distribution of model prediction so that output uncertainty is defined. The probabilistic distribution is enhanced by the selection of the acceptable parameter space because only physically reasonable parameter sets are used. The disadvantage of GLUE is that it does not explicitly consider the effects of individual parameters on model predictions because it considers only sets of parameter values. However, sensitivity analysis may be performed to determine the relative importance of model parameters and the processes they represent. The same sensitivity analysis can also be used to reduce the number of parameters used in GLUE by selecting only those that show significant effects on model predictions (Zak and Beven 1999). Another disadvantage is that the applicability of GLUE is limited by the need for observations against which model predictions can be compared in the likelihood analysis. Such observations are often unavailable in scaling projects. GLUE has been used primarily in hydrological modeling (Zak and Beven 1999, Brazier et al. 2000).

3.4.2.5 Bayesian statistics

Bayesian statistical methods quantify uncertainty by calculating probabilistic predictions. The procedure has three stages: (1) determination of the prior probability distribution for model parameters, (2) construction of a likelihood function for the statistical model, and (3) derivation of the posterior probability distribution for the parameters by using the Bayes rule to adjust the prior distribution based on the observed data (Katz 2002). The Bayes rule states that the posterior probability distribution is proportional to the prior probability distribution multiplied by the likelihood, i.e.,

$$P\left(Z \mid Z_{Obs}\right) = P(Z) \cdot P\left(Z_{Obs} \mid Z\right) \Big/ P\left(Z_{Obs}\right) \tag{3.20}$$

where $P\left(Z \mid Z_{Obs}\right)$ is the posterior probability distribution for the predicted variable Z given the observations Z_{Obs}, $P(Z)$ is the prior probability distribution of Z, $P\left(Z_{Obs} \mid Z\right)$ is the conditional distribution of the observations, and $P(Z_{Obs})$ is the marginal distribution of the observations. The probabilistic predictions generated by Bayesian statistical methods are used to define modeling uncertainty.

One example of such methods is the Bayesian Forecasting System (BFS) developed for deterministic hydrologic models by Krzysztofowicz (1999a, 1999b). The BFS first identifies the random inputs whose uncertainty significantly affects the model outputs and varies from forecast to forecast. Then, the BFS decomposes the total uncertainty into input uncertainty and model uncertainty. Krzysztofowicz (1999a, 1999b) proposed that model uncertainty should be the combined uncertainty from all other sources, including imperfections of the model, incorrect estimates of parameters, and incorrect estimates of deterministic inputs. The BFS has three components (steps): (1) the input uncertainty processor, which runs simulations with parameter values of the random inputs and defines the uncertainty in model output caused by the input uncertainty; (2) the model uncertainty processor, which yields the posterior density based on the prior density and the likelihood function; and (3) the integrator, which integrates input uncertainty and model uncertainty into a predictive distribution. The characteristics of Bayesian statistical methods in general and BFS in particular are that they are a process of learning from data, require prior probability distributions and observations, and provide probabilistic predictions in the form of posterior distributions.

3.4.2.6 Sequential partitioning

The sequential partitioning method is not a new technique, but rather a hybrid approach based on a new strategy (Rastetter et al. 1992). It may apply a combination of probability theory, Taylor series, and Monte Carlo methods in clearly defined steps or modules in the process of modeling or scaling up. The sequential partitioning method should be useful to deal with uncertainty in complex models when other methods are difficult to implement. The procedures employed vary from situation to situation, but the general approach is as follows. Complex models are first divided into independent compartments or modules, and appropriate techniques are used to identify the most critical variables in each module. Various methods are then employed to examine uncertainty associated with the critical variables from different compartments. For example, to study uncertainty in net ecosystem productivity, one may first use the Monte Carlo method to determine uncertainties related to the processes of photosynthesis and respiration separately, and then apply the probability theory method to combine the findings of these analyses. As all of the previously discussed methods of uncertainty analysis have limitations and as ecological models are getting too complex, the sequential partitioning method may be a promising alternative, given that successful division of complex models into independent compartments can be achieved. This approach requires the full access to the source codes of the models, a requirement that often cannot be met (Urban et al., Chapter 13).

3.4.3 Presentation of Prediction Uncertainty

Uncertainty from different sources is manifested in model output, and effective communication of the manifestations of uncertainty in modeling is a critical

component of uncertainty analysis. In this section we use some examples to show how results of uncertainty analysis can be summarized and reported as output uncertainty, prediction accuracy, and error contributions of critical factors.

Output uncertainty is usually presented quantitatively as the probability distribution or statistical characteristics of model predictions. In the BFS and GLUE approaches, the uncertainty of each model prediction is defined by the probability distribution of model output. Thus, no additional analysis is needed. In other approaches (e.g., Taylor series, Monte Carlo simulation), the variance of an output variable is a common measure of error (Heuvelink 1998b, Rypdal and Winiwarter 2001). Often, the confidence levels for predicted values of system variables are presented as indicators of degree of uncertainty, and coefficient of variation (CV) is used to compare uncertainties associated with different variables or different applications. There are no simple criteria for judging the acceptability of estimated prediction uncertainty in model applications. However, a clear indicator of unacceptable uncertainty is that the simulation results cannot be used to determine the basic status of the system being modeled. For example, high uncertainty may prevent researchers from unequivocally answering the following questions: Is the system a carbon sink or source? Does a system perturbation generate a positive effect or negative effect or no effect at all on the processes of interest? Failure to answer such questions will limit effective applications of simulation modeling to resource management.

Many measures of prediction accuracy can be used in uncertainty analysis (Kvalseth 1985, Armstrong and Collopy 1992, Mayer and Butler 1993, Beldring 2002). For example, the Nash-Sutcliffe modeling efficiency index, R^2, is often used to assess the goodness of fit between model predictions and observations. The index is defined as (Nash and Sutcliffe 1970):

$$R^2 = 1 - \sum_i \left(Z_i - \hat{Z}_i\right)^2 \Big/ \sum_i \left(Z_i - \mu_Z\right)^2 \tag{3.21}$$

where Z_i is the observed value of Z and $\hat{Z}_i$ is the simulated value of Z. The values of R^2 range from minus infinity to 1.0, with higher values indicating better agreement. Other error measures include root mean square error (RMSE) and mean absolute percentage error (MAPE). In addition, the goodness of fit may also be revealed by graphic displays of results, such as the plot of observed (Z) against predicted ($\hat{Z}$) values or the plot of residual ($Z_i - \hat{Z}_i$) against predicted values. The selection of an error measure depends on the situation. None of the error measures is best in all circumstances (Armstrong and Collopy 1992).

Determining how much error each of the critical factors contributes to the total uncertainty is important because this information may indicate how uncertainty can best be reduced. The questions to resolve are: What is the ranking of the relative contributions of factors to the total uncertainty of model output? Which factor is the most critical uncertainty source? The usual approach is to use the variance of the output distribution as a measure of prediction uncertainty because the variance can

often be decomposed into meaningful parts. When the output variance can be partitioned, the analysis of uncertainty contributions becomes essentially a form of analysis of variance (Heuvelink 1998b). Although partitioning of variances is not always possible, it can be achieved by many techniques of uncertainty analysis. Notice that, for models with multiple outputs, rankings of input parameters as sources of uncertainty are not unique, but output specific (Klepper 1997). If a single set of ranks is of interest, one can use the ranks of output variables as weights to derive a composite ranking of input parameters for the model. Below, we discuss methods of partitioning variance and determining error contributions in association with sensitivity analysis, probability theory, Taylor series, and Monte Carlo simulation procedures.

Klepper (1997) described a simple technique for determining the relative importance of output variables to input parameters in sensitivity analysis. The basic procedure is to run Monte Carlo simulations with parameter values sampled from the parameter space, obtain a linear regression model of the output variable on the corresponding parameters, and calculate relative sensitivity as an aid to interpretation of the results. For the simple system defined in Equation 3.1, the regression model is in the form of

$$Z = a + b \cdot X + c \cdot Y \tag{3.22}$$

where a, b, and c are the regression coefficients. The relative sensitivity is given by the standardized regression coefficient, i.e.,

$$S_X = b \cdot \sigma_X / \sigma_Z \tag{3.23}$$

$$S_Y = c \cdot \sigma_Y / \sigma_Z \tag{3.24}$$

where S_X and S_Y are the relative sensitivity indices of variables X and Y, respectively. Equations 3.22, 3.23, and 3.24 can easily be extended to a general system with more variables and parameters. This approach is effective if the coefficient of determination of the regression model is high; otherwise, additional analyses are needed.

For the analytical methods of uncertainty analysis (i.e., probability theory, Taylor series), variance partitioning is straightforward because the prediction error is already decomposed and treated as a function of the variances of the independent variables. The relative contribution of a variable to the total uncertainty is defined by the fraction of the terms associated with its variance (e.g., Equations 3.4, 3.7, and 3.9). For example, the relative contribution by variable X as calculated by the Taylor series method (see Equations 3.8 and 3.9) can be expressed as

$$U_X = f\left(\sigma_X^2\right) \Big/ \sigma_Z^2 \tag{3.25}$$

$$f\left(\sigma_X^2\right) = \left(\partial\Phi / \partial X\right)^2 \cdot \sigma_X^2 + \left(\partial\Phi / \partial X\right) \cdot \left(\partial\Phi / \partial Y\right) \cdot \sigma_{XY} \tag{3.26}$$

where U_X is the relative contribution to the total uncertainty by variable *X*, and $f(\sigma_X^2)$ is uncertainty associated with variable *X*. Equation 3.26 assumes that the covariance is shared equally by the two variables involved. Similar equations can be derived for *Y*. This analysis can also be extended to a general system with more variables and parameters.

When the Monte Carlo simulation method is employed, error contributions of critical factors must be determined indirectly because the variances are not given in an analytical form. Similar to the regression method discussed for sensitivity analysis, correlation analysis can be used with simulation results. Melching and Bauwens (2001) used correlation coefficients between model parameters (or variables) and model outputs to rank the importance of the parameters in predicting pollutant loading in streams. A more complicated method requires that, in addition to the Monte Carlo simulation that considers the uncertainty of all factors (i.e., σ_Z^2), new simulations be conducted to define the uncertainty caused by a particular factor (Katz 2002).

Katz (2002) described two approaches: the absence effect approach and the presence effect approach. The absence effect approach requires a new simulation with the uncertainties of all factors but one (e.g., X) to calculate the top-marginal variance, which is defined as the expected reduction of prediction variance if the uncertainty of factor *X* is assumed to become perfectly known (Katz 2002). The procedure is represented by the following equations:

$$U_{-X} = \sigma_{-X}^2 \big/ \sigma_Z^2 \tag{3.27}$$

$$\sigma_{-X}^2 = \sigma_Z^2 - \sigma_{ALL-X}^2 \tag{3.28}$$

where U_{-X} is the relative error contribution of the factor *X*, σ_{-X}^2 is top-marginal variance of *X*, σ_Z^2 is the variance caused by all factors, and σ_{ALL-X}^2 is the output variance calculated without factor *X*. Similarly, the presence effect approach uses a different new simulation to consider the uncertainty only from factor X in calculating the bottom-marginal variance, which is defined as the prediction variance caused by the factor of interest when all other factors are assumed to be perfectly known (Katz 2002). The equation for the error contribution is given by

$$U_{+X} = \sigma_{+X}^2 \big/ \sigma_Z^2 \tag{3.29}$$

where U_{+X} is the relative error contribution of the factor *X* and σ_{+X}^2 is the bottom-marginal variance of *X*. For linear models with independent sources of uncertainty, U_{-X} and U_{+X} are the same and the sum of uncertainty contributions from all sources is equal to the total prediction variance. For complex models, however, both

the absence and presence effect approaches should be used to define the range of the relative error contribution of a particular factor (Katz 2002).

3.5 CONCLUDING REMARKS

Future research should focus on bridging data gaps and developing new techniques of uncertainty analysis so that uncertainty in complex models can be assessed effectively. Lack of good data is the most critical obstacle to uncertainty analysis (Berk et al. 2002). Conducting well-designed field experiments and observations, especially at large scales, to meet the data requirements for uncertainty analysis should be a top priority. However, the problem of inadequate data may be partially resolved by improving the way data are reported: the variability of key variables and parameters (e.g., variance, range) should be presented together with the mean values. Another challenge in uncertainty analysis is the high complexity of models needed to address environmental assessment and resource management issues at large scales. This difficulty may be resolved by the sequential partitioning method discussed above and the disintegrated uncertainty analysis approach recommended by Katz (2002). Both approaches imply a strategy of keeping uncertainty analysis simple and doable by assessing uncertainties in individual model components separately. Nonetheless, effective techniques should be developed to conduct uncertainty analysis with complex systems when data are incomplete and models are insufficiently verified. For example, quantitative information about spatial heterogeneity should be incorporated into scaling procedures to reduce uncertainty and improve predictions.

Development of new techniques or innovative ways of using existing techniques should also be directed at creating capabilities of providing ideal outputs of uncertainty analysis – those that can fully characterize the uncertainty involved in modeling or scaling. These desirable outputs of uncertainty analysis include: (1) measures of model adequacy, (2) full probability distributions of model outputs (e.g., density function, probability-weighted values), (3) reliability of model results (e.g., accuracy, confidence level, error), (4) relative contribution or importance of each factor as an error source to total uncertainty, (5) the likelihood of different scenarios (probability or ranking), and (6) identification of the least understood or predictable components of the model (critical factors).

It is imperative that prediction uncertainty be treated as a critical issue and uncertainty analysis as a mandatory component in scaling because uncertainty in scaling is inevitable and should be assessed thoroughly to ensure the credibility and reliability of scaling results. Important sources of uncertainty in scaling include scaling algorithms, model parameters, quality and natural variability of data, and heterogeneous environment. These uncertainties must be quantified and reduced to ensure that scaling results are used effectively in policy- and decision-making. The existing techniques of uncertainty analysis (e.g., Taylor series expansion, Monte Carlo simulation, Bayesian statistics) can provide basic information about prediction accuracy, effects of uncertainty from different sources on scaling results, and the relative importance of individual sources of uncertainty even though they have some major limitations. The use of uncertainty analysis in ecological studies has been

rather limited. However, with recognition of the importance of uncertainty analysis in both research and application, uncertainty analysis will become an integral part of modeling and scaling.

ACKNOWLEDGEMENTS

The authors wish to thank Helen Regan, Devendra Amatya, and Debra Peters for their valuable suggestions to an earlier draft of the manuscript. JW acknowledges support from EPA (STAR R827676-01-0) and NSF (DEB 9714833, CAPLTER). This paper has not been subjected to the Agencies' required peer and policy review and therefore does not necessarily reflect the views of the agencies and no official endorsement should be inferred.

REFERENCES

Allen, M. R., P. A. Stott, J. F. B. Mitchell, R. Schnur, and T. L. Delworth. 2000. Quantifying the uncertainty in forecasts of anthropogenic climate change. Nature 407:617-620.

Armstrong, J. S., and F. Collopy. 1992. Error measures for generalizing about forecasting methods: empirical comparisons. International Journal of Forecasting 8:69-80.

Beck, M. B. 1987. Water quality modeling: a review of the analysis of uncertainty. Water Resources Research 23(8):1393-1442.

Berk, R. A., P. Bickel, K. Campbell, S. Keller-McNutly, E. Kelly, and J. Sacks. 2002. Workshop on Statistical Approaches for the Evaluation of Complex Computer Models. Statistical Science 17: 173-192.

Beldring, S. 2002. Multi-criteria validation of a precipitation-runoff model. Journal of Hydrology 257:189-211.

Beven, K., and A. Binley. 1992. The future of distributed models: model calibration and uncertainty prediction. Hydrological Processes 6:279-298.

Blöschl, G., and M. Sivapalan. 1995. Scale issues in hydrological modelling: a review. Hydrological Processes 9:251-290.

Brazier, R. E., K. J. Beven, J. Freer, and J. S. Rowan. 2000. Equifinality and uncertainty in physically based soil erosion models: application of the GLUE methodology to WEPP – the water erosion prediction project – for sites in the UK and USA. Earth Surface Processes and Landforms 25: 825-845.

Bubier, J. L., S. Frolking, P. M. Crill, and E. Linder. 1999. Net ecosystem productivity and its uncertainty in a diverse boreal peatland. Journal of Geophysical Research 104:27683-27692.

Cale, W. G., R. V. O'Neill, and H. H. Shugart. 1983. Development and application of desirable ecological models. Ecological Modelling 18:171-186.

Clark, J. S., S. Carpenter, M. Berber, S. Collins, A. Dobson, J. A. Foley, D. M. Lodge, M. Pascual, R. Pielke, W. Pizer, C. Pringle, W. V. Reid, K. A. Rose, O. E. Sala, W. H. Schlesinger, D. H. Wall, and D. Wear. 2001. Ecological forecasts: an emerging imperative. Science 293:657-660.

Forest, C. E., P. H. Stone, A. P. Sokolov, M. R. Allen, and M. D. Webster. 2002. Quantifying uncertainties in climate system properties with the use of recent climate observations. Science 295:113-117.

Franks, S. W., and K. J. Beven. 1997. Estimation of evapotranspiration at the landscape scale: a fuzzy disaggregation approach. Water Resources Research 33(12):2929-2938.

Funtowwicz, S. O., and J. R. Ravetz. 1990. Uncertainty and Quality in Science for Policy. Kluwer Academic Publishers, Dordrecht.

Gardner, R. H., and R. V. O'Neill. 1983. Parameter uncertainty and model predictions: a review of Monte Carlo results. Pages 245-257 *in* M. B. Berk and G. V. Straten, editors. Uncertainty and Forecasting of Water Quality. Springer-Verlag, New York.

Gardner, R. H., V. H. Dale, and R. V. O'Neill. 1990. Error Propagation and uncertainty in process modeling. Pages 208-219 *in* R. K. Dixon, R. S. Meldahl, G. A. Ruark, and W. G. Warren, editors. Forest Growth: Process Modeling of Response to Environmental Stress. Timber Press, Portland.

Hansen, S., M. Thorsen, E. J. Pebesma, S. Klesschulte, and H. Svendsen. 1999. Uncertainty in simulated nitrate leaching due to uncertainty in input data. A case study. Soil Use and Management 15:167-175.

Hargrove, W. W., R. H. Gardner, M. G. Turner, W. H. Romme, and D. G. Despain. 2000. Simulating fire patterns in heterogeneous landscapes. Ecological Modelling 135:243-263.

Heuvelink, G. B. M. 1998a. Uncertainty analysis in environmental modelling under a change of spatial scale. Nutrient Cycling in Agroecosystems 50:255-264.

Heuvelink, G. B. M. 1998b. Error Propagation in Environmental Modeling with GIS. Taylor & Francis, London.

Hunsaker, C. T., M. F. Goodchild, M. A. Friedl, and T. J. Case, editors. 2001. Spatial Uncertainty in Ecology: Implications for Remote Sensing and GIS Applications. Springer-Verlag, New York.

Jain, A. K., H. S. Kheshgi, and D. J. Wobbles. 1997. Is there an imbalance in the global budget of bomb-produced radiocarbon? Journal of Geophysical Research 102:1327-1333.

Jansen, M. J. W. 1998. Prediction error through modeling concepts and uncertainty from basic data. Nutrient Cycling in Agroecosystems 50:247-253.

Jenkins, J. C., D. W. Kicklighter, S. V. Ollinger, J. D. Aber, and J. M. Melillo. 1999. Sources of variability in net primary production predictions at a regional scale: a comparison using PnET-II and TEM4.0 in Northeastern US forests. Ecosystem 2:555-570.

Katz, R. W. 2002. Techniques for estimating uncertainty in climate change scenarios and impact studies. Climate Research 20:167-185.

Klepper, O. 1997. Multivariate aspects of model uncertainty analysis: tools for sensitivity analysis and calibration. Ecological Modelling 101:1-13.

Kros, J., E. J. Pebesma, G. J. Reinds, and P. A. Finke. 1999. Uncertainty assessment in modeling soil acidification at the European scale: a case study. Journal of Environmental Quality 28:366-377.

Krzysztofowicz, R. 1999a. Bayesian theory of probabilistic forecasting via deterministic hydrologic model. Water Resources Research 35:2739-2750.

Krzysztofowicz, R. 1999b. Bayesian Forecasting via Deterministic Model. Risk Analysis 19(4):739-749.

Kvalseth, T. O. 1985. Cautionary note about R^2. The American Statistician 39(4):279-285.

Lowell, K., and A. Jaton, editors. 1999. Spatial Accuracy Assessment – Land Information Uncertainty in Natural Resources. Ann Arbor Press, Chelsea.

Mayer, D. G., and D. G. Butler. 1993. Statistical validation. Ecological Modelling 68:21-32.

McKay, M. D., W. J. Conover, and R. J. Beckman. 1979. A comparison of three methods for selection values of input variables in the analysis of output from a computer code. Technometrics 2:239-245.

Melching, C. S., and W. Bauwens. 2001. Uncertainty in coupled nonpoint source and stream water quality models. Journal of Water Resources Planning and Management 127:403-413.

Mitchell, T. D., and M. Hulme. 1999. Predicting regional climate change: living with uncertainty. Progress in Physical Geography 23:57-78.

Morgan, M. G., and M. Henrion. 1990. Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis. Cambridge University Press, Cambridge.

Nash, J. E., and J. V. Sutcliffe. 1970. River flow forecasting through conceptual models. Part 1 – a discussion of principle. Journal of Hydrology 10:282-290.

Nauta, M. J. 2000. Separation of uncertainty and variability in quantitative microbial risk assessment models. International Journal of Food Microbiology 57:9-18.

O'Neill, R. V. 1979. Natural variability as a source of error in model predictions. Pages 23-32 *in* G. S. Innis and R. V. O'Neill, editors. Systems Analysis of Ecosystems. International Co-operative, Fairland, MD.

O'Neill, R. V., and R. H. Gardner. 1979. Sources of uncertainty in ecological models. Pages 447-463 *in* B. P. M. Zeigles, M. S. Elzas, G. J. Klir, and T. I. Oren, editors. Methodology in Systems Modelling and Simulation. North Holland, Amsterdam, Holland.

Oreskes, N., K. Shrader-Frechette, and K. Belitz. 1994. Verification, validation, and confirmation of numerical models in the earth sciences. Science 263:641-646.

Pate-Cornell, M. E. 1996. Uncertainties in risk analysis: six levels of treatment. Reliability Engineering and System Safety 54:95-111.

Petersen, A. C. 2000. Philosophy of climate science. Bulletin of the American Meteorological Society 81(2):265-271.

Rastetter, E. B., A. W. King, B. J. Cosby, G. M. Hornberger, R. V. O'Neill, and J. E. Hobbie. 1992. Aggregating fine-scale ecological knowledge to model coarser-scale attributes of ecosystems. Ecological Applications 2:55-70.

Reckhow, K. H. 1994. Water quality simulation modeling and uncertainty analysis for risk assessment and decision making. Ecological Modelling 72:1-20.

Regan, H., M. Colyvan, and M. A. Burgman. 2002. A taxonomy and treatment of uncertainty for ecology and conservation biology. Ecological Applications 12:618-628.

Reynolds, J. F., D. W. Hilbert, and P. R. Kemp. 1993. Scaling ecophysiology from the plant to the ecosystem: a conceptual framework. Pages 127-140 *in* J. R. Ehleringer and C. B. Field, editors. Scaling Physiological Processes: Leaf to Globe. Academic Press, San Diego, CA.

Rykiel, E. J. 1996. Testing ecological models: the meaning of validation. Ecological Modelling 90: 229-244.

Rypdal, K., and W. Winiwarter. 2001. Uncertainties in greenhouse gas emission inventories – evaluation, comparability and implications. Environmental Science & Policy 4:107-116.

Saltelli, A., K. Chan, and E. M. Scott, editors. 2000. Sensitivity Analysis. Wiley, Chichester, UK.

Scherm, H. 2000. Simulating uncertainty in climate-pest models with fuzzy numbers. Environmental Pollution 108:373-379.

Scherm, H., and A. H. C. van Bruggen. 1994. Global warming and nonlinear growth: how important are changes in average temperature? American Phytopathological Scoiety 84:1380-1384.

Schulze, R. 2000. Transcending scales of space and time in impact studies of climate and climate change on agrohydrological responses. Agriculture, Ecosystem and Environment 82:185-212.

Stott, P. A., and J. A. Kettleborough. 2002. Origins and estimates of uncertainty in predictions of twenty-first century temperature rise. Nature 416:723-726.

Wiwatenadate, P., and H. G. Claycamp. 2000. Error propagation of uncertainties in multiplicative models. Human and Ecological Risk Assessment 6:355-368.

Wobbles, D. J., A. Jain, J. Edmonds, D. Harvey, and K. Hayhoe. 1999. Global change: state of the science. Environmental Pollution 100:57-86.

Wu, J. 1999. Hierarchy and scaling: extrapolating information along a scaling ladder. Canadian Journal of Remote Sensing 25:367-380.

Zak, S. K., and K. J. Beven. 1999. Equifinality, sensitivity and predictive uncertainty in the estimation of critical load. The Science of the Total Environment 236:191-214.

Zhang, J., and C. T. Haan. 1996. Evaluation of uncertainty in estimated flow and phosphorus loads by FHANTM. Applied Engineering in Agriculture 12(6):663-669.

CHAPTER 4

MULTILEVEL STATISTICAL MODELS AND ECOLOGICAL SCALING

RICHARD A. BERK AND JAN DE LEEUW

4.1 INTRODUCTION

A useful way to conceptualize ecological processes operating at different spatial scales is through what Wu (1999) calls hierarchical patch dynamics. A key notion is that a few parts of a large hierarchical structure can be studied in isolation insofar as these parts are distinguished from the rest by "near-decomposability." In effect, a segment of special interest within the hierarchical structure interacts weakly with the rest and then only asymmetrically. In this chapter, we focus on a particular kind of segment comprised of nested elements; higher levels are composed of the components of the level below. We consider multilevel statistical models that can be used to describe how variables characterizing higher levels affect processes operating at lower levels.

For simplicity, consider a subset of a hierarchy with two levels. The basic idea is to have a regression equation characterizing relationships at the lower (or micro) level and then have one or more of the regression coefficients at the micro level a function of predictors at the macro level. At the micro level, for instance, taxa richness may be a function of stream velocity (and other things). Then at the macro level, the regression coefficient linking stream velocity to taxa richness may be a function of proximity of the stream to land used for agriculture. Thus, one can address how the relationship between stream velocity and taxa richness varies (or not) in different locations, here with locale characterized by proximity to land use for agriculture. That is, one can learn when to generalize over sites and when not to generalize over sites. One can also learn how different scales are related.

These sorts of relationships can easily be formulated as interaction effects within a conventional regression analysis. However, the usual estimation procedures will not properly characterize the uncertainty in the output, so that the confidence intervals and hypothesis tests will not perform properly. A key problem is that the model's errors (or disturbances) are not likely to behave as if drawn independently from a single distribution. Special estimation procedures are required. Such

J. Wu, K.B. Jones, H. Li, and O.L. Loucks (eds.),
Scaling and Uncertainty Analysis in Ecology: Methods and Applications, 67–88.

procedures, often constructed within a multilevel framework, are well known and widely available in existing software (Raudenbush and Bryk, 2002). Our goal, therefore, is to summarize some recent extensions of multilevel models to more complicated and realistic situations common in ecological research. In the next section we provide an overview of the work. Technical details follow in subsequent sections (the more formal theoretical work here is primarily the work of de Leeuw).

4.2 EXTENSIONS OF MULTILEVEL MODELING

Our first extension of traditional multilevel modeling allows for spatial autocorrelation in the residuals of multilevel models. More proximate spatial units at the micro level can be expected to have model disturbances that are more alike than spatial units at the micro level that are more distant from one another. Thus, transects that are closer together will likely have disturbances that are more similar than transects that are farther apart. Failing to take this spatial autocorrelation into account will generally lead to biased estimates of the standard errors, and hence, inaccurate confidence intervals and hypothesis tests. Uncertainty will be characterized incorrectly.

Formally, a good solution to this problem for linear regression can be found in a classic paper by Ord (1975). For the usual sorts of regression models, one constructs a matrix capturing the distance between all micro units within each macro unit (e.g., transects within research sites) and builds that information into the estimation process. We initially adopted this approach, introduced it into a multilevel formulation, and applied it to two data sets. One data set was collected to study biodiversity in streams located in Ventura County, California, and the other was collected to study the impact of marine preserves on biodiversity and total fish biomass in coral reefs in the Philippines.

Our early results were disappointing. First, there was essentially no theory or empirical research in ecology or related disciplines to inform in sufficient detail the construction of the distance matrix. One difficulty was, for example, that it was not clear how to measure distance given ocean currents, which transport nutrients more readily between some locals than others. “Distance” was a function of spatial proximity and the direction and speed of prevailing current; locations formally closer together could easily have less in common than locations farther apart. Another difficulty was that there are a number of different distance functions that could have been used in the distance matrix (e.g., exponential decay with increasing distance) and, again, there is no guidance from the scientific literature. We believe that similar problems are common for a wide variety of environmental applications.

Second, except for very simple and somewhat unrealistic models, the numerical methods used in the estimation did not perform well. There were several technical reasons, but a key obstacle was that the regression coefficients and the distance matrix were “competing” for much the same information. This was because the predictors necessarily also contained spatial information. Micro units that were closer were likely to be more similar not just in their disturbances, but in the values of their predictors. Such predictors could include composition of the streambed and the amount of shading from trees along the banks, for instance. Because of the

competition for spatial information, the output from the statistical models tended to be very unstable. Small changes in the model or the data could introduce large changes in the output, which is a sure sign of trouble.

Finally, we planned to move beyond multilevel linear models to multilevel generalized linear models. In generalized linear models outcome variables can be counts or proportions. Thus, we would be able to include popular procedures, such as logistic regression for binary outcomes and Poisson regression for count data. Unfortunately, the Ord's approach led to effectively intractable mathematical problems when applied to generalized linear models.

These difficulties forced us to reconsider the entire enterprise and indeed, the usual philosophy by which spatial modeling is undertaken. To begin, we suspect that for spatial regression models, far too much is made about the exact form of the distance matrix. With scant scientific guidance about how the distance matrix should be formulated, any one of several competing formulations can be applicable. But, there is no way to know which is the best. In addition, the distance matrix by itself is rarely of much scientific interest. Its usual role is to allow for more accurate estimates of the regression coefficients that are the real focus of scientific concern. In statistical parlance, the distance matrix represents a set of "nuisance parameters."

At a deeper level, George Box's famous dictum applies: "all models are wrong, but some are useful." Given the current state of subject-matter knowledge, it is naive to aim for the "right" model. And in the absence of the right model, many of the usual statistical concerns become relatively unimportant. In particular, confidence intervals and tests no longer have much probative value. Rather, one should develop models that are informative and relatively simple and that capture in broad-brush strokes the essential features of the empirical world at hand (Berk 2003).

These and other considerations led us to consider methods by which the distance matrix could be well approximated and in a manner that eliminated much of the instability produced by taking the Ord approach. Two methods now seem to be especially effective. The first method extracts the eigenvectors of the distance matrix and uses the first few to adjust for spatial autocorrelation. This still requires, however, that a distance matrix be specified. The second method constructs simple functions of the spatial coordinates (e.g., longitude and latitude) and uses these to adjust for spatial autocorrelation. For example, one might include longitude, latitude, and their product. Analyses of real data and our own simulations indicate that both methods work well, although the second method is somewhat simpler to implement. Moreover, one can in both cases improve the approximation of the distance matrix as much as desired by using more of its eigenvectors or more complicated functions of the spatial coordinates. That is, one can make the approximations arbitrarily close to the specified distance matrix, although at some point the instabilities reappear. Finally, we have developed novel algorithms for estimating multilevel linear models with spatial autocorrelation that have been implemented in our software. The formal properties of these procedures have also been derived.

With our new approach, we can now easily turn to multilevel generalized linear models with spatial autocorrelation. All of the pieces are now in place. It is important to emphasize again, however, that we have in important ways reformulated the manner in which the modeling is approached; we are no longer

seeking the right model but rather, a useful model (For a rich elaboration on this point, see Berk 2003, p. 206-218).

4.3 THE FORMAL STRUCTURE OF MULTILEVEL MODELS

We have built on several existing traditions in statistics. Spatial regression models (Anselin 1988) are heteroscedastic linear models with correlated disturbances, in which the covariance between the disturbances depends on the spatial distance of the sites. Random coefficient models (Longford 1993) are heteroscedastic linear models with correlated disturbances, in which the covariance between the disturbances depends on the predictor similarity of the sites. Multilevel models (Kreft and de Leeuw 1998) are random coefficient models in which the predictor similarity is determined by the fact that sites are grouped into clusters. Disturbances between clusters are uncorrelated, but within clusters the covariance depends on the predictor similarity of the sites. Since distance and similarity are closely related constructs, one would expect a relationship between these three classes of models.

Spatial regression models and random coefficient models both have correlated disturbances, with the size of the correlation depending on the similarity of the sites. Similarity can be defined spatially or, more generally, in terms of similarity of the sites on a number of predictors which may not be spatial. Multilevel models simplify the overall correlation structure by assuming that sites in different clusters are uncorrelated, which means that the covariance matrix of the sites is block-diagonal, and presumably sparse.

It should come as no surprise, then, that much of our work relies on the earlier work of many others. But, we make four new contributions as follows: (1) we combine autoregressive models with multilevel models; (2) we consider spatial effects both as functions of non-spatial covariates with random coefficients and as autocorrelated disturbances; (3) we usefully approximate autocorrelated disturbance structures by using spatial regressors with random coefficients; and (4) we develop augmentation and majorization methods to estimate generalized multilevel autoregressive models. These are iterative computational methods dealing with the nonlinearities in such models (De Leeuw 1994).

The first three contributions are summarized below. Our work on the fourth contribution is available upon request. We also have a single, broad "take-home message." The development of statistical tools for environmental applications and the use of those tools should forego the traditional search for the "correct model" and focus instead on building one or more "useful models."

4.3.1 Basics

We assume *multilevel data*. In the simplest case with two levels, the units of level one (which we call the *one-units*) are nested in units of level two (the *two-units*). Because of our concentration on spatial examples, we will often use the terminology of (research) "sites" and "transects" for the units in our levels. Transects are nested in sites. In the two-level case, we have m two-units, and within two-unit j, we have

n_j one-units. For each two-unit j, there is a vector z_j, of length p, of regressors describing the two-units. This implies that there will be p regression coefficients, excluding the intercept, for two-units. There are also (n_j x q) matrices X_j of regressors describing one-units. This implies that there will be q regression coefficients, excluding the intercept, for one-units. The total number of one-units in all m two-units is n.

We usually allow for an intercept in regression models. So, we add a column of n_j x 1 columns of 1's to X_j, with 1 as the lead element in z_j. Then, the standard two-level model assumes that within each two-unit j we have a *random-coefficient regression model* of the form

$$\underline{y}_{ij} = \sum_{s=0}^{q} x_{ijs}\underline{\beta}_{js} + \underline{\varepsilon}_{ij}. \tag{4.1}$$

Here, i is the index used for one-units (I = 1, 2, …, n_j), which are nested in the two-units. We follow conventional practice and assume that the disturbances $\underline{\varepsilon}_{ij}$ are uncorrelated with the predictors; there are no "omitted variables" at the one-unit level, and the functional forms are appropriate. In practice, these assumptions must be carefully examined and justified. Often they will be found wanting. Note that random variables are always underlined and that we use element-wise notation initially, but matrix notation further on.

The *q+1 random regression coefficients* $\underline{\beta}_{js}$ in Equation 4.1 express the relationship between the *first-level predictors* and the *outcomes*. These random coefficients, of which there are p+1 for each two-unit j, are themselves outcomes of a second regression model, with fixed regression coefficients, shown in the equation

$$\underline{\beta}_{js} = \sum_{r=0}^{p} z_{jr}\gamma_{rs} + \underline{\delta}_{js}, \tag{4.2}$$

in which the random regression coefficients are outcomes predicted by *second-level predictors*. Again following convention, we assume that the disturbances, $\underline{\delta}_{js}$, are uncorrelated with the predictors; there are no "omitted variables" at the two-unit level, and the functional forms are appropriate. Of course, both assumptions are likely to be substantially wrong in practice, which again underscores the need to focus on useful models, not correct models. More will be said about this later.

In the spatial case, the first level predictors describe properties of the transects. They can be spatial, in the sense that they are functions of the coordinates of the transects, or non-spatial. The second level predictors describe properties of the sites, and again they can be spatial or non-spatial.

One can substitute Equation 4.2 into Equation 4.1 to write the model as a single equation:

$$\begin{aligned} \underline{y}_{ij} &= \Sigma_{s=0}^{q} x_{ijs} \{\Sigma_{r=0}^{p} z_{jr}\gamma_{rs} + \underline{\delta}_{js}\} + \underline{\varepsilon}_{ij} \\ &= \Sigma_{s=0}^{q}\Sigma_{r=0}^{p} x_{ijs} z_{jr}\gamma_{rs} + \Sigma_{s=1}^{q} x_{ijs}\underline{\delta}_{js} + \underline{\varepsilon}_{ij}. \end{aligned} \tag{4.3}$$

Thus, we see that the fixed part for two-unit j has the form

$$E(\underline{y}_{ij}) = \sum_{r=0}^{p}\sum_{s=0}^{q} \gamma_{rs} z_{jr} x_{ijs} \tag{4.4}$$

with $(p+1)(q+1)$ fixed predictors, each a product of a first-level and a second-level variable, often called "interaction variables." The random part has the form

$$\underline{y}_{ij} - \mathrm{E}(\underline{y}_{ij}) = \sum_{s=0}^{q} x_{ijs}\underline{\delta}_{js} + \underline{\varepsilon}_{ij}. \tag{4.5}$$

We now need some additional assumptions on the distribution of the disturbance terms. Some very general ones are: $\mathrm{E}(\underline{\varepsilon}_{ij}) = 0$; $\mathrm{E}(\underline{\delta}_{js}) = 0$; $C(\underline{\varepsilon}_{ij}, \underline{\varepsilon}_{k\ell}) = 0$ if $j \neq \ell$; $C(\delta_{js}, \delta_{\ell t}) = 0$ if $j \neq \ell$; and $C(\varepsilon_{ij}, \delta_{\ell s}) = 0$.

Thus, first-level disturbances for different two-units are uncorrelated, and so are second level disturbances. The dispersion matrices of the first-level disturbances are

$$\mathrm{E}(\underline{\varepsilon}_j \underline{\varepsilon}'_j) = \sigma_j^2 \Lambda_j, \tag{4.6}$$

and those of the second-level disturbances are

$$E(\underline{\delta}_j \underline{\delta}'_j) = \sigma_j^2 \Omega_j. \tag{4.7}$$

The dispersion matrix, $\sigma_j{}^2\Lambda_j$, allows the one-unit disturbances $\underline{\varepsilon}_{ij}$ for a given two-unit to have different variances and to be correlated with one another. The dispersion matrix, $\sigma_j^2\Omega_j$, allows the disturbances δ_{js} for a given two-unit to have different variances and to be correlated with one another. The former is where spatial dependence not captured by the regressors is likely to be seen. The latter will reflect dependence between the random coefficients that is not spatial, but a result of chance processes not captured by the two-unit model.

As a practical matter, it will be impossible to estimate the values of Λ_j and Ω_j. These matrices contain weights that determine the disturbance variances and covariances and as such, there are far too many parameters to estimate. Often to simplify we suppose that Ω_j are the same for all two-units, and usually σ_j^2 are supposed to be the same too. Still, in most cases (see the examples below), Ω_j and

Λ_j are assumed to depend on a small number of parameters θ, which may again be constant over two-units.

4.3.2 Example

A simple spatial example may help clarify the model. It is not intended to be realistic, but to illustrate some key concepts. The one-units are observation stations, and the two-units are one of three counties. We suppose that rainfall at station *i* in county *j* depends on altitude (*alt*) and distance from the ocean (*dfo*).

$$\underline{rain}_{ij} = \underline{\beta}_{0j} 1_{ij} + \underline{\beta}_{1j} alt_{ij} + \underline{\beta}_{2j} dfo_{ij} + \underline{\varepsilon}_{ij}, \tag{4.8}$$

where 1_{ij} is the intercept, which is equal to one for all one-units. We do not assume that the regression coefficients are the same for all three counties. In fact, they vary according to a second regression model, for which we use indicator variables coding for the counties in the study. Thus, for s = 0, 1, 2,

$$\underline{\beta}_{js} = \gamma_{0s} 1_j + \gamma_{1s} LA_j + \gamma_{2s} SB_j + \underline{\delta}_{js}, \tag{4.9}$$

where again 1_{ij} is the intercept, now equal to one for all two-units. All observation stations in Los Angeles County (*LA*) have the same random coefficient distribution, and so do the observation stations in San Bernadino County (*SB*) and those in neither Los Angeles nor San Bernadino County.

If one substitutes the equations at the county level into the equations at the station level, for $i \neq k$ and assuming for notational simplicity that σ_j^2 and Ω_j are the same for all two-units,

$$C(\underline{rain}_{ij}, \underline{rain}_{kj}) = \sigma^2 [1 \quad alt_{ij} \quad dfo_{ij}] \begin{bmatrix} \omega_{00} & \omega_{01} & \omega_{02} \\ \omega_{10} & \omega_{11} & \omega_{12} \\ \omega_{20} & \omega_{21} & \omega_{22} \end{bmatrix} \begin{bmatrix} 1 \\ alt_{kj} \\ dfo_{kj} \end{bmatrix} \tag{4.10}$$

Thus, the covariance between the one-units in the same two-unit is determined by the similarity of predictor values of the one-units, where similarity is measured by their inner product in the matrix Ω. This is a key insight, which shows why estimation of the parameters in spatial multilevel models can be difficult when one believes that certain sets of disturbances are correlated as well.

4.3.3 Matrix Notation

Define the matrix Z_j as the direct sum of q copies of the row vector z_j^t. Thus, it is q by qp, and it looks like

$$Z_j = \begin{bmatrix} z'_j & 0 & 0 & \dots & 0 \\ 0 & z'_j & 0 & \dots & 0 \\ 0 & 0 & z'_j & \dots & 0 \\ 0 & 0 & 0 & z'_j & 0 \\ 0 & 0 & 0 & 0 & z'_j \end{bmatrix} \tag{4.11}$$

Using this matrix and stacking the γ_{rs} in a single vector γ, we can rewrite Equation 4.2 as

$$\underline{\beta}_j = Z_j \gamma + \underline{\delta}_j, \tag{4.12}$$

If we substitute Equation 4.12 into Equation 4.1, we find

$$\underline{y}_j = U_j \gamma + X_j \underline{\delta}_j + \underline{\varepsilon}_j, \tag{4.13}$$

$$E(\underline{y}_j) = U_j \gamma, \quad \text{with } U_j \triangleq X_j Z_j, \tag{4.14a}$$

$$V(\underline{y}) = \sigma_j^2 (X_j \Omega_j X'_j + \Lambda_j). \tag{4.14b}$$

It is convenient to write $\sum_j$ for $X_j \Omega_j X'_j + \Lambda_j$. Now U_j is of the form

$$U_j = [x_{j1} z'_j \,|\, \dots \,|\, x_{jq} z'_j], \tag{4.15}$$

where x_{jr} is column r of X_j. Thus, in Equation 4.14a, the predictors in U_j are products of a first-level predictor from X and a second-level predictor from Z. In principle, all these *cross-level interactions* are part of the model, but we can eliminate some of them by setting the corresponding element of γ equal to zero. Also observe that often the first column of both the X_j and of Z is an *intercept* column with all elements equal to +1. If we form all cross-level interactions, this implies that the columns of X and Z themselves also occur as predictors, because they are the intersections with the intercept at the other level.

4.3.4 Generalizations

4.3.4.1 More than two level

In a more-than-two-level model, there are one-units, two-units, and so on, nested within each other. For instance, we can have transects nested within streams nested within watersheds. For this case we can adopt a more general notation. Suppose we have n_r observations on level r, and q_r predictors on that level. Thus, we have

$n_r \times (q_r + 1)$ matrices $X^{(r)}$ with predictors. We also use indicator matrices $G^{(r)}$, which are $n_r \times n_{r+1}$, and which indicate how the r-units map into the $(r+1)$-units.

The first two equations defining our multilevel model are

$$\underline{y}^{(1)}_{i_1} = \sum_{s_1=0}^{q_2} x^{(1)}_{i_1 s_1} \sum_{i_2=1}^{n_2} g^{(1)}_{i_1 i_2} \underline{y}^{(2)}_{i_2 s_1} + \underline{\varepsilon}^{(1)}_{i_1}, \tag{4.16a}$$

$$\underline{y}^{(2)}_{i_2 s_1} = \sum_{s_2=0}^{q_2} x^{(2)}_{i_2 s_2} \sum_{i_3=1}^{n_3} g^{(2)}_{i_2 i_3} \underline{y}^{(3)}_{i_3 s_1 s_2} + \underline{\varepsilon}^{(s)}_{i_2 s_1}. \tag{4.16b}$$

Thus, we have n_1 random variables in $\underline{y}^{(1)}$. These are the observed outcomes. We have $n_2 \times q_1$ unobserved random variables in $\underline{y}^{(2)}$, these are the random regression coefficients from our previous formulation. Then we have $n_3 \times q_1 \times q_2$ unobserved random coefficients in $\underline{y}^{(3)}$, and so on.

In the same way as before, we can combine equations to form single equations, which rapidly become unwieldy. For both mathematical reasons and ease of interpretation, it is wise to work with the fewest levels that can be justified. In practice, complex models also can have very unstable results and often will not converge at all. It is far better to have a model that is too simple than a model than is too complex. From Equation 4.16 we find, for example,

$$\underline{y}^{(1)}_{i_1} = \sum_{s_1=0}^{q_1} x^{(1)}_{i_1 s_1} \sum_{i_2=1}^{n_2} g^{(1)}_{i_1 i_2} [\sum_{s_2=0}^{q_2} x^{(2)}_{i_2 s_1 s_2} \sum_{i_3=1}^{n_3} g^{(2)}_{i_2 i_3} y^{(3)}_{i_3 s_1 s_2} + \varepsilon^{(s)}_{i_2 s_1}] + \underline{\varepsilon}^{(1)}_{i_1}. \tag{4.17}$$

4.*3.4.2 Multivariate outcomes*

If there is more than one outcome variable, we can use a simple trick to force the model into the multilevel framework. We use *variables* as the first level. Thus variables are nested in transects, transects in sites, and so on. For example, if there are three outcomes contained in three columns of the data set, one can reorganize the data so that within each one-unit there are three rows, one for each outcome. In each of these rows is the value for each of the three outcome variables respectively, with the values in the columns for predictor variables duplicated three times. Having multiple outcomes just adds a level to the hierarchy. In addition, missing data on the outcomes can be incorporated without difficulty, because some transects simply have fewer units (i.e., variables) than others.

We suspect that because there are often several ecologically interesting response variables for a given analysis, this approach to multiple outcomes can be widely useful. That is, even if there is no need for multilevel models because of a particular

hierarchical structure, multilevel level models can be used when there is a need to consider more than one outcome at a time.

4.3.4.3 Non-independent two-units

In our models, we usually assume that Ω_j are the same for all sites. With this assumption, it is possible to use a simple model for correlated sites, which has

$$C(\underline{y}_j, \underline{y}_\ell) = \sigma_{j\ell}(X_j \Omega X'_\ell + \Lambda_j^{1/2} \Lambda_\ell^{1/2} \tag{4.18a}$$

for all $j \neq \ell$, and

$$C(\underline{y}_j, \underline{y}_j) = \sigma_{jj}(X_j \Omega X'_j + \Lambda_j) \tag{4.18b}$$

for all j, where $\sigma_{j\ell}$ are the covariances between sites.

4.3.4.4 Generalized MAR models

In the same way as linear models are generalized to generalized linear models, one can try to construct generalized mixed linear models from mixed linear models. The trick is simply to condition on the random effects. In generalized linear models, first-level observations are independent given the random effects, and thus, the conditional distribution is a simple product of univariate Poisson, binomials, or gammas. But in generalized mixed linear models with autocorrelated or spatially correlated first-level disturbances, one no longer can use independence, and there is a need to assume that the disturbances within sites have multivariate Poisson, binomial, or gamma distributions. There is no agreement in statistics about how to define such multivariate distributions, and the definitions that are popular do not have many of the simplifying properties of the univariate versions.

We shall see below, however, that models with correlated first-level disturbances can be approximated by models with additional random effects and uncorrelated first-level disturbances. In these approximations, conditioning on the random effects makes the observations independent again, and the results developed for generalized mixed linear models apply again. This is perhaps the key technical point of this paper. To see why this works, we need to consider in greater detail the nature of the disturbances in the models.

4.4 MODELS FOR DISTURBANCE DISPERSIONS

The dispersion matrices Λ_j of first-level disturbances can take many different forms. Generally, they are a function of a number of unknown parameters, collected in a vector ρ. Estimation simplifies considerably if the Λ_j are known, and in particular in the homoscedastic case with uncorrelated disturbances in which

$\Lambda_j = I_j$, the identity matrix of order n_j. But in spatial situations the assumption that the disturbances are uncorrelated often is difficult to defend.

This is why a great deal of attention has been paid to modeling the dependence of spatial observations, taking as the main inspiration the literature on time series models. The key paper in spatial autoregressive (SA) modeling is Ord (1975). Also compare Griffith (2002b) and Anselin (2001). There are various forms of these SA models, but the most important ones are one-parameter models, in which the single parameter ρ is interpreted as spatial autocorrelation. It indicates the strength of the spatial effects.

In multilevel models, restrictions are often placed on Ω_j. For instance, it is common to assume that they are equal or that specific elements are zero. We shall discuss these restrictions later, and concentrate here on the first-level disturbances.

4.4.1 The Spatial Lag Model

The spatial lag model is also known as the AR, or autoregressive response model. It specifies

$$\underline{y}_j = \rho_j W_j \underline{y}_j + X_j \underline{\beta}_j + \underline{\varepsilon}_j, \tag{4.19}$$

where $\underline{\varepsilon}_j$ is homoscedastic with variance σ_j^2. With $\underline{y}_j$ on both sides of the equal sign, this is an AR model with

$$E(\underline{y}_j \big| \beta_j) = (I_j - \rho_j W_j)^{-1} X_j \beta_j \tag{4,20a}$$

$$V(\underline{y}_j \big| \beta_j) = \sigma_j^2 [(I_j - \rho_j W_j)(I_j - \rho_j W_j')]^{-1}. \tag{4.20b}$$

In this formulation, the autoregression is defined directly in terms of the outcomes. The spatial dependence is built into the model in a structural manner. That is, the data analyst will typically have a subject-matter rationale for why and how values of the outcome variable are related. For example, if water quality in a lake is the outcome of interest, there may be diffusion of pollution from any one location to locations nearby. Depending on the value of ρ_j, the diffusion effects might be large or small, or perhaps even be negative. Note also that to isolate the role of the predictors, adjustments have to be made for the diffusion process, which links the outcome across locations. A failure to make such adjustments may mean that effects attributed to one or more of the predictors are really just a result of the movement of pollution from one place to another.

4.4.2 The Spatial Error Model

The spatial error model is also known as the SAR or simultaneous autoregressive model (Anselin 2001). It is expressed as

$$\underline{y}_j = X_j \underline{\beta}_j + \underline{\zeta}_j, \tag{4.21a}$$

and it assumes an autoregression structure for the error terms. Thus

$$\underline{\zeta}_j = \rho_j W_j \underline{\zeta}_j + \underline{\varepsilon}_j, \tag{4.21b}$$

where the $\underline{\zeta}_j$ are homoscedastic with variance σ_j^2. This leads to

$$E(\underline{y}_j \big| \beta_j) = X_j \beta_j, \tag{4.22a}$$

$$V(\underline{y}_j \big| \beta_j) = \sigma_j^2 [(I_j - \rho_j W_j)(I_j - \rho_j W_j')]^{-1}. \tag{4.22b}$$

This formulation implies that the spatial dependence is not potentially confounded with the predictors. It derives solely from dependence among the disturbances themselves. Disturbances that are more proximate in space, for instance, may tend to be more alike than disturbances that are farther apart. The reasons for the dependence are usually not of much interest. As such, the dependence is a mere nuisance and/or beyond current subject matter interest. The goal is to "mop up" the spatial dependence in the disturbances so that it does not affect the precision of estimates of the β_j or estimates of their standard errors.

4.4.3 The Conditional Autoregression Model

Under the conditional autoregression model (CAR), also discussed in Anselin (2001), we let

$$\underline{y}_j = X_j \underline{\beta}_j + (I_j - \rho_j W_j)^{-1/2} \underline{\varepsilon}_j, \tag{4.23}$$

where W_j is now a symmetric weight matrix, and $\underline{\varepsilon}_j$ is homoscedastic with variance σ_j^2. This implies

$$E(\underline{y}_j \big| \beta_j) = X_j \beta_j, \tag{4.24a}$$

$$V(\underline{y}_j \big| \beta_j) = \sigma_j^2 [(I_j - \rho_j W_j)]^{-1}. \tag{4.24b}$$

If dependence in the disturbances can be treated as a mere nuisance, the model that one uses for the disturbances is of little importance as long as the dependence is taken into account when the regression coefficients are estimated. In this context, the conditional autoregressive model can be seen as an alternative to the spatial error model, and it has some of the same look and feel. Larger values of ρ_j imply more dependence among the disturbances. And just as for the spatial error model, the dependence may be a function of distance; closer disturbances may tend to be more alike. The main advantage of CAR model is that it can be as effective in mopping up dependence in the disturbances as the spatial errors model, but will be far easier to compute.

4.4.4 Weight Matrices

How to choose W_j has been discussed many times in the geostatistics literature. A good review is Bavaud (1998; see also Cressie 1991). Although it is possible to give some general indications, choosing a precise and appropriate W_j is difficult, probably even more difficult than choosing a correct set of predictors. The usual problem is that there is too little a priori knowledge to inform the choice and at best some general clues in the data.

4.4.4.1 Choice of weights

For W_j in spatial situations, we assume that its elements are similarities of transects in site j. The more similar (the closer) the transects, the larger the corresponding element in W_j. If we do not have a good reason to choose a specific W_j, we can make it some (decreasing) function of the transect distances, but again choosing the function is often disturbingly arbitrary. In many cases, moreover, we even want to replace simple Euclidean distance by other distances (measured along a network or stream, for instance), which take the actual spatial setting into account. Throughout, we suppose the elements of W_j are non-negative.

4.4.4.2 Large matrices

In spatial analysis we often encounter situations in which the order of W_j is very large, maybe 10^5 or 10^6. Obviously in such cases, it will generally not be possible to store floating-point matrices of this size, let alone compute their determinants, inverse, or eigen-decomposition.

There are several ways around this problem. The first is to use patterned weight matrices of zeroes and ones (coding adjacency or nearest neighbor, for instance), with a determinant or an inverse available in analytical form (Pace and Zou 2000). The second is to use sparse matrix techniques for weight matrices with a very large proportion of zeroes (Pace and Barry 1997a, 1997b, 1997c) (again, adjacency matrices come to mind). We have also seen that multilevel analysis suggests partitioning transects or sites into clusters, and making the between cluster covariance equal to zero. This also introduces a great deal of sparseness. And finally, fast

numerical approximations to the loss function are also a possibility. Specifically, techniques for approximating the determinant in the normal log-likelihood for all AR, SAR, and CAR models are in Smirnov and Anselin (2001) and Griffith (2002a).

In the models discussed in this paper, we have the additional complication that the dispersion matrix is made up of two components: a part based on similarity of the regressors and a part based on spatial information, coded in the weight matrices. This makes patterned weight matrix and sparse matrix techniques more difficult to use, and we have to resort to other types of approximations.

4.4.4.3 Normalizing the weights

It is computationally convenient if the weight matrices in the SAR and AR models are symmetric. Then, we get a more simple formulation,

$$(I_j - \rho_j W_j)(I_j - \rho_j W_j') = (I_j - \rho_j W_j)^2,$$

which is easier to work with. Unfortunately, in many applications an asymmetric set of weights may make more sense (think of the influence of stream flow or hillside slope on ecological distance, for instance).

Consider why having symmetric matrices is convenient. If the W_j are known symmetric matrices, one can compute the spectral decomposition, $W_j = K_j \Phi_j K_j'$, and we find

$$\Lambda_j(\rho_j) = \sum_s \frac{1}{(1-\rho_j \phi_{js})^2} k_{js} k_{js}' \text{ for SAR} \tag{4.25a}$$

$$\text{and } \Lambda_j(\rho_j) = \sum_s \frac{1}{1-\rho_j \phi_{js}} k_{js} k_{js}' \text{ for CAR.} \tag{4.25b}$$

Thus, the eigenvectors of $\Lambda_j(\rho_j)$ are the same as those of W_j, and the eigenvalues are simple functions of the eigenvalues of W_j. If ρ_j changes, only the eigenvalues change; the eigenvectors remain the same.

For interpretation purposes, one can normalize the weights in such a way that the rows of W_j sum to unity. This makes the weight matrix stochastic, and by Frobenius theorem implies that the largest eigenvalue of W_j is equal to +1. This means that the smallest eigenvalue of $I_j - \rho_j W_j$ is $1 - \rho_j$, and thus $I_j - \rho_j W_j$ is positive definite as long as $\rho_j < 1$, which helps in the interpretation of ρ as a type of autocorrelation coefficient.

In some cases, it is desirable for W_j to be both symmetric and normalized (i.e., doubly stochastic). This is discussed for CAR models in Page and LeSage (2002). We have developed an algorithm and computer codes to normalize non-negative

symmetric matrices in such a way that they become doubly stochastic (the codes available upon request).

4.4.5 Special Case: Time Series Models

If the outcomes are one-dimensional (e.g., if transects are arranged in lines), then it makes sense to use a time series model for the first-level disturbances (Hedeker 1989, Hedeker and Gibbons 1996). We discuss these models here briefly because they show where the SA models come from, and because they are more likely to be familiar.

A first obvious choice for a time-series model is the *random walk*, which has

$$\underline{\varepsilon}_j = W_j \underline{\varepsilon}_j + \underline{\zeta}_j, \tag{4.26}$$

where W_j has all elements equal to zero, except for those immediately below the main diagonal, which are one. It follows that

$$\underline{\varepsilon}_j = T_j \underline{\zeta}_j, \tag{4.27}$$

where T_j has all elements on and below the main diagonal equal to one and all elements above the main diagonal equal to zero. Thus,

$$\Lambda_j = T_j T_j', \tag{4.28}$$

which means that element (s, t) is equal to *min(s, t)*.

In an AR(p) process,

$$\underline{\varepsilon}_j = W_j \underline{\varepsilon}_j + \underline{\zeta}_j, \tag{4.29}$$

where W_j has a band of width p below the diagonal and zeroes elsewhere. There are p parameters, the autoregression coefficients, in W_j. The AR(1) model is much like the random walk, except that the element below the diagonal is the single parameter ρ_j.

An MA(q) process also uses a banded matrix with parameter values, but now

$$\underline{\varepsilon}_j = W_j \underline{\zeta}_j, \tag{4.30}$$

where W_j has diagonal one, and a band of width q in each row below the diagonal. Thus, an MA(1) has diagonal one, and ρ_j below the diagonal.

It is easy to extend this to ARMA(p, q) and even more complicated processes, but this is comparatively straightforward and it may be overkill in many practical situations. For our purposes, the most interesting models are AR(1) and MA(1),

which can be defined in term of the backshift matrix B_j, which has elements equal to one below the diagonal only. Then for AR(1), we have

$$\Lambda_j(\rho_j) = (I_j + \rho_j B_j)^{-1}(I_j + \rho_j B_j')^{-1}, \tag{4.31a}$$

and for MA(1) we have

$$\Lambda_j(\rho_j) = (I_j + \rho_j B_j)(I_j + \rho_j B_j'). \tag{4.31b}$$

The random walk is AR(1) with $\rho_j = 1$.

4.5 MODEL APPROXIMATION

Now we turn to approximations of various AR models that can lead to practical computational results. Because, as noted earlier, the concept of "the true model" is at least obscure and because even if we know how to think about "the true model," we usually do not have very precise information about which W_j produces it, it makes sense to employ an approximation of the dispersion matrix that is computationally convenient. We will first simplify the model by an approximation that works well for small ρ_j, and then we approximate the model by another formulation with homoscedastic first-level disturbances (i.e., a model with $\Lambda_j = I_j$).

4.5.1 Simplified AR

Consider again the SAR model described in the section of the spatial error model above. Recall that the variance-covariance matrix of the disturbances was $\sigma_j^2[(I_j - \rho_j W_j)(I_j - \rho_j W_j)]^{-1}$, where all of the terms to the right of σ_j^2 represent $\Lambda_j(\theta)$. In the *Simplified* AR Model (SIMAR), assume

$$\Lambda_j(\theta) = I_j + \rho_j W_j, \tag{4.32}$$

where the off-diagonal elements of the symmetric matrix W_j are again some decreasing function of the Euclidean distances between the transects or, more generally, of the spatial dissimilarities.

In the CAR model, if ρ_j is small,

$$(I_j - \rho_j W_j')^{-1} = I_j + \rho_j W_j + o(\rho_j), \tag{4.33}$$

and in the SAR and CAR models,

$$\begin{aligned}\Lambda_j(\rho_j) &= (I_j - \rho_j W_j)^{-1}(I_j - \rho_j W_j')^{-1} \\ &= I_j + \rho_j(W_j + W_j') + o(\rho_j),\end{aligned} \tag{4.34}$$

which are both of the SIMAR form.

For both AR(1) and MA(1), and small ρ_j,

$$\Lambda_j = I_j + \rho_j(B_j + B_j') + o(\rho_j), \tag{4.35}$$

which is again of the required SIMAR form.

4.5.2 Spatial Effects As Random Coefficients

By using random coefficients in appropriate ways, one can emulate the covariance structure of the SIMAR without assuming correlated disturbances for the first-level units. Thus, one can maintain $\Lambda_j = I_j$. The trick is really quite simple. In our spatial multilevel models

$$\underline{y}_j = U_j\gamma + X_j\underline{\delta}_j + \underline{\varepsilon}_j, \tag{4.36}$$

$$V(\underline{\varepsilon}_j) = \sigma_j^2(I_j + \rho_j W_j). \tag{4.37}$$

Now suppose $W_j = K_j\Phi_j K_j'$ is the spectral decomposition of W_j. Then,

$$\underline{y}_j = U_j\gamma + X_j\underline{\delta}_j + K_j\underline{\eta}_j + \underline{\zeta}_j, \tag{4.38}$$

where $\underline{\delta}_j$ and $\underline{\eta}_j$ are uncorrelated, and where

$$V(\underline{\eta}_j) = \sigma_j^2\rho_j\Phi_j, \tag{4.39a}$$

$$V(\underline{\zeta}_j) = \sigma_j^2 I_j. \tag{4.39b}$$

But, Equations 4.38 and 4.39 can be interpreted as a simple multilevel model in which the covariance matrix of the random effects is of the form

$$\begin{bmatrix}\Omega & 0 \\ 0 & \rho_j\Phi_j\end{bmatrix}. \tag{4.40}$$

First-level disturbances are homoscedastic, and the regression coefficients corresponding with the eigenvector-predictors K_j only have a random part and a

vanishing fixed part. Moreover, the random parts are uncorrelated, with a diagonal dispersion matrix proportional to the eigenvalues of W_j. In short, one can write the SIMAR model as a multilevel model with restrictions on the covariance matrix of the random effects.

4.5.3 Positive Definite Variances

One problem with this formulation is that it is not guaranteed that the eigenvalues Φ_j of W_j are non-negative. If there are negative eigenvalues, then Equation 4.39a is difficult to interpret. One can use the fact, however, that $I_j + \rho_j W_j$ must be positive definite. Suppose $\rho_j > 0$, and write y_j for the smallest eigenvalue of W_j. Then,

$$I_j + \rho_j W_j = (1 + \rho_j \psi_j) I_j + \rho_j K_j (\Phi_j - \psi_j I_j) K_j', \tag{4.41}$$

and we can rewrite Equation 4.39 as

$$V(\underline{\eta}_j) = \sigma_j^2 \rho_j (\Phi_j - \psi_j I_j), \tag{4.42a}$$

$$V(\underline{\zeta}_j) = \sigma_j^2 (1 + \rho_j \psi_j) I_j. \tag{4.42b}$$

These are somewhat more complicated restrictions, but they always give positive semidefinite dispersion matrices.

4.5.4 Using Fewer Eigenvalues

A second problem with our approximation is that we replace a very large spatial disturbance covariance matrix with a very large number of random effects. The number of random effects added is equal to the order of the spatial covariance matrix. We attack this problem by using only a small number of eigenvectors of W_j, those corresponding with the largest eigenvalues (in modulus). Thus, we use a principal component type approximation to the random effects. With spatial information in W_j using some function of the distances, two or three principal components are likely to give a rather good approximation.

4.5.5 General Approach

Instead of approximating the SA models by SIMAR, and then approximating SIMAR by using eigenvectors, one can employ a more straightforward approach that can reduce the computational burdens. Consider the following multilevel model for site *j*

$$\underline{y}_j = X_j \underline{\beta}_j + Z_j \eta_j + \underline{\varepsilon}_j, \tag{4.43}$$

where X_j contains regression coordinates, and Z_j contains functions of the spatial coordinates. For our second level model, we use

$$\underline{\beta}_j = A_j\gamma + \underline{\delta}_j, \tag{4.44a}$$

$$\underline{\eta}_j = B_j\kappa + \underline{\xi}_j. \tag{4.44b}$$

This implies

$$\underline{y}_j = X_jA_j\gamma + Z_jB_j\kappa + \underline{\nu}_j, \tag{4.45a}$$

where

$$\underline{\nu}_j = X_j\underline{\delta}_j + Z_j\underline{\xi}_j + \underline{\varepsilon}_j, \tag{4.45b}$$

and thus, with suitable uncorrelatedness assumptions,

$$E(\underline{y}_j) = X_jA_j\gamma + Z_jB_j\kappa, \tag{4.46a}$$

$$V(\underline{y}_j) = \sigma_j^2(X_j\Omega_jX_j' + Z_j\Theta_jZ_j' + I_j). \tag{4.46b}$$

This becomes an approximate multilevel Ord model if we let $B_j = 0$ (i.e., the spatial regression coefficients do not have a fixed part), and we let $\Theta = \rho_j^2 I_j$ (i.e., the spatial regression coefficients are uncorrelated). Then,

$$E(\underline{y}_j) = X_jA_j\gamma, \tag{4.47a}$$

$$V(\underline{y}_j) = \sigma_j^2[X_j\Omega_jX_j' + (I_j + \rho_j^2Z_jZ_j')]. \tag{4.47b}$$

Moreover, to get closer to SA, one can choose Z_j in clever ways, using the results that we discussed earlier in this section. If the W_j matrix in the Ord model is a function of the spatial distances, then it obviously is a function of the coordinates, and thus all its eigenvectors are functions of the coordinates. If we choose Z_j as a low-rank (principal component) approximation of W_j, using the eigenvectors, then we can get very close to the Ord model.

With the practical approximation of the Ord Model, applications to the generalized linear model follow easily and directly. Work on these extensions is nearly completed, and software development has begun.

4.6 EMPIRICAL EXAMPLE

We have data from coral reefs along Olango Island in the Phillipines (unpublished data provided by Craig Schuman and Domingo Ochavillo). There are 33 sites with four transects in each. There are 14 sites in areas that are protected; fishing is prohibited. There are 19 sites that are in unprotected areas; fishing is allowed (and is common). And the fishing can include very destructive practices such as poisoning fish. The data that we analyze is an aggregate over four equally spaced observations along each transect. Thus, transects are our lowest level, and the second level is sites in the multilevel spatial model.

To keep the example simple, we use the same formulation illustrated in the example section above. The main difference is that the specific response is the number of different fish species. Therefore, the results reported are for a Poisson response variable within our multilevel framework. At the lowest level, the number of fish species is a function of how sandy the bottom is. The larger the percent of the bottom that is sandy, the fewer species one would expect. This relation depends on the intercept (β_0) and the slope of regression (β_1) at the level of the site. Then, the intercept is a function of whether the reef is protected via γ_0. Note that a reef is coded 1 if the reef is protected and 0 otherwise. The slope is also a function of whether the reef is protected via γ_1. Thus, both of the parameters at the level of the site are treated as random coefficients with a structural component determined by whether the reef is protected. In Table 4.1, we used an augmentation algorithm, but the results are much the same for any of the simplifications that we have discussed.

Table 4.1. *Model for species counts estimated by augmentation algorithm (N = 132).*

Predictor	*Coefficient*	*Standard Error*
Protected (γ_0)	5.58	3.75
% Sandy Bottom (β_1)	-0.18	0.04
Protect *X* % Sandy (γ_1)	0.05	0.08
Constant (β_0)	27.6	2.10
θ (AR parameter)	0.44	–

Focusing first on the regression coefficients from Table 4.1, one can see that if a reef is unprotected there are on average nearly 28 distinct fish species at a site. At these unprotected sites, for each additional percent of the bottom that is sandy, the number of species drops by 0.18; for every additional 10%, the number of species drops by 1.8. In the protected sites, the number of fish species is greater by 5.58. Finally, in the protected sites, the negative impact of a sandy bottom on the number of species is a bit less pronounced. The regression coefficient of -0.18 is now -0.13. For every 10% increase in sandy bottom, the number of species is reduced by 1.3.

The autoregressive parameter is 0.44, which is of moderate size. There is some meaningful spatial autocorrelation in the residuals. When this is taken into account, we see in Table 4.1 that the percent of the bottom that is sandy is easily twice the

standard error. The impact of protecting a reef is about 1.5 times its standard error, statistically significant at the 0.10 level for a one-tailed test. The coefficient for the interaction effect is less than its standard error. One should treat any formal tests with great caution, in part because the data were not collected by random sampling, and there is no compelling model-based sampling alternative. Also, it is virtually certain that important explanatory variables have been overlooked. But if one chooses to take formal tests seriously, the interaction effect can be discarded.

4.7 SOFTWARE

The results shown in Table 4.1 were produced by developmental software that we wrote in R. But, one can obtain consistent estimates of all the regression parameters using any software for Poisson regression, usually as a special case of the generalized linear model. One just has to substitute the higher level equations into the lower level equations, simplify, and proceed as usual. And if confidence intervals and tests are not formally justified (which is often the case), one does not have to proceed any farther. Moreover, all of the conventional regression diagnostics for the generalized linear model apply (Cook and Weisberg 1999).

Getting the uncertainty right is more difficult and, at this point, requires special software. We hope to have ours available soon. In the meantime, there are two good options that can be employed with existing software. First, if one can justify the spatial lag model discussed above, then once the substitution of the two-unit model into the one-unit model is completed, the GLM version of Equation 4.19 (e.g., logistic regression or Poisson regression) can be estimated in all of the major statistical packages with their routine GLM procedures. Consistent parameter estimation follows. Second, and far more generally, there exists, at least in SAS PROC MIXED (Littell et al. 1996), MLwiN (Goldstein 1995), GLLAMM (Rabe-Hesketh et al. 2002), and HLM (Raudenbush and Bryk 2002), the ability to do mixed effects generalized linear models. For a comparison of these packages we refer to De Leeuw and Kreft (2001). In each of them one can include functions of the spatial coordinates for the one-units but with their regression coefficients constrained to be equal to zero. We have had good success including just the horizontal coordinate, the vertical coordinate (e.g., longitude and latitude), and their product as one-unit predictors. Most of the spatial autocorrelation in the model will likely be "soaked-up." The standard errors should then be sufficiently accurate for most purposes.

4.8 CONCLUSIONS

In this paper, we discussed tools for the construction of multilevel linear models with ecological data. Extensions to multilevel generalized linear models followed directly. With these tools, one can examine how variables at one level are related to processes at another level; one can study the interactions between phenomena at different spatial/temporal scales. If one can also make the case that the structure of a model is very nearly right, and one has either random sampling or credible-model

based sampling (Berk, 2003), conventional ways of representing uncertainty apply. Our suggestions for obtaining useful estimates of the standard errors are then appropriate. However, we favor a more realistic approach in which description is the primary goal.

ACKNOWLEDGEMENTS

This work was supported by a grant from the EPA STAR program.

REFERENCES

Anselin, L. 1988. Model validation in spatial econometrics: a review and evaluation of alternative approaches. International Regional Science Review 11:279-316.

Anselin, L. 2001. Spatial Regression. http://geog55.gis.uiuc.edu/~luc/talks/spreg.pdf.

Bavaud, F. 1998. Models for spatial weights: a systematic look. Geographical Analysis 30:153-171.

Berk, R. 2003. Regression Analysis: A Constructive Critique. Sage Publications, Newbury Park, CA.

Cook, R. D., and S. Weisberg. 1999. Applied Regression Including Computing and Graphics. Wiley.

Cressie, N. 1991. Statistics for Spatial Data. Wiley.

de Leeuw, J. 1994. Block relaxation methods in statistics. Pages 275-289 in H. H. Bock, W. Lenski, and M. M. Richter, editors. Information Systems and Data Analysis. Springer, New York.

de Leeuw, J., and I. G. G. Kreft. 2001. Software for multilevel analysis. Pages 187-204 in A. H. Leyland and H. Goldstein, editors. Multilevel Modelling of Health Statistics. Wiley, New York.

Goldstein, H. 1995. Multilevel Statistical Medels. Arnold.

Griffith, D. A. 2002a. Quick but not so Dirty. ML Estimation of Spatial Autoregressive Models. Technical Report. Department of Geography, Syracuse University, Syracuse, NY.

Griffith, D. A. 2002b. Spatial Autoregression. Syracuse University, Syracuse, NY.

Hedeker, D. R. 1989. Random Regression Models with Autocorrelated Errors: Investigating Drug Plasma Levels and Clinical Response. Ph.D. Thesis. University of Chicago at Illinois, Chicago, IL.

Hedeker, D. R., and R. D. Gibbons. 1996. MIXREG: a computer program for mixed-effects regression analysis with autocorrelated errors. Computer Methods and Programs in Biomedicine 49:229-252.

Kreft, I., and J. de Leeuw. 1998. Introducing Multilevel Modeling. Sage Publications, Newbury Park, CA.

Littell, R. C., G. A. Milliken, W. W. Stroup, and R. D. Wolfinger. 1996. SAS System for Mixed Models. SAS Institute, Cary, NC.

Longford, N. 1993. Random Coefficient Models. Oxford University Press, Oxford.

Ord, J. K. 1975. Estimation methods for models of spatial interactions. Journal of the American Statistical Association 70:120-127.

Pace, R. K., and R. Barry. 1997a. Fast CARs. Journal of Statistical Compution and Simulation 59:123-147.

Pace, R. K., and R. Barry. 1997b. Quick computation of spatial autoregressive estimators. Geographical Analysis 29:232-246.

Pace, R. K., and R. Barry. 1997c. Sparse spatial autoregressions. Statistics and Probability Letters 33:191-197.

Pace, R. K., and D. Zou. 2000. Closed-form maximum likelihood estimates of nearest neighbor spatial dependence. Geographic Analysis 32:154-172.

Page, R. K., and J. P. LeSage. 2002. Conditional Autoregressions with Doubly Stochastic Weight Matrices (manuscript).

Rabe-Hesketh, S., A. Skrondal, and A. Pickles. 2002. Reliable estimation of generalized linear mixed models using adaptive quadrature. The Stata Journal 2:1-21.

Raudenbush, S. W., and A. S. Bryk. 2002. Hierarchical Linear Models, second edition, second edition. Sage Publications, Newbury Park, CA.

Smirnov, O., and A. Anselin. 2001. Fast maximum likelihood estimation of very large spatial auto-regressive models. Computational Statistics and Data Analysis 35:301-319.

Wu, J. 1999. Hierarchy and scaling: extrapolating information along a scaling ladder. Canadian Journal of Remote Sensing 25:367-380.

CHAPTER 5

DOWNSCALING ABUNDANCE FROM THE DISTRIBUTION OF SPECIES:

Occupancy Theory and Applications

FANGLIANG HE AND WILLIAM REED

5.1 INTRODUCTION

One of the most important contributions to our understanding of how and why species distribute in landscapes is to document the significant correlation between abundance and distribution of species across a broad range of scales (Brown 1984, 1995, Gaston and Blackburn 2000). The correlation suggests that there is a general tendency that locally abundant species are more widely distributed in space than rare species, which forms a positive distribution-abundance (or occupancy-abundance) relationship. While the observed relationship of this macroecological pattern begs for ultimate biological accounts (Brown 1984, Hanski et al. 1993, Gaston 1994, Kolasa and Drake 1998, Gaston and Blackburn 2000), the mathematical forms of the relationship derived from physical, statistical and geometrical considerations have greatly advanced the study on the topics and have indeed provided a solid ground for fermenting biological explanation further (Maurer 1990, Wright 1991, Hanski et al. 1993, Leitner and Rosenzweig 1997, Hartley 1998, Kunin 1998, He and Gaston 2000, Kunin et al. 2000, Harte et al. 2001, He et al. 2002; see Holt et al. 2002 for a review). An important implication of the distribution-abundance correlation is to allow for the derivation of species abundance from information on species distribution, a downscaling process (Wu and Li, Chapters 1 and 2). Here we will follow this premise to derive abundance by examining the spatial distribution of species in landscapes based on the combinatorial theory of occupancy.

The combinatorial theory of occupancy can date back as far as Pierre Laplace (Barton and David 1962) and has a long application in physics (Feller 1967). Laplace's classical example of occupancy considers the following birth game. Assume that there are N births taking place within a year and that each birth has the same chance to occur in any of the 365 days. What Laplace wanted to know was how many days out of the 365 would have no births, i.e., the number of empty days.

J. Wu, K.B. Jones, H. Li, and O.L. Loucks (eds.),
Scaling and Uncertainty Analysis in Ecology: Methods and Applications, 89–108.

Similarly, in statistical mechanics physicists are interested in knowing how N particles occupy a space composed of M small cells. The most well-known models that describe the number of empty cells (without particles) include Maxwell-Botltzmann and Bose-Einstein models. In this chapter, however, we wanted to know the reverse: not how many cells are empty, but how many particles are there provided that the number of empty cells is known. Specifically, let's consider a real example illustrated in Figure 5.1a in which a 50 ha plot in a rain forest of Malaysia is evenly divided into 800 cells of 25 × 25 m each (Figure 5.1b). The distribution (or

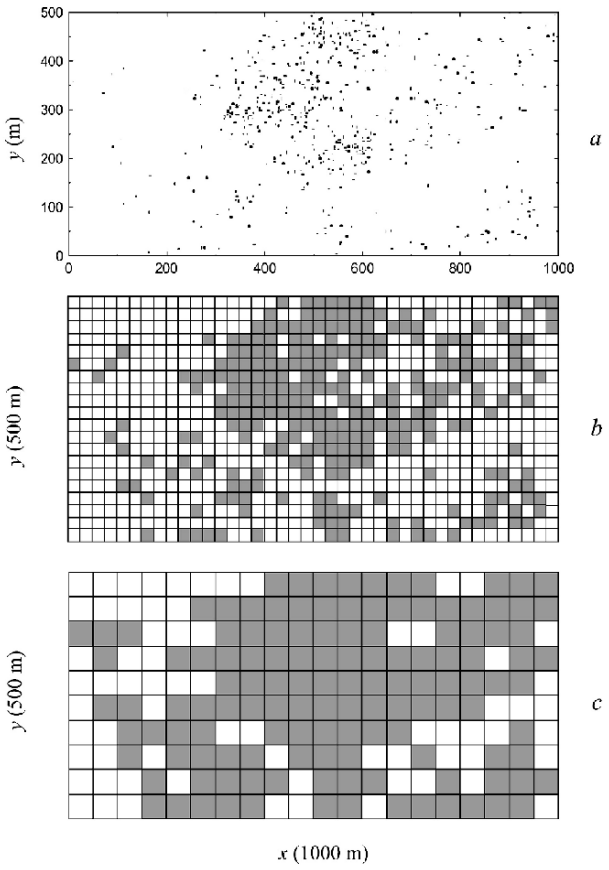


Figure 5.1. *Example of distribution of canopy tree species Dacryodes rubiginosa in a 500 × 1000 m tropical rain forest plot of Malaysia. (a) The actual distribution of 591 stems of the species in the plot. (b) The lattice representation of the species distribution with a map resolution of 25 × 25 m. The area of occupancy A_{a1} by the species is 171875 m^2. (c) The coarse-scale lattice map generated by aggregating four adjacent cells in (b) with a map resolution of 50 × 50m. The area of occupancy A_{a2} by the species is 325000 m^2.*

occurrence map, binary map, or atlas) is so generated that a cell is grey if the species is present and white if it is absent. Thus, a grey cell has at least one tree, but can have many more. Given such a map, we want to find out how many trees there are; of course, for Figure 5.1 we already know the number of trees and their locations in the plot. Note that real distribution maps are usually not as regularly bordered as Figure 5.1, but, for simplicity, statistical derivations dealt with in this study will be based on a map with assumed regular borders. It will become clear later that the models so derived are equally applicable to irregular maps.

In its mathematical form, a distribution can be defined as

$$\boldsymbol{x} = (x_1, x_2, ..., x_M), \tag{5.1}$$

where the subscript $(1, 2, \ldots, M)$ is a (spatial) location index for the M cells, x_i is represented by either 0 or 1, depending on the absence or occurrence of the species in the cell. The vector $\boldsymbol{x}$ can be a random or systematic sample from a study area, or an exhaustive sample (census) that covers entire area of interest as illustrated in Figure 5.1b. Although random or systematic sampling is important, this study concentrates on exhaustive sampling.

In Equation 5.1, when $x_i = 1$, we know for sure that there is at least one individual occurring in that cell. Therefore, for an observation $\boldsymbol{x}$, we know that there are at least $\sum x_i$ individuals occurring in the M cells. But how many are actually there? This chapter was designed to answer this question. The reminder of the chapter is organized into three sections:

(1) We start from a classical occupancy model to derive an abundance estimate by assuming that the individuals of a species are randomly and independently placed in space. The classical occupancy estimate was evaluated by simulations and real data from a tropical rain forest of Malaysia.
(2) Following the same approach for the classical occupancy model, we derive an abundance estimate by assuming contagious distribution of the individuals. The estimate was also evaluated for the same data from the tropical rain forest.
(3) We show the connection of the estimates to species detectability in population sampling and derived a variance estimate to quantify uncertainty in detectability.

5.2 OCCUPANCY MODELS OF RANDOM PLACEMENT

5.2.1 The Classical Occupancy Model

The individuals of a species in an area can be distributed in many ways, which range from aggregated to regular patterns. Different spatial distributions result in different occurrence maps (Equation 5.1), even though the number of occupied cells may

remain the same. In total, there are $\binom{M}{m}$ or $\binom{M}{u}$ possible maps for $\sum_{i=1}^{M} x_i = m$ and $u = M - m$, where m is the number of occupied cells, u is the number of empty cells. Figure 5.2 shows an example for $M = 4$, $m = 2$. Amongst these maps, the simplest case arises when all N individuals of a species are randomly distributed in a study area, A. This is equivalent to the situation of N individuals randomly placed into the M cells that comprise the area. Several models can be used to describe this random placement. In the statistical literature they are known as the "classical occupancy model" (Barton and David 1962, Kolchin et al. 1978).

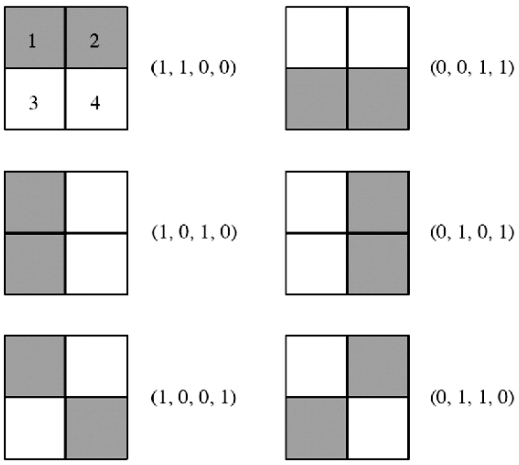


Figure 5.2. *Six possible maps of distributions for M = 4, m = 2 occurrences. Location index s = (1, 2, 3, 4) is shown in the upper left map.*

In the classical model, a species with N individuals is assumed to be randomly and independently distributed among the total number of M cells. The cell size is denoted as a, which defines the resolution of a map, or is called scale or grain in landscape ecology. It is clear that the probability that an individual falls in a given cell is simply $1/M$ or a/A and the number of organisms, n, in a given cell follows a binomial distribution, i.e.,

$$p(n) = \binom{N}{n} \left(\frac{1}{M}\right)^n \left(1 - \frac{1}{M}\right)^{N-n}, \; n = 0, 1, 2, ..., N. \tag{5.2}$$

This model can be equally written in terms of areas as

$$p(n) = \binom{N}{n} \left(\frac{a}{A}\right)^n \left(1 - \frac{a}{A}\right)^{N-n}, \; n = 0, 1, 2, ..., N. \tag{5.3}$$

A realization of Equation 5.2 or Equation 5.3 produces an occurrence map as given in Equation 5.1. What we are interested in here is to estimate N given the occupancy in the map. The problem can be thought of as equivalent to placing N balls randomly and independently into M cells. Some cells will end up with no balls, some will have one ball, and others have several balls.

It can be shown that the random placement process will lead to the moment estimate of abundance N as (see the Appendix)

$$\hat{N} = \frac{\ln(1 - m/M)}{\ln(1 - 1/M)} \tag{5.4}$$

with approximate variance given as

$$V(\hat{N}) = \frac{V(u)}{[M\ (1 - 1/M)^N \ln(1 - 1/M)]^2} \tag{5.5}$$

where $V(u)$ is given by Equation A5 in the Appendix.

Equation 5.4 relates abundance N to the number of occupied cells m and the total number of cells M. A more desirable expression that explicitly links N to mapping scale a can be readily obtained as

$$\hat{N} = \frac{\ln(1 - A_a / A)}{\ln(1 - a / A)} \tag{5.6}$$

where A_a is the total occupied area (= $a \times m$), A is the total study area (= $a \times M$).

While Equation 5.4 was derived from regularly shaped maps, Equation 5.6 is suitable for both regular and irregular maps because the data on areas are used. Equation 5.6 was obtained previously by He and Gaston (2000) by a different approach and the derivation here is more rigorous. The variance given by He and Gaston (2000) is incorrect although it differs from Equation 5.5 by a small term.

The estimate given in Equation 5.6 can be further simplified for abundant species distributed in a large study area. It is easy to show that, when the study area $A \to \infty$ (or the total number of cells $M \to \infty$), Equation 5.6 becomes

$$\hat{N} = -\frac{A}{a} \ln(1 - \frac{A_a}{A}). \tag{5.7}$$

Its variance can be similarly derived from Equation 5.5 when $N>>M$ as

$$V(\hat{N}) = \frac{A}{a} \exp(\frac{Na}{A}). \tag{5.8}$$

5.2.2 Simulation Test and Applications

5.2.2.1 Simulation

The performance of Equation 5.4 or 5.6 was evaluated by generating a random distribution of a known number of "trees" (points) in an area. We simulated the distribution of "species" in a study plot of 500 × 1000 m. Three "species" were generated. The first one had 500 "trees" that were randomly located within the study plot. The plot was then divided into a lattice with scale a = 50 × 50 m to create a distribution map. Based on this map, the number of trees was estimated by using Equation 5.4. The simulation was repeated 100 times. The estimates are shown in Figure 5.3 (Species 1), along with the upper and lower bounds of the 95% confidence defined by $N \pm 1.96\sqrt{V(\hat{N})}$, where $V(\hat{N})$ is given by Equation 5.5.

The second "species" had 2000 "trees", but this time the distribution was converted into a map with scale a = 25 × 25 m (as illustrated in Figure 5.1b). The third "species" had 5000 "trees" in a map at the same scale as for the second species. The results in Figure 5.3 show that Equation 5.4 estimates the abundances reasonably well for the randomly and independently distributed species. It appears that with the increase in N the approximate 95% confidence intervals constructed using the asymptotic variance of Equation 5.5 are too liberal when $N >> M$.

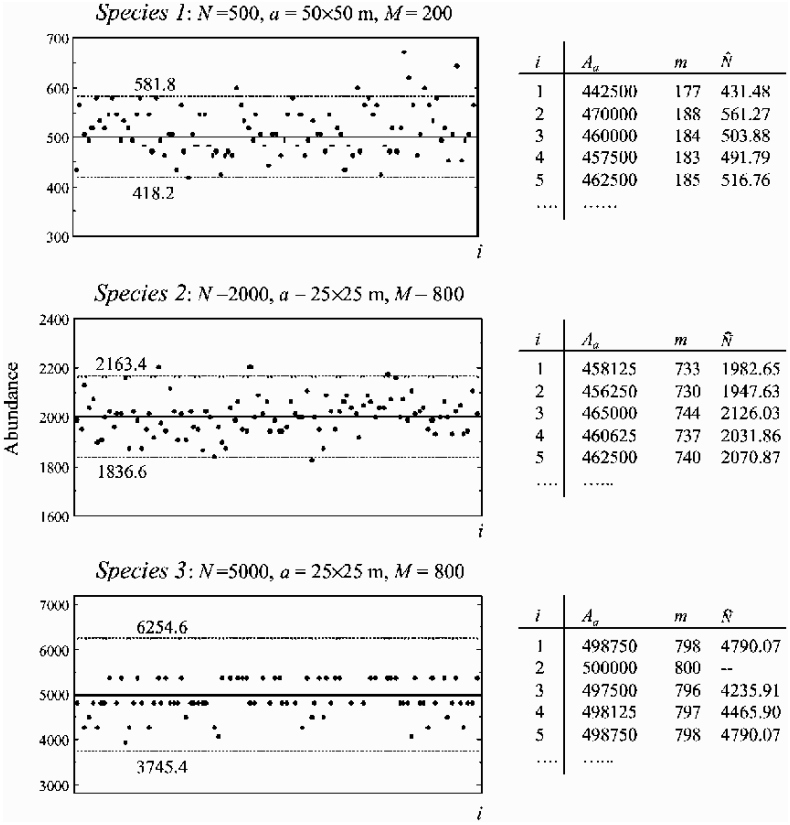
There was no estimation in the second simulation for species 3 (see the last table in Figure 5.3). This happens when a species fills up the entire area of a study. In this case, $m = M$, there is no solution to Equation 5.4.

5.2.2.2 *Applications*

We now apply Equation 5.4 to estimate the abundances of tree species in a lowland rainforest of Malaysia. The forest is located in the Pasoh Forest Reserve of Malaysia (2°55' N, 102°18' W). A 50 ha rectangular plot (500 × 1000 m) was initially established in 1987 and the census was repeated in 1990 and 1995 (Manokaran et al. 1999). The data from the 1995 census were used in this study. In each census, all free-standing trees and shrubs with diameter at breast height ≥ 1 cm were located by geographical coordinates on a reference map, and identified to species. In the 1995 survey, there were a total of 378224 trees belonging to 824 species. The most abundant species had 10470 individuals. Figure 5.1a is the distribution for one of the 824 species. The spatial patterns of the species surveyed in1990 were analyzed by He et al. (1997) which showed that about 80% of the species were aggregated, 20% had random distributions, and only one displayed a regular distribution. Because the abundance of each species was known, these census data allowed us to test the models that we developed.

Thirty-five of the 824 species were selected for analysis to represent the abundance range and spatial distribution patterns of the forest. The observed (true) abundances for the 35 species are listed in Table 5.1 together with the areas of

occupancy at four scales (a = 10 × 10, 12.5 × 12.5, 25 × 25 and 50 × 50 m). Note that the total area of the study A is 500000 m^2 and that the area of occupancy A_a can be read from Table 5.1 for a given scale a. Substitute these three values into Equation 5.6, the abundance for each species could be estimated, and its corresponding variance can also be obtained from Equation 5.5. The results are shown in Table 5.2. The results for the simplified Equation 5.7 at a = 25 × 25 m are also presented in Table 5.2. It is clear that the simplified Equation 5.7 differs very little from Equation 5.6 even for rare species.



i	A_a	m	$\hat{N}$
1	442500	177	431.48
2	470000	188	561.27
3	460000	184	503.88
4	457500	183	491.79
5	462500	185	516.76
....	.....		

i	A_a	m	$\hat{N}$
1	458125	733	1982.65
2	456250	730	1947.63
3	465000	744	2126.03
4	460625	737	2031.86
5	462500	740	2070.87
....	.....		

i	A_a	m	$\hat{N}$
1	498750	798	4790.07
2	500000	800	--
3	497500	796	4235.91
4	498125	797	4465.90
5	498750	798	4790.07
....	.....		

Figure 5.3. *Estimation of abundance for three simulated "species" in a 500 × 1000 m plot. The figures on the left-hand column are the outputs of 100 simulations for each species. The dashed lines are* $N \pm 1.96\sqrt{V(\hat{N})}$, *where* $V(\hat{N})$ *is given by Equation 5.5. The tables on the right-hand column are the outputs of the first five simulations for each species.*

Table 5.1. *Observed (true) abundance for 35 of 824 species in the Pasoh plot, and their area (m^2) of occupancy at four scales: 10 × 10, 12.5 × 12.5, 25 × 25 and 50 × 50 m.*

Species	*Abundance*	*Cell Size (m^2)*			
		10 × 10	*12.5 × 12.5*	*25 × 25*	*50 × 50*
1	1	100	156.25	625	2500
2	10	900	1562.50	5000	15000
3	13	1300	2031.25	8125	32500
4	22	1900	2968.75	11250	45000
5	27	2600	4062.50	15625	55000
6	30	2900	4531.25	14375	32500
7	50	5000	7500	27500	90000
8	98	9300	14375	51250	155000
9	115	9700	15156.25	45625	130000
10	122	11700	18437.50	65625	202500
11	155	14600	22031.25	79375	235000
12	157	14700	22812.5	83750	255000
13	177	16700	25312.50	82500	245000
14	207	19700	30468.75	105625	290000
15	302	27600	42031.25	140625	322500
16	325	30800	46406.25	158125	385000
17	333	31100	47968.75	159375	362500
18	384	33200	50312.50	162500	377500
19	405	36100	55156.25	175625	395000
20	490	44700	64687.50	195000	390000
21	520	43900	63750	161875	302500
22	522	45000	68593.75	199375	407500
23	537	47300	70625	203125	357500
24	742	63900	92968.75	262500	445000
25	874	72900	109062	290000	477500
26	891	74600	110469	286875	462500
27	1371	111400	156875	353750	482500
28	1419	115200	166562	376250	497500
29	2190	168000	231094	428750	492500
30	2793	166200	222812	410625	490000
31	3181	190300	246094	421250	492500
32	6031	308200	371562	478125	495000
33	7202	173400	200000	287500	392500
34	8571	186400	207656	275000	340000
35	10470	383300	433750	496875	500000

Table 5.2. *Estimated abundances for the 35 species in Table 5.1 using random placement Equation 5.6 and its simplified Equation 5.7 at four scales. The last row measures the "goodness-of-estimation" Δ of Equation 5.9.*

Cell Size (m^2)		*10 × 10*	*12.5 × 12.5*	*25 × 25*		*50 × 50*
Species	*True*	*Equation 5.6*	*Equation 5.6*	*Equation 5.6*	*Equation 5.7*	*Equation 5.6*
1	1	1	1	1	1	1
2	10	9.0	10.0	8.0	8.0	6.1
3	13	13.0	12.0	12.1	13.1	13.4
4	22	19.0	19.1	18.2	18.2	18.8
5	27	26.1	26.1	25.4	25.4	23.3
6	30	29.1	29.1	23.3	23.3	13.4
7	50	50.3	48.4	45.2	45.3	39.6
8	98	93.9	93.3	86.5	86.5	74.0
9	115	97.9	98.5	76.5	76.6	60.1
10	122	118.4	120.2	112.5	112.6	103.6
11	155	148.2	144.2	138.2	138.3	126.7
12	157	149.2	149.4	146.6	146.7	142.3
13	177	169.8	166.2	144.2	144.3	134.3
14	207	201.0	201.2	189.7	189.8	173.1
15	302	283.9	280.9	264.0	264.2	206.6
16	325	317.9	311.7	303.9	304.1	293.2
17	333	321.1	322.7	306.9	307.1	257.6
18	384	343.5	339.3	314.2	314.4	280.6
19	405	374.7	374.0	346.0	346.2	311.4
20	490	468.2	443.3	395.2	395.4	302.1
21	520	459.4	436.4	312.8	312.8	185.3
22	522	471.5	472.1	406.7	407.0	336.6
23	537	496.8	487.2	416.8	417.0	250.4
24	742	683.6	658.2	595.2	595.6	440.4
25	874	787.9	787.3	693.6	694.0	618.7
26	891	807.8	798.8	681.8	682.2	516.8
27	1371	1260.2	1204.7	982.8	983.4	668.8
28	1419	1309.3	1296.3	1116.4	1117.1	1057.0
29	2190	2047.2	1984.5	1557.8	1558.7	837.8
30	2793	2020.1	1887.4	1376.6	1377.4	780.5
31	3181	2394.8	2168.1	1477.7	1478.7	837.8
32	6031	4790.3	4348.6	2501.9	2503.4	918.7
33	7202	2129.2	1634.4	684.1	684.5	306.7
34	8571	2332.2	1717.1	638.4	638.8	227.3
35	10470	7274.3	6466.7	4057.6	4060.1	–
Δ		1.206	1.368	1.996	1.994	2.649

To measure the "goodness-of-estimation", we define

$$\Delta = \sqrt{\sum b^2(\hat{N}_i)} \tag{5.9}$$

where $b(\hat{N}_i) = \dfrac{\hat{N}_i - N_i}{N_i}$ for species *i*.

The results in Table 5.2 show that except for those rare species there is considerable underestimation and that the underestimation becomes stronger with the increase of scale as evident from the measurement of Δ of Equation 5.9 (see the last row of Table 5.2). This is expected because few species in nature would present a truly random and independent distribution. Except at very low abundance, individuals of most species are typically aggregated. The underestimation of the random placement Equation 5.6 is largely due to the aggregation of a species, (overestimation would be more common if a species is actually at regular distribution). In other words, if the individuals of a species are not randomly and independently distributed, Equation 5.6 is biased. To reduce the bias, we need a method to take account of species aggregation.

5.3 OCCUPANCY MODELS OF CONTAGIOUS PLACEMENT

5.3.1 The Contagious Occupancy Model

Aggregated pattern arises when the distribution of individuals of a species among cells is produced by contagious processes. In this situation, the assumption of randomness and independence no longer holds; instead, a cell that already has an individual would be more likely to contain more individuals, and an occupied cell would be more likely to be adjacent to another occupied cell (and vice versa for empty cells). Barton and David (1959) show that contagious processes can either be modeled by a negative hypergeometric distribution arising from a Polyà urn model (they termed this model as pseudo-contagious process) or by a model of true contagion in which cells to be occupied are first selected at random and the number of individuals in the selected cells are then determined as realizations of logarithmically distributed random variables. It is well known that such a process generates the negative binomial distribution. For this latter model it can be shown that the moment estimate of abundance *N* of a species is (see Appendix B) given by

$$\hat{N} = Mk\left[\left(1 - \frac{m}{M}\right)^{-1/k} - 1\right] \tag{5.10}$$

where *k* is the aggregation parameter of the negative binomial distribution that takes positive values. Aggregated species have small *k* while random species have large *k*. Equation 5.10 can also be expressed in terms of areas as

$$\hat{N} = \frac{Ak}{a}\left[\left(1 - \frac{A_a}{A}\right)^{-1/k} - 1\right] \tag{5.11}$$

with approximate variance (see Equation B5)

$$V(\hat{N}) = \left(1 + \frac{N}{Mk}\right)^{2k+2} V(u) \tag{5.12}$$

where u is the number of empty cells and $V(u)$ is given by Equation B4 in the Appendix.

5.3.2 Test for the Contagious Abundance Estimate

Given a distribution map, there are two unknown parameters N and k to be estimated in Equation 5.10 or 5.11. To use the method of moments one would normally equate the observed first and second moments with their theoretical mean and variance (see Equation B3 and B4 in the Appendix). In this case, however, a map only has one single observation on u, and its variance $V(u)$ is not available. Splitting the map to create more observations will not work, because there will be a new unknown parameter for each part of the map (the number of organisms in that part). A possible alternative is to group cells to produce a coarser scale map (Kunin 1998, He and Gaston 2000). This will lead to two equations of Equation 5.10 or 5.11 for two unknown variables N and k with the assumption that the aggregation parameter k remains the same at both scales. We realize that this assumption does not necessarily hold in reality (Pielou 1957, Taylor et al. 1978). However, if the difference in scale for the two maps is not large, this assumption may be plausible.

From the fine-scale map we can read the area of occupancy A_{a1} at scale a_1. The second map can be produced as follows. If any of the adjacent cells at the fine-scale map are occupied, then the aggregated cell at coarse-scale is occupied; otherwise, it is left empty. The second map has a coarse scale a_2 and an area of occupancy A_{a2} (see Figure 5.1c for an example).

With the two maps so generated, N and k in Equation 5.10 can be evaluated numerically using, e.g., Newton-Raphson method. The estimated abundances for 35 of the 824 species are shown in Table 5.3. The first one is calculated in terms of two maps: the fine-scale map with $a_1 = 12.5 \times 12.5^2$ m and the coarse-scale map with $a_2 = 25 \times 25$ m^2. The second map pair is the two maps with scales at $a_1 = 25 \times 25$ m^2 and $a_2 = 50 \times 50$ m^2. The results show that Equation 5.10 works fairly well. Compared with the random placement Equation 5.4 in the previous section, the estimation is substantially improved (compared the Δ's in the last rows of Tables 5.2 and 5.3). Estimation for rare species (e.g., abundance $\leq$ 2000) appears to work particularly well, which is indeed the strength of the method (Equation 5.10)

because the abundance information on rare species is the major concern of conservation.

Table 5.3. *Estimated abundance for the 35 species in Table 5.1 using the contagious occupancy Equation 5.11 in terms of two map pairs: 12.5 × 12.5 – 25 × 25 m, and 25 × 25 – 50 × 50 m. The last row measures the "goodness-of-estimation" Δ of Equation 5.9.*

Species	*True*	*12.5 × 12.5 – 25 × 25*	*25 × 25 – 50 × 50*
1	1	1	1
2	10	11.0	9.1
3	13	13.0	13.0
4	22	19.4	18.0
5	27	26.4	26.2
6	30	32.1	34.0
7	50	49.5	47.6
8	98	96.0	91.7
9	115	110.3	85.1
10	122	123.1	115.9
11	155	146.3	142.7
12	157	150.4	148.1
13	177	175.8	147.9
14	207	205.4	196.3
15	302	287.2	294.3
16	325	314.3	307.7
17	333	328.5	329.5
18	384	348.9	328.0
19	405	384.7	359.9
20	490	463.0	446.8
21	520	515.1	439.5
22	522	500.8	439.8
23	537	518.5	576.3
24	742	683.4	687.0
25	874	826.6	724.2
26	891	851.0	774.8
27	1371	1312.8	1206.3
28	1419	1375.2	1138.1
29	2190	2209.1	2492.7
30	2793	2200.8	2043.2
31	3181	2659.5	2193.3
32	6031	6341.9	14253.4
33	7202	4803.7	1603.1
34	8571	9078.3	4603.2
35	10470	8548.8	3528.3
Δ		0.542	1.920

It is apparent that a considerable degree of underestimation still remains for those very abundant species. This underestimation is also observed for abundant insects (Warren et al. 2003). However, in that study they did not estimate abundances for rare species which are actually very simple to compute, their conclusions are thus unfortunately biased.

Similar to Equation 5.4, the accuracy of the estimation of Equation 5.10 also depends on the scale of observation. The results as measured by Δ of Equation 5.9 (the last row of Table 5.3) show that the estimation becomes progressively poorer with the increase in the scale from the map pair of $12.5 \times 12.5 - 25 \times 25$ m^2 to that of $25 \times 25 - 50 \times 50$ m^2. In addition to the effect of scale on the accuracy of abundance estimates, the precision (i.e., the variances of Equation 5.5 and 5.12) of the estimates also varies with scale. Figure 5.4 shows the effect of spatial aggregation on the variance-scale relationship for Equations 5.5 and 5.12. At random distribution (Equation 5.5), variance in abundance monotonically increases with scale (the dashed line), while the variance of Equation 5.12 can be hump-shaped for aggregated species. The practical implications of the variance-scale relationship are that sampling scale for randomly distributed species should be as small as possible if high precision is to be achieved, and that, for aggregated species, the model scales that lead to high variance should be avoided in order to achieve high precision.

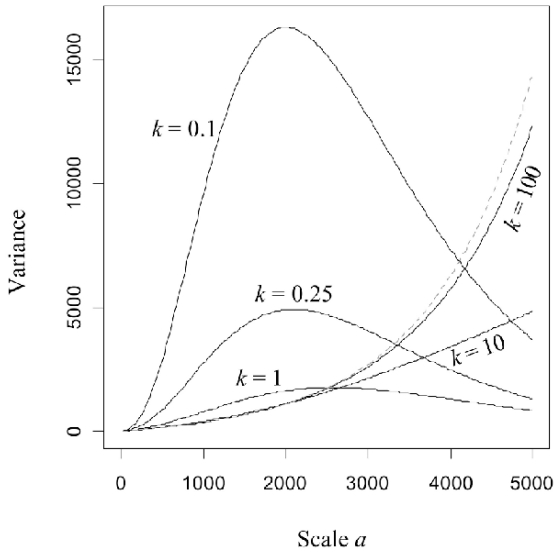


Figure 5.4. *Variance-scale relationships. The dashed curve is the variance for the classical random placement estimate (Equation 5.5). The solid curves are the variances of Equation 5.12 with the aggregation parameter k varying from 0.1 to 100. The plot is produced by setting N=500 and a from 0 to 5000. Note Equation 5.12 approaches Equation 5.5 at large k.*

5.4 UNCERTAINTY IN OCCUPANCY AND SPECIES DETECTABILITY

Equation 5.11 is an important occupancy-abundance model in ecology (Wright 1991, Hanski et al. 1993, Hartley 1998, He and Gaston 2000), which is typically written in the standard form as

$$p = 1 - \left(1 + \frac{Na}{Ak}\right)^{-k} \tag{5.13}$$

where p is the proportion of occupied area, or A_a/A.

More than any other occupancy-abundance models in the literature (see Holt et al. 2002 for a review), Equation 5.13 unifies occupancy (p), species abundance (N), the spatial pattern of the species (k), mapping scale (a), and the extent of study area (A) into a single mathematical form. In addition, Equation 5.13 is mathematically very flexible in that many other occupancy-abundance models are its special cases (He and Gaston 2000, He et al. 2002). For example, it is easy to show that the random placement Equation 5.6 is a special case of Equation 5.11 at $k = -N$. Equation 5.13 also provides a basic tool for investigating other biodiversity patterns, such as species-area curves (He and Legendre 2002, He et al. 2002) and beta diversity patterns (Plotkin and Muller-Landau 2002).

When used for sampling populations, occupancy p is often referred to as species detectability. The occupancy-abundance model (Equation 5.13) suggests that the detectability depends not only on the abundance of the species but also on its spatial distribution and the size of sampling unit. This finding is useful in sampling design. For instance, for a given abundance we know from Equation 5.13 that strong aggregation (i.e., small positive k) leads to small detectability and randomly distributed species (large k) have large detectability. So, in order to retain a high level of detectability for an aggregated species, it is necessary to use a large sample area (i.e., large a). Similarly, Equation 5.13 would help us calculate the size of sample areas for rare and common species for a predetermined detectability p.

Another important sampling issue is that the presence of a species in a site may or may not be observed in the field, i.e., there is always an uncertainty associated with detectability. The nondetection may mean that the species is truly absent or that it is missed because of insufficient survey efforts or sampling errors (MacKenzie et al. 2002). The latter scenario will inevitably lead to underestimation of occupancy rates. This uncertainty in occupancy p can be quantified according to the theory of occupancy in the Appendix. Because $p = m/M = 1 - u/M$, where the number of empty cells u is a random variable, it is easy to show that p has a variance of $V(p) = V(u)/M^2$, where the variance $V(u)$ is given either by Equation A5 or B4, depending on whether the random placement model or the contagious model is used. $V(p)$ provides information on the uncertainty in the detectability.

5.5 CONCLUSIONS

Based on the discussions in previous sections, we make the following conclusions:

(1) The widely recognized distribution-abundance macroecological pattern suggests that distribution and abundance of species are closely correlated so that we may infer about one from the other. This premise provides an essential basis for deriving information on abundance in terms of the distribution of species. In this chapter, we approached this problem by modeling the distribution of species in landscapes with the occupation process that N balls are placed into M cells following the theory of combinatorial occupancy.

(2) Two abundance estimates were derived from the theory of occupancy. The first one, as given by Equation 5.6, was derived under the assumption that the N unknown balls are randomly and independently placed into M cells. The second estimate (Equation 5.11) was derived from the contagious process that generates aggregated distribution of species (the negative binomial distribution). The random placement Equation 5.6 is a special case of Equation 5.11 at $k = -N$.

(3) While Equation 5.6 predicted very accurately the abundance of randomly placed species, it underestimated the abundance of aggregated species. Equation 5.11 greatly improved the accuracy of the estimation for real species because it accounts for aggregation with the addition of parameter k. Nevertheless, both simulated and observed data showed that the accuracy of the estimates consistently decreases with scale (i.e., the mapping resolution). The underestimation is particularly serious for abundant species as evidenced in Tables 5.2 and 5.3. Similarly, high intensity of aggregation would result in a poor estimation. In the extreme case, if a species is so highly aggregated that all of the individuals are clustered in a single cell, none of the methods could differentiate this species from the one that has only one individual and occurs also in a single cell.

(4) Equation 5.13 is a fundamental occupancy-abundance model that unifies occupancy (p), species abundance (N), the spatial pattern of the species (k), mapping scale (a) and the extent of study area (A) into a single mathematical form. The model suggests that occupancy (or species detectability) depends not only on the abundance of the species but also on its spatial pattern and the size of sample unit. This finding would help us understand the factors that may influence the detectability of species in a field survey and, thus, design the survey in order to maintain a desirable level of detectability (e.g., to calculate the size of sample unit). The derived variance for the occupancy p can be used to quantify the uncertainty in the detectability.

(5) Two obvious questions need to be answered: How can we further incorporate the information on scale (i.e., mapping resolution) and aggregation to improve the estimation? What scale should be used for mapping a distribution to ensure a certain level of accuracy? The questions about scale appear to be more challenging. Answers necessarily depend on the life history properties of organisms. For example, for insects with small body sizes and highly aggregated distribution, a small mapping scale compatible with the size of the insects should be used (e.g., in centimeters or a few meters), while for large body trees, a relatively large mapping

scale may be used (e.g., in 10 or 100 meters). The underestimation caused by strong aggregation may be solved by some ad hoc methods. The aggregated mapping method used in this chapter (Figure 5.1b, c) assumed that k in Equation 5.11 was constant. This assumption may be relaxed by correcting the k by comparing the observed map with its random counterpart.

(6) This study deals only with exhaustive survey of a distribution map as defined by Equation 5.1. It will also be interesting and useful to consider Equation 5.1 as a random sample from a distribution map. If one knows that a species is randomly and independently distributed, the classical occupancy estimate Equation 5.6 can be applied to estimating abundance in this sampling scheme. But in other situations where the random and independent assumption does not hold, the estimation of abundance is a challenging task. This is certainly an interesting problem deserving further investigation.

APPENDIX: DERIVATION OF OCCUPANCY-ABUNDANCE MODELS

A. The Random Placement Occupancy Model

Assume a distribution map of m occupied cells out of M total number of cells. Let u ($= M - m$) be the number of empty cells, and let E_i be the event that the i^{th} cell is empty and $\overline{E}_i$ be the event complementary to E_i. Then the probability that *one particular*, say the first, cell is empty is $p(E_1) = (1 - \frac{1}{M})^N$ which is equivalent to $n = 0$ in the binomial distribution Equation 5.2, or obtained by replacing A in Equation 5.3 by $a \times M$.

The probability that *two particular*, say the first two, cells are empty is $p(E_1E_2) = (1 - \frac{2}{M})^N$. This probability can again be derived from Equation 5.2 with $n = 0$ by replacing 1 by 2 since there are *two* empty cells, or by replacing a by $2 \times a$ in Equation 5.3. Similarly, the probability that u *particular* cells are empty is $p(E_1 ... E_u) = (1 - \frac{u}{M})^N$.

Here we shall not be interested in a *particular* set of cells but the number of u empty cells given N balls being placed into M cells. From Figure 5.2, we know there are $\binom{M}{u}$ possible combinations for u (out of M) empty cells. Thus the probability that there are u empty cells is $p(u) = \binom{M}{u} p(E_{i_1} ... E_{i_u} \overline{E}_{j_u} ... \overline{E}_{j_{M-u}})$. It is equivalent to

$$p(u) = \binom{M}{u} p(E_1 ... E_u \overline{E}_{u+1} ... \overline{E}_M). \tag{A1}$$

Because E_1, …, E_u, $\overline{E}_{u+1}$, …, $\overline{E}_M$ are independent events, by some probability operations, we arrived at

$$\begin{aligned} p(E_1 \ldots E_u \overline{E}_{u+1} \ldots \overline{E}_M) &= \left(1-\frac{u}{M}\right)^N \sum_{i=0}^{M-u} (-1)^i \binom{M-u}{i} \left(1-\frac{i}{M-u}\right)^N \\ &= \sum_{i=0}^{M-u} (-1)^i \binom{M-u}{i} \left(1-\frac{u+i}{M}\right)^N . \end{aligned} \tag{A2}$$

Finally, the probability that there are u empty cells given N balls randomly and independently placed into M cells is derived by substituting Equation A2 into Equation A1, i.e.,

$$p(u) = \binom{M}{u} \sum_{i=0}^{M-u} (-1)^i \binom{M-u}{i} \left(1-\frac{u+i}{M}\right)^N , \text{ for } u = 0, 1, \ldots, M. \tag{A3}$$

The factorial moment of the number of u empty cells of the probability mass function equation A3 is known to be (Johnson et al. 1993, p. 415)

$$\lambda_{[r]} = \frac{M!}{(M-r)!}\left(1-\frac{r}{M}\right)^N,$$

where $\lambda_{[r]} = E\left(\frac{u!}{(u-r)!}\right)$. Thus the expectation and variance of the number of empty cells are

$$E(u) = \lambda_{[1]} = M\left(1-\frac{1}{M}\right)^N \tag{A4}$$

$$\begin{aligned} V(u) &= \lambda_{[2]} + E(u) - E^2(u) \\ &= M(M-1)\left(1-\frac{2}{M}\right)^N + M\left(1-\frac{1}{M}\right)^N - M^2\left(1-\frac{1}{M}\right)^{2N}. \end{aligned} \tag{A5}$$

The variance for the number of occupied cells m is the same as Equation A5 for a given map with fixed M because $V(u) = V(M-m) = V(m)$.

Given an occurrence map, it is obvious that the moment estimate of $E(u)$ is simply $M - m$. Hence, the estimate of N can be solved from Equation A4 as

$$\hat{N} = \frac{\ln(u/M)}{\ln(1-1/M)} = \frac{\ln(1-m/M)}{\ln(1-1/M)}. \tag{A6}$$

The approximate variance of the abundance estimate $\hat{N}$ in Equation A6 can be easily obtained by applying the delta method to Equation A6, i.e.,

$$V(\hat{N}) = \left[\frac{\partial(\hat{N}(u))}{\partial u}\right]^2_{u=E(u)} V(u), \tag{A7}$$

where $V(u)$ is as Equation A5 and $\hat{N}$ is as Equation A6. The derivative is evaluated at $E(u)$ of Equation A4. The variance so obtained is

$$V(\hat{N}) = \frac{V(u)}{\left[M(1-1/M)^N \ln(1-1/M)\right]^2}. \tag{A8}$$

B. The Contagious Occupancy Model

For the contagious process that generates the negative binomial distribution, Barton and David (1959) show that the distribution of the number of empty cells u has probability mass function:

$$p(u) = \frac{\binom{M}{u}}{(kM+N-1)^{(N)}} \sum_{i=0}^{M-u} (-1)^i \binom{M-u}{i} [kM-k(u+i)+N-1]^{(N)}, \tag{B1}$$

where $i^{(j)} = \frac{i!}{(i-j)!}$, k is the aggregation parameter of the negative binomial distribution, N is the (unknown) number of organisms of a species distributed in a defined area with size A, M is the total number of cells dividing A.

The r^{th} factorial moment of u is

$$\lambda_{[r]} = M^{(r)} \frac{(kM-1)!(kM-kr+N-1)!}{(kM+N-1)!(kM-kr-1)!} \tag{B2}$$

from which the expectation and variance of u can be found. However, Barton and David (1959) show that, even for relatively small M and N, the pmf given by Equation B1 can be well approximated by a normal distribution with mean and variance:

$$E(u) = M\left(1+\frac{N}{Mk}\right)^{-k} \tag{B3}$$

$$V(u) = Me^{-2\mu}\left(e^{\mu} - 1 - \mu\right) \tag{B4}$$

where $\mu = \frac{N}{M}$.

Given a binary map, it is straightforward that the observed first moment estimate of $E(u)$ is $M - m$. Therefore, from Equation B3 the moment estimate of N, Equation 5.10, is resulted. The variance of the estimate $\hat{N}$ of Equation 5.10 can be derived from Equation B3 and Equation B4 using the delta method following Equation A7:

$$V(\hat{N}) = \frac{1}{M^2}\left(1 + \frac{N}{Mk}\right)^{2k+2} V(u) \tag{B5}$$

where $V(u)$ is given by Equation B4.

REFERENCES

Barton, D. E., and F. N. David. 1959. Contagious occupancy. Journal of the Royal Statistical Society, Series B. 21:120-133.

Barton, D. E., and F. N. David. 1962. Combinatorial Chance. Charles Griffin, London, UK.

Brown, J. H. 1984. On the relationship between abundance and distribution of species. American Naturalist 124:255-279.

Brown, J. H. 1995. Macroecology. University of Chicago Press, Chicago.

Feller, W. 1967. An Introduction to Probability Theory and Its Applications. Wiley, New York.

Gaston, K. J. 1994. Rarity. Chapman & Hall, London, UK.

Gaston, K. J., and T. M. Blackburn. 2000. Pattern and Process in Macroecology. Blackwell, London, UK.

Hanski, I., J. Kouki, and A. Halkka. 1993. Three explanations of the positive relationship between distribution and abundance of species. Pages 108-116 *in* R. E. Ricklefs and D. Schluter, editors. Species diversity in ecological communities: historical and geographical perspectives. University of Chicago Press, Chicago.

Harte, J., T. Blackburn, and A. Ostling. 2001. Self-similarity and the relationship between abundance and range size. American Naturalist 157:374-386.

Hartley, S. 1998. A positive relationship between local abundance and regional occupancy is almost inevitable (but not all positive relationships are the same). Journal of Animal Ecology 67:992-994.

He, F., and K. J. Gaston. 2000. Estimating species abundance from occurrence. American Naturalist 156:553-559.

He, F., and P. Legendre. 2002. Species diversity patterns derived from species-area models. Ecology 83:1185-1198.

He, F., K. J. Gaston, and J. Wu. 2002. On species occupancy-abundance models. Ecoscience 23:503-511.

He, F., P. Legendre, and J. V. LaFrankie. 1997. Distribution patterns of tree species in a Malaysian tropical rain forest. Journal of Vegetation Science 8:105-114.

Holt, A. R., K. J. Gaston, and F. He. 2002. Occupancy-abundance relationships and spatial distribution: a review. Basic & Applied Ecology 3:1-13.

Johnson, N. L., S. Kotz, and A. W. Kemp. 1993. Univariate Discrete Distributions. Wiley, New York.

Kolasa, J. 1989. Ecological systems in hierarchical perspective: breaks in community structure and other consequences. Ecology 70:36-47.

Kolchin, V. F., B. A. Sevast′yanov, and V. P. Chistyakov. 1978. Random Allocations. Winstons & Sons, Washington, D.C.

Kunin, W. E. 1998. Extrapolating species abundance across spatial scales. Science 281:1513-1515.

Kunin, W. E., S. Hartley, and J. L. Lennon. 2000. Scaling down: on the challenging of estimating abundance from occurrence patterns. American Naturalist 156:560-566.

Leitner, W. A., and M. L. Rosenzweig. 1997. Nested species-area curves and stochastic sampling: a new theory. Oikos 79:503-512.

MacKenzie, D. I., J. D. Nichols, G. B. Lachman, S. Droege, J. A. Royle, and C. A. Langtimm. 2002. Estimating site occupancy rates when detection probabilities are less than one. Ecology 83:2248-2255.

Manokaran, N., J. V. LaFrankie, K. M. Kochummen, E. S. Quah, J. E. Klahn, P. S. Ashton, and S. P. Hubbell. 1999. The Pasoh 50-ha Forest Dynamics Plot: 1999 CD-ROM Version. Forest Research Institute of Malaysia, Kepong, Malaysia.

Maurer, B. A. 1990. The relationship between distribution and abundance in a patchy environment. Oikos 58:181-189.

Pielou, E. C. 1957. The effect of quadrat size on the estimation of the parameters of Neyman's and Thomas' distributions. Journal of Ecology 45:31-47.

Plotkin, J. B., and H. C. Muller-Landau. 2002. Sampling the species composition of a landscape. Ecology 83:3344-3356.

Taylor, L. R., I. P. Woiwod, and J. N. Perry. 1978. The density dependence of spatial behaviour and the rarity of randomness. Journal of Animal Ecology 47:383-406.

Warren, M., M. A. McGeoch, and S. L. Chown. 2003. Predicting abundance from occupancy: a test for an aggregated insect assemblage. Journal of Animal Ecology 72:468-477.

Wright, D. H. 1991. Correlations between incidence and abundance are expected by chance. Journal of Biogeography 18:463-466.

CHAPTER 6

SCALING TERRESTRIAL BIOGEOCHEMICAL PROCESSES:

Contrasting Intact and Model Experimental Systems

MARK A. BRADFORD AND JAMES F. REYNOLDS

6.1 INTRODUCTION

Planet Earth is undergoing enormous change, clearly discernable on time scales of decades to centuries. This is largely a result of human activities, especially the emission of greenhouse gases and pollutants, the clearing of global forests, the urbanization of agricultural cropland, and many other extensive modifications in the land surface. Significantly, these rates of change are unprecedented in the Earth's geological history (Committee on Global Change 1988), and one of the great scientific challenges of the 21st century is to forecast future behaviors of global ecosystems under the constant pressure of human insults (Clark et al. 2001). This requires that we better understand feedbacks and interactions of the major patterns and processes of the key components of planet Earth: the atmosphere, oceans, freshwater, rocks, soils, and biosphere. In an attempt to meet this challenge, an interdisciplinary approach to studying systems dynamics on a planetary-scale has emerged, known as Earth System Science (ESS) (Schellnhuber 1999, Lawton 2001).

Biogeochemical processes of terrestrial ecosystems are at the core of ESS research (Schellnhuber 1999). Hence, considerable effort has been invested towards understanding the relative importance of biotic and abiotic regulators and controllers. Of special interest are how natural and human-induced perturbations may affect the rates and directions of biogeochemical processes in terrestrial ecosystems, especially in terms of potential feedbacks to climate systems (Walker and Steffen 1996, Pielke 2002). Given the complexity of global systems and their many interconnections, one of the main scientific challenges of ESS is to document change, diagnose underlying causes, and develop plausible projections of how natural variability and human actions may affect global biogeochemical cycles in the future. With regard to the latter, once we have the requisite quantitative understanding of process rates, as well as a detailed understanding of key regulatory

J. Wu, K.B. Jones, H. Li, and O.L. Loucks (eds.),
Scaling and Uncertainty Analysis in Ecology: Methods and Applications, 109–130.

mechanisms, the goal is to extrapolate findings obtained at one temporal and spatial scale to another.

Typically extrapolations for terrestrial ecosystems are done using mathematical models (see Goudriaan et al. 1999, Prinn et al. 1999, Wu and Li, Chapters 1 and 2). However, in this chapter we focus on extrapolations via empirical experimentation: we discuss experimental designs that inform about process rates and regulatory factors at spatial and temporal scales *greater* than the one on which the experiment is conducted. Paradoxically, while most experiments are in fact *intended* to further understanding and knowledge at scales *beyond* the ones at which they are actually being conducted (e.g., 1 m^2 plots or a forested watershed), most fail to incorporate the spatial and temporal scale considerations necessary to justify such an extrapolation (Gardner et al. 2001). A number of issues are germane to this discussion, including the specific characteristics of the factors under investigation, the importance of nonlinear responses, the type of treatment imposed (e.g., step vs. gradual), and whether the goal is spatial or temporal extrapolation.

We discuss experiments conducted using two general types of systems: intact systems and model systems. Our objective is to compare and contrast these approaches in the context of their potential for contributing to our predictive understanding of process rates and their regulators in terrestrial biogeochemical cycles. In this chapter we will show that: (1) *intact ecosystem experiments* can provide process rates, mechanistic understanding and absolute/relative treatment effects suitable for direct extrapolation, but rarely do; and (2) *model ecosystem experiments* can provide the sign (positive or negative) of treatment effects and insights into their mechanistic basis. However, data obtained on process rates and absolute/relative treatment effects are not suitable for extrapolation. We concur with Gardner et al. (2001) that there is a need for much greater "scale awareness" in ecology, especially with regard to the role of experimental design and execution. Our primary objective is to raise awareness of the importance of spatial and temporal scale considerations in the design and interpretation of experiments, so that findings at the scale of an experimental plot and duration may be extrapolated with *known* confidence.

6.2 DEFINITIONS

For the purposes of our discussion, empirical experimental systems are grouped into two classes: (i) intact and (ii) model. Prior to embarking on our discussion it is important to note that although neither mathematical modeling nor observation are explicitly discussed here, we recognize the critical importance of these approaches, which are adequately covered elsewhere (e.g., Goudriaan et al. 1999, Gardner et al. 2001). Also, much discussion on scaling in experimental ecology is semantic and, albeit this is an important debate, it is beyond the scope of our discussion. The literature is replete with terminology that is often contradictory and/or case-specific; hence, we cite established terminology where possible and provide definitions when introducing context-specific terms.

6.2.1 Intact Ecosystem Experiments

Intact ecosystem experiments (IEEs) utilize plots where the established ecosystem of interest, e.g., a forest, agricultural field or desert, is sampled "intact" by the experimenter. Within the ecosystem of interest, a defined area is selected (arbitrarily, pragmatically or using a statistical rationale) and one or more factors are manipulated or perturbed across replicate plots *in situ*. Treatment effects are quantified by way of comparison with non-manipulated "control" or "reference" plots. Examples of IEEs are the exclusion of exotic, mammalian herbivores using fenced areas across 30 locations (Wardle et al. 2001); surrounding six 30 m diameter forest plots with free air CO_2 enrichment (FACE) rings and exposing half to ambient and half to elevated atmospheric CO_2 concentrations (DeLucia et al. 1999); and establishing ten alpine tundra plots and adding nitrogen fertilizer to five (Neff et al. 2002).

6.2.2 Model Ecosystem Experiments

In contrast, model ecosystem experiments (MEEs) are conducted either within the field or laboratory. We define MEEs as those where the experimental system is either wholly or partially "constructed" and one or more treatments are applied. Well-known field examples are the biodiversity manipulations at Cedar Creek (Tilman et al. 1996) and the pan-European Biodepth study (Hector et al. 1999), where plant communities of varying species richness and composition were constructed. Laboratory-based experimental systems span ranges of biotic complexity from simple Gaussian style systems (e.g., McGrady-Steed et al. 1997), through individual plants in pots of soil grown in controlled environment chambers (e.g., Fernández and Reynolds 2000), to multi-species, multi-trophic systems constructed as analogues of intact systems (e.g., Lawton 1996). A vegetated soil monolith placed within the laboratory would, although an intact unit, be defined as an MEE because its climate is constructed.

6.3 MEASURING TREATMENT EFFECTS USING INTACT ECOSYSTEM EXPERIMENTS

6.3.1 Plot Size and Spatial Extrapolation

We define *internal regulators* as factors that affect some ecosystem process of interest and which, in contrast to *external regulators*, vary within the system. Examples of each are soil fertility and air temperature, respectively. The key to extrapolating results from experimental plots to larger spatial scales is an understanding, or at least appreciation, of which internal factors regulate the process under investigation (e.g., how soil nitrogen affects plant growth; how soil water affects leaf photosynthesis) and how these factors are spatially distributed. To illustrate this point, we use an example whereby we assume that the values of the factors are invariant over time and exert their control internally on the process of

interest. If we have an understanding of what internal regulatory factors are important – and their spatial distribution in the environment – we can opt to select the size of our plots to capture the spatial heterogeneity of these regulatory factors (i) within a single plot (e.g., Figure 6.1a), or (ii) through use of multiple plots, each of which is placed within a spatially homogenous patch of the regulatory factors (e.g., Figure 6.1b).

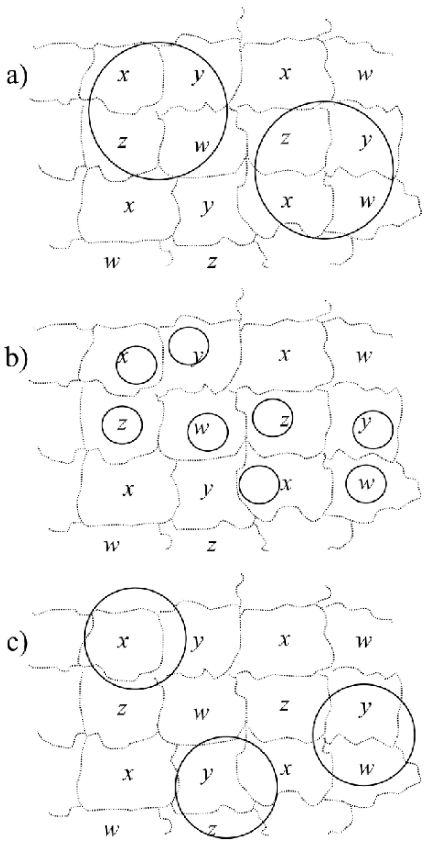


Figure 6.1. *Selection of plot sizes (open circles) to capture the spatial heterogeneity of a regulatory factor and, thus, permit spatial extrapolation of the measured process from the plot to the landscape. Letters in italics denote a fixed level of the regulatory factor. There are four levels, w-z, and the spatial distribution of the regulatory factor is repeated every two homogenous areas. In (a) and (b) plot sizes permit extrapolation to the landscape. In (a) the plot size captures the heterogeneity within a single plot. In (b) four plots are required to capture the same heterogeneity but the advantage is that the collected data can be used to construct the shape of the relationship between the rates of the process and the levels of the regulatory factor. Plots can be placed anywhere within the homogenous area of a single value of the regulatory factor. In (c) the plots are inappropriately sized to capture the spatial heterogeneity and so the results from the plots cannot be extrapolated to the landscape. In the scenario depicted in (c) the plots could be positioned to capture an equal proportion of w-z within them but in reality the heterogeneity is unlikely to be so uniformly distributed.*

To capture the spatial heterogeneity of the regulatory factors within our plots we first need to know over what "target" area we plan to extrapolate our results and then measure the spatial heterogeneity of the regulatory factors across this landscape. The advantages of capturing the spatial heterogeneity within our plot size (Figure 6.1a) are: (i) measured process rates can be directly extrapolated across the landscape of interest; (ii) there is an inherent assumption that large plot sizes inform more about how a system would naturally respond to a perturbation than smaller plot sizes; and (iii) as plots increase in size within a fixed spatial area the variation between replicate plots decreases. The disadvantage is that we do not generate data necessary to construct the shape of the relationship between process rates and

regulatory factors. This information is required to parameterize predictive models given that many processes respond nonlinearly to variation in regulatory factors (Figure 6.2a-c). From the relationships depicted in Figure 6.2a-c it is clear that a single mean value of the regulatory factor (as would be obtained by "capturing" the spatial heterogeneity within single plots) cannot be used to determine the shape of the response of the process of interest across varying regulator values (Figure 6.2d).

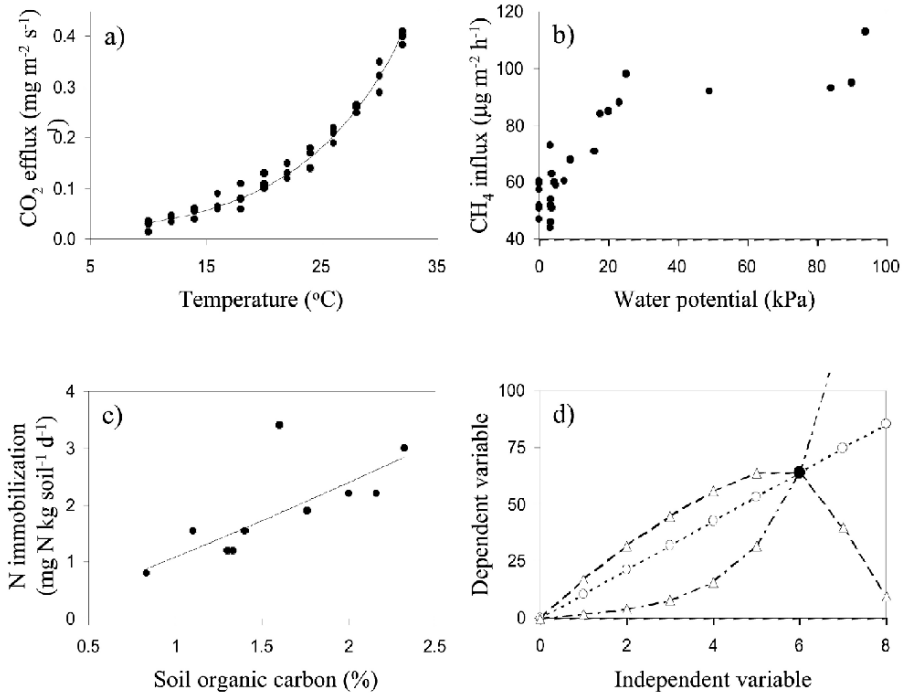


Figure 6.2. *Examples of nonlinear relationships between a process (Y) and regulatory factor (X). For the purposes of our discussion, nonlinearity exists when the relationship between X and Y is not constant (i.e., $dY/dX \neq$ constant). (a) A positive, exponential relationship exists between respiration from a forest soil and soil temperature across 10 to 32°C. Redrawn from Fang and Moncrieff (2001); not all single measurements (filled circles) are shown. (b) Bradford et al. (2001) found that soil water potential explains 78% of the variation in net, soil methane oxidation rates measured across one year in a temperate forest. (c) Barrett and Burke (2000) report a strong linear correlation between soil organic carbon content and gross nitrogen immobilization but in fact the relationship is better described by a power function (shown as solid line). (d) Examples of relationships (broken lines with open symbols) that could be constructed were measurements (filled circle denotes the mean) of a process (y-variable) taken at only one mean value of the regulatory factor (x-variable), such as if plots were sized to capture the spatial heterogeneity in regulatory factors within single plots (see Figure 6.1a). With only a single x-value there is no way to establish the shape of the relationship between X and Y.*

Theoretically, an alternative approach – to establish plots within spatially homogenous patches (Figure 6.1b) – generates the same rate estimates as obtained using larger plots and enables the shape of the relationship between process and regulatory factor to be quantified. A second advantage is that smaller plot sizes are often more tractable for manipulation but the trade-off is that greater replication is needed. Selection of a plot size that falls somewhere between (Figure 6.1c), the size that captures the spatial heterogeneity in a single plot (Figure 6.1a) and smaller plots that contain homogenous areas (Figure 6.1b), can only inform about the response of the specific experimental plot being investigated; the experimental results will have little quantitative relevance to areas outside of the experimental plot.

Our discussion of what constitutes a suitable plot size for valid spatial extrapolation belongs in an "ideal world" – in reality, a number of complications come to bear. First, large plots that capture spatial heterogeneity will rarely have smaller homogenous areas evenly proportioned within them and, therefore, rate estimates may differ between large and small plot experiments. Appropriate replication can reduce these differences. Second, single factors do not regulate biogeochemical processes although, in many cases, a single factor can explain a significant part of the observed variation in measured process rates (e.g., Figure 6.2a-c); in such instances, basing plot size on the spatial heterogeneity of this factor will provide more information than if spatial considerations were ignored. Third, spatial patterns may vary with time. This limitation is perhaps most relevant when variation is stochastic (e.g., localized rainfall within a desert) as opposed to non-stochastic (nutrients accumulate under desert shrubs across time, increasing differences in soil fertility between shrubs and interspaces; Schlesinger et al. 1996). Fourth, the identity of the landscape unit next to the one of interest, and/or the degree of connectedness between similar units (e.g., corridors in fragmented landscapes), can markedly influence processes (see Peters et al., Chapter 7, Groffman et al., Chapter 10, Lloyd et al., Chapter 14). Fifth, in many experiments multiple processes are measured within an experimental plot (e.g., Shaver et al. 1998, DeLucia et al. 1999, Bradford et al. 2002) and their regulatory factors may not only be different but also distributed, in terms of their heterogeneity, across distinctly different spatial areas. As a result there may not be a single plot size that enables all the measured processes to be extrapolated legitimately. And finally, in many cases there may be no *a priori* knowledge of which factor(s) regulates a process until an experiment has been conducted.

The above scenarios illustrate the need to consider the appropriate plot size in designing IEE studies. For example, one may decide that issues of spatial scale cannot be incorporated into an experiment due to limiting resources or knowledge about regulatory factors. If so, at least this is a conscious decision and as such the ability to extrapolate the results spatially can be critically assessed. Such "scale awareness" is a marked improvement on the current norm (Gardner et al. 2001).

6.3.2 Variability in External Regulatory Factors

In this section we develop our discussion of spatial scale in IEEs. What temporal scale considerations are necessary when biogeochemical processes are primarily

regulated by external factors, and secondarily by internal factors? To illustrate, we use an excellent IEE example in the recent literature where both *temporal* and *spatial* patterns are shown to determine the absolute and relative magnitude of treatment effects. Smith et al. (2000) exposed an intact Mojave Desert ecosystem to elevated atmospheric CO_2 and showed that new shoot production of a dominant perennial shrub doubled in a high rainfall year but was not affected by elevated CO_2 in the subsequent year, which was characterized by low rainfall. Further, in the high rainfall year production of annual and exotic grasses, which failed to germinate in the drought year, was higher under elevated CO_2 and the magnitude of the treatment response was greater within fertile "shrub-islands" than in open interspaces between perennial plants. Precipitation – an external factor that varies temporally – is the dominant control on system productivity and when delivered in sufficient amounts the spatial pattern of soil fertility across the landscape (an internal factor) becomes an important determinant of the system response to elevated CO_2. Such information on the spatial and temporal controls on productivity is necessary to predict how the Mojave might respond, at the landscape level, to future environmental change.

Spatial and temporal patterns in deserts are of course typified by extremes. Identifying and then incorporating similar scale considerations into other types of ecosystems (e.g., tropical and temperate forests) that appear more spatially and temporally homogenous in terms of factors, such as soil nutrient availability and rainfall, may be more difficult but potentially no less pertinent. For example, the magnitude of plant biomass responses to elevated CO_2 appears dependent on soil nutrient availability for the majority of species tested from across biomes (Bazzaz and Catovsky 2002). The primacy of this control is likely to be dependent on other factors, such as climate (e.g., Smith et al. 2000), but commonly there is detailed information on variation in such factors across years, which can be used to determine the degree of confidence that one can extrapolate results in time and space. However, published relationships between processes and regulatory factors should be assessed with caution prior to designing spatially- and temporally-sound experiments. For example, actual evapotranspiration (AET) explains marked variation in plant litter decomposition rates across latitudes (Aerts 1997), but within a particular system litter quality or soil community composition may be a better predictor (Aerts 1997, González and Seastedt 2001). Were such scale considerations to be ignored, predictions about how decomposition will respond to perturbation might very well be flawed. Probably the best example of such scale-ignorance is the plant diversity-productivity debate. Here, the hump-shaped relationship of diversity with increasing productivity (see Mittelbach et al. 2001), generated from inter-community/site comparisons across a landscape, has been used to suggest that a similar relationship should occur within a single community/site (Loreau et al. 2001).

6.3.3 Temporal Scale

Many of the same considerations used to determine appropriate spatial scales for extrapolation are relevant when deciding upon the necessary temporal grain and extent to extrapolate experimental findings. After defining *grain* and *extent*, we

discuss these considerations with respect to temporal scale and at the same time continue discussion appropriate to spatial scaling.

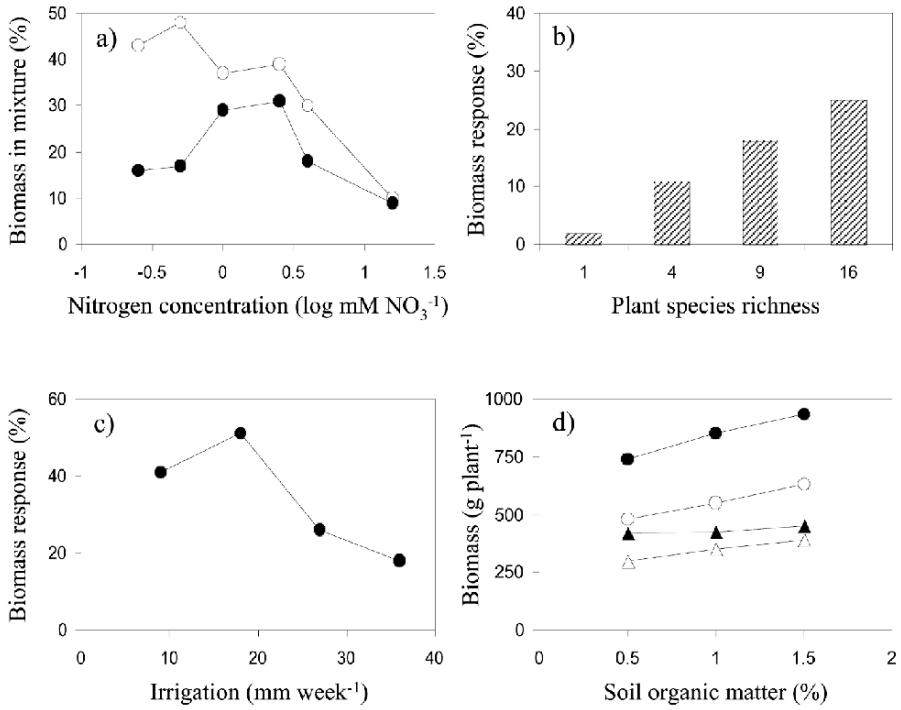


Figure 6.3. *Examples of nonlinear relationships between the magnitude of treatment effect on a process and different levels of the regulatory factor. (a) Effect of CO_2 concentration (ambient, open circles; elevated, filled circles) on the proportion of biomass that is Trifolium repens in four species plant mixtures, across a gradient of nitrogen availability. At low nitrogen concentrations there is a marked effect of CO_2 but at high nitrogen, the CO_2 concentration does not affect representation of T. repens in the mixtures (Navas et al. 1999). (b) The impact of nitrogen addition (compared to the control) on the biomass response of plant communities of increasing species richness. More diverse communities show a greater response to nitrogen treatment (Reich et al. 2001). (c) Volk et al. (2000) show that the relative response of plant biomass in two species mixtures to an increase in CO_2 concentration markedly decreases when amended with higher levels of simulated rainfall. (d) Soil organic matter content (a surrogate for soil fertility) regulates the aboveground biomass of individual, two-year old, loblolly pine seedlings but planting density moderates its regulatory role (Zutter et al. 1997). One (filled circles), two (open circles), three (filled triangles) and four (open triangles) seedlings were planted per m^2.*

Grain is the smallest spatial or temporal resolution at which data are collected (i.e., the limit to which individual measurements can be resolved), whereas extent is the spatial size (length, area or volume) of the study and the total duration over which it is observed (Wu and Li, Chapter 1). See Kemp et al. (2001) for discussion

of how these definitions are modified depending on whether the data of interest are observed in nature, collected through experimental manipulations, or measured as intrinsic scales of a natural system. In the previous section we used the Smith et al. (2000) study to highlight how temporal variability in regulatory factors can often affect the magnitude of treatment response of a process. This builds on the argument that relationships between processes and regulatory factors may be nonlinear (Figure 6.2) because it can be demonstrated that a linear relationship between the magnitude of treatment effect on a process and the regulatory factor(s) is unlikely (Figure 6.3). Thus, the temporal extent of IEE studies must capture enough variability to construct a robust relationship between regulatory factors, processes and treatment effects. This will be dependent on (i) variability in the regulatory factor(s), (ii) the process measured, (iii) the presence of synergies, feedbacks and thresholds, and (iv) how each treatment is applied. We discuss each in turn, highlighting the four main areas that we believe are necessary for consideration when deciding the temporal grain and extent of an experiment.

6.3.3.1 Variability in regulatory factors

Similar to selection of plot size (i.e., spatial extent), there is a strong relationship between the amount of heterogeneity in the regulatory factor that the experimenter is able to capture (to explain variation in the measured process) and the robustness of the mechanistic understanding generated. That is, the more levels of the regulatory factor at which the process is measured, the better quantified the relationship. Ideally, experiments should be run for as long as is feasible to enable this. When choosing spatial extent one can measure the spatial arrangement of regulatory factors; however, in setting the temporal extent one must presuppose how the regulatory factor will vary over time. This might be based on past records (if they exist), speculations, hunches, analogies with other studies, etc.: in other words, predict the future! Here, perhaps, is when the experimenter is most at the mercy of factors outside experimental control.

6.3.3.2 Frequency of measurement: integrative vs. active processes

The frequency with which to measure a process is the next decision to be made. The term *frequency* is synonymous with that of *sampling density*. The frequency chosen for a study will be dependent upon the process of interest (see below) and there is little to be gained by setting the frequency to the same resolution as the temporal extent of the IEE. For example, if we measured tree growth once at the end of a three year experiment it is unlikely that we could ever extrapolate this result in time. We would need exactly the same pattern of climatic conditions to manifest over a three year period to do so, which is unlikely. However, had we increased the frequency of our measures we would have been able to plot the treatment effect against values for regulatory temporal factors (e.g., climatic variables) and therefore construct a relationship potentially useful for extrapolation – as done by Smith et al. (2000). These authors measured aboveground production of the dominant shrubs at

their desert site in two successive years that differed markedly in rainfall and, as a consequence, rate of shrub growth. Had they measured production only at the end of the second year, then they could not have determined that production was dependent upon rainfall within a year because they would have had a single production estimate that integrated across two years of different rainfall patterns.

Commonly, the process of interest determines the frequency that we choose. For example, estimates of net primary productivity (NPP) constitute a measure of plant biomass production across time that is influenced by external factors such as climate and nitrogen deposition, and by internal factors such as site fertility, herbivory, competition, species identity and facilitation. All of these factors operate over the growing season but NPP is usually measured only once (e.g., during peak growth) or twice (rarely) to assess treatment effects (e.g., Hector et al. 1999, Shaw et al. 2002). We define measures such as NPP as a *temporally-integrative* process, in contrast to a *temporally-active* process, such as soil CO_2 efflux. The former may have units of, for example, kg m^{-2} y^{-1} whereas the latter, μmole m^{-2} s^{-1}. These distinctions muddy the notion of frequency to a certain extent because it does not simply define how often a process is measured. For example, if we measure NPP and soil CO_2 efflux once a year then the frequency is the same; however, the information gathered regarding the response of the process to the treatment will reflect very different time-scales (and therefore temporal grains!). Further, the integrative vs. active nature of a process is likely to be on a continuous rather than discrete scale, and the exact positioning for a process may depend on what ecosystem is being investigated. For example, soil CO_2 efflux in a well-drained soil may be a truly active process, but in a wetland, where movement of gas through soil is restricted, changes in atmospheric pressure and freeze-thaw cycles can lead to release of large build-ups of trace gases (Bubier et al. 2002). Hence the active vs. integrative nature of a specific process may be temporally, as well as ecosystem, dependent.

Such considerations directly impinge on the frequency with which we decide to measure a process and what mechanistic understanding we construct. For temporally-active processes, which have the potential to exhibit marked short-term variation if regulated by, say, a single but temporally dynamic factor, frequent quantification is necessary to generate robust estimates (e.g., Bradford et al. 2001). The advantage is that frequent quantification will permit strong relationships between process and regulator to be constructed within relatively short time-scales. In contrast, the mechanistic basis for variation in a temporally-integrative process such as NPP can be difficult to obtain because so many factors have had the opportunity to affect the rate estimate obtained at a single time point. Even after a number of years of measurement, causative mechanistic-relationships may be hard to identify (Shaw et al. 2002, Field et al. 2003, Mitchell and Reich 2003), unless regulatory factors exert their influence in a temporally predictive manner (e.g., litter decomposition shifting from nitrogen- to lignin-controlled as in Taylor et al. 1989). Identification of causation may be further compounded by changes induced by treatment on the system of study over time that then interact with the treatment to modify the original treatment impacts on the process of interest, as discussed next.

6.3.3.3 Synergies, feedbacks and thresholds

The magnitude and direction of treatment effects in IEEs may change over time not simply through variation in the magnitude of a regulatory factor (Figure 6.3) but because of synergies, feedbacks and thresholds (Reynolds et al. in press). Synergies arise when two or more processes/factors interact in such a way that the outcome is greater than the sum of their separate effects (e.g., multiplicative rather than additive). Feedbacks occur when a system or process is modified by changes in its own influence or size. If the initial (direct) response is enhanced, the feedback is considered positive; if decreased, the feedback is negative. Thresholds occur where critical values or set points are reached. They often result in major nonlinear changes in a process, and are products of the complex of synergies and feedbacks that exist in both climate and biological systems (Harte 1996). Thus, in short, the temporal responses of processes to treatment will be the result of variation in regulatory factors modified by the action of synergies, feedbacks and thresholds (examples are provided in Table 6.1) – a highly complex situation indeed!

How one can reconcile the potential complexity of mechanisms that result in observed treatment effects with the necessary understanding for temporally predictive science is a moot point. Schellnhuber (1999) challenges us to consider: "Where are the limits of scientific predictability in complex systems?" Holling et al. (2002) argue that it is essential to distinguish what is understandable or predictable (even if uncertain) from that which is inherently unpredictable. It is clear that prediction (i.e., temporal extrapolation) operates at the frontier of what is amenable to the scientific method today and this must be borne in mind when questions of temporal scaling are posed.

6.3.3.4 Press vs. pulse and step vs. gradual

Intact ecosystem experimental treatments can be categorized as press or pulse. Based on the definition of Bender et al. (1984), a *press* experiment is where the perturbation is sustained (often at a constant level) across time (e.g., exposure of intact communities to elevated CO_2). In contrast, *pulse* experiments involve a relatively brief period of perturbation (e.g., single application of nitrogen fertilizer). The investigator is typically interested in how a system/process deviates from its pre-treatment state (press and pulse) and then how quickly, if at all, it returns to its pre-treatment state (pulse only). With pulse experiments the transient response of a process to treatment is usually the response of most interest. For accurate quantification of process values across time, knowledge of the timing of process response to treatment is required. Williams et al. (1999) demonstrated that N_2O efflux from grassland soils amended with bovine urine was practically instantaneous (<4 h) and that the flux values observed were approximately six times higher within the first 24 hours than after them. This marked initial response was missed in earlier studies, leading to underestimates of soil N_2O flux. Where more integrative processes are measured (e.g., Pfisterer and Schmid 2002), a single process measurement may integrate across declining or increasing responses of the process to treatment over time but at a temporal grain finer than that measured. Thus, the

sensitivity of the measured process response to the pulse perturbation may be under- or over-estimated, respectively.

Table 6.1. *Examples of ecological nonlinearities on biogeochemical processes caused by thresholds, feedbacks and synergisms.*

Thresholds	*Feedbacks*	*Synergisms*
Al stress-acid deposition: The calcium/aluminum (Ca/Al) molar ratio of the soil solution is an ecological indicator of thresholds, beyond which the risk of forest damage from Al stress and nutrient imbalances increases (Cronan and Grigal 1995). Based on data from 89 forested catchments across Europe, Dise et al. (2001) found thresholds for N and S depositions, above which Al is released from forests at maximum rates.	*Herbivores-plants-N cycling*: Herbivores indirectly decelerate N cycling in savannah by decreasing the abundance of plant species with nitrogen-rich tissues. They may also decelerate succession by reducing the cover and biomass of woody plants (Ritchie et al. 1998).	*Vegetation pattern-NPP*: In arid ecosystems clumping or banding of vegetation results in an overall increase of production and plant diversity due to the combined benefits of the redistribution of precipitation and nutrients into patches (Aguiar and Sala 1999).
Weed density-productivity: When weed populations reach a threshold density, crop production is reduced (Brown et al. 1999).	*Elevated CO_2-respiration*: Plants grown in elevated CO_2 increase C input to the rhizosphere, which in turn feedbacks to affect tree growth and soil C dynamics. For example, additional soil C inputs increase rhizosphere respiration (Cheng 1999).	*Litter quality-decomposition*: Mixing litters of different nitrogen concentration has negative synergistic effects on decomposition rate (Smith and Bradford 2003).
Canopy chemistry-N cycling: Estimates of soil C:N ratios indicate that 63% of the White Mountains in New Hampshire fall below 22, a critical threshold for the onset of nitrification (Ollinger et al. 2002).	*Canopy chemistry-N cycling*: Foliar N increases with increasing net N mineralization, suggesting a positive feedback between foliar chemistry and soil N status (Ollinger et al. 2002).	*Insects-fire-C sequestration*: Climatic change will affect the frequency and intensity of pest outbreaks, which causes a considerable loss of wood and affects fire frequencies. This synergism between outbreaks, fires, and warming must be accounted for in models if realistic carbon
	Herbivore-root growth: Herbivores have a	

Table 6.1 (contd.)

Thresholds	*Feedbacks*	*Synergisms*
Plant cover-soil loss: Minor changes in shrub or grass cover resulting from climatic variability and grazing can lead to sudden increases in soil erosion (Breshears and Barnes 1999). *Bark thickness-fire susceptibility*: Forest fragmentation and intentional burning of grasslands in the dry, subhumid tropical forests of Bolivia is increasing. Hence, the frequency of fires in local forests, which did not develop with fires, is likely to increase. A bark thickness below 18 mm makes the tree species highly susceptible to fire damage and 81% of the species are below this threshold and, therefore, likely to experience cambial injury from low intensity fires; thus, biomass C storage in this region is at grave risk (Pinard and Huffman 1997).	positive feedback on root growth, stimulating growth 7× more than shoot production (Frank et al. 2002). *Logging-fires-C loss*: For the 1997-98 Indonesian fires, the largest fire disaster ever observed, forest fires primarily affected recently logged forests while primary forests or those logged long ago were less affected. These results support the hypothesis of positive feedback between logging and fire occurrence (Siegert et al. 2001), both of which affect C loss from tropical forests. *Soil N-plant growth-CO_2*: Populus grown under elevated CO_2 and low-N soil have reduced photosynthetic capacity, but in high-N soil, photosynthesis is enhanced (Zak et al. 1993).	sequestration forecasts in a warming climate are to be made (Volney and Fleming 2000). *Plant diversity-soil processes*: Two-species mixtures of ten different grassland plants had higher or lower soil microbial biomass, respiration: biomass ratio and plant litter decomposition than the additive prediction from monoculture treatments (Wardle and Nicholson 1996). *Soil nutrients-UV radiation-plant growth*: Enhanced UV-B radiation (equivalent to a 15% ozone depletion) and high soil nutrient levels interact synergistically to accelerate growth of *Phlomis fruticosa*, a fast-growing, drought-deciduous Mediterranean plant species (Levizou and Manetas 2001).

Given that the main focus of pulse experiments is to quantify transient dynamics, they tend to be less criticized than press experiments for being short-term. However, the criticism that short-term experiments may quantify transient dynamics which may be opposite to long-term effects (Tilman 1989), is equally applicable to press and pulse manipulations. These "transient dynamics" may result from ecological mechanisms that would be observed under natural circumstances, such as the crossing of critical thresholds (see Table 6.1 for examples). Alternatively, they may be artifacts of the way in which treatment is applied. For example, in the Duke

Forest FACE study, where an entire stand of loblolly pine is being exposed to elevated atmospheric CO_2 concentrations, the growth of the forest in the first three years was significantly enhanced. However, after three years this stimulation was lost, a result of the negative feedback of insufficient soil nutrients (Oren et al. 2001). In a modeling exercise, parameterized using data from the same FACE site, Luo and Reynolds (1999) demonstrated that had the treatments been applied in a gradual increment in CO_2 concentration – rather than a step from ambient to 550 ml l^{-1} – then the marked and transient increase in productivity would not have been observed. This was due to differences in forest N demand between a gradual and step increase in CO_2. Of course, many treatments are applied in such a "step" fashion – we simply do not have the time to "realistically" mimic a gradual change in CO_2 concentration, mean annual temperature, or rainfall variability, as predicted under IPCC scenarios (Houghton et al. 2001). What is important is to recognize the implications of such step changes on our ability to extrapolate results in time and how the mechanistic understanding generated aids future prediction (e.g., Luo 2001). For some treatments the step vs. gradual categorizations are less relevant, if at all. For example, one can realistically impose treatments to investigate the response of an agricultural field to long-term fertilizer application (Hütsch et al. 1993) or the novel occurrence of exotic species within ecosystems (Alvarez and Cushman 2002). Regardless of whether the "step" categorization applies, the whole concept of step vs. gradual application has stimulated important debate and experimentation into the effects of different levels of applied treatment.

The responses of processes to different levels of treatment are typically nonlinear, even when applied over the same spatial and temporal extent. For example, in a Texas grassland net N mineralization rates in soils decrease exponentially with increasing CO_2, with the largest changes occurring at subambient concentrations (Gill et al. 2002). In a separate study, Granados and Körner (2002) examined the potential for elevated CO_2 to accelerate the growth of climbing vine species of tropical forests, and hence tree turnover. In both species, the increase from 280 (pre-industrial) to 420 ml l^{-1} CO_2 had a much larger effect on growth than the increase from 420 to 560 ml l^{-1}. Further, the 560 to 700 ml l^{-1} enhancement caused a massive reduction of the stimulation. Yet most experiments use only one treatment level, potentially masking nonlinear relationships. If we wish to extrapolate treatment effects in time or space, across which the magnitude of the press or pulse of the perturbation simulated by the experimental treatment may differ, then we need to quantify the shape of the relationship between processes and treatment levels.

6.4 MEASURING TREATMENT EFFECTS USING MODEL ECOSYSTEM EXPERIMENTS

6.4.1 Laboratory and Field MEEs

We define MEEs as those where the experimental system is either wholly or partially constructed. Our definition includes three types of MEEs. In type (1) the

climate is constructed and the biotic component remains "intact" – i.e., laboratory experiments using, for example, vegetated soil monoliths that are removed from the field. In type (2) the climate is "intact" and the biotic component is constructed – i.e., field experiments where, for example, the plant community is grown from seed and controlled for richness and functional-type composition. Lastly, in type (3) both the climate and the biotic component are constructed. These are laboratory experiments using biotic systems that are not "intact", e.g., those typically conducted in controlled environment chambers (e.g., Lawton 1995, Reynolds 2001).

Type (1) MEEs isolate an "ecosystem-unit" from the rest of its landscape and introduce a suite of questions regarding the "reality" of process rates and treatment effects measured under laboratory conditions. For example, what impact will the absence of immigration have? Has the disturbance altered soil nutrient availabilities? Do diurnal patterns of temperature need to be simulated? Questions with respect to disturbance effects of the removal into the laboratory have a different basis to those that question the effects of obstructing processes such as immigration and diurnal variation in climate. Disturbance effects are non-target changes in the system of interest, whereas isolation from biotic and abiotic factors that would not be under experimental control in intact ecosystems is a deliberate attempt to maximize the relative contribution of treatment to variation in measured process rates. In short, the aim is to increase the chances of detecting a signal, if one exists, by reducing noise.

The rationale of maximizing the signal-to-noise ratio also underlies the use of type (3) MEEs. In addition, many of the treatments imposed on these MEEs may not be feasible under field conditions. For example, to manipulate the soil community in the field, and maintain it over biological meaningful temporal and spatial scales, presents many logistic difficulties (Kampichler et al. 2001). Thus, partial or complete construction of an experimental system is often necessary to test the posed hypotheses (e.g., the model soil communities created by Bradford et al. 2002 and De Deyn et al. 2003). The same "feasibility" criterion underlies the use of the majority of field MEEs. For example, construction of plant communities of set initial density and varying diversity simply would not be possible without a constructionist approach (e.g., Hooper and Vitousek 1997). Given that the biotic system has been partially constructed, which clearly is not a natural situation, the decision to conduct the experiment in the field rather than in the laboratory is often an attempt to increase the realism of the experimental system. These constructed vs. real components to MEE design are properties associated with internal and external validity.

6.4.2 Internal and External Validity

Manly (1992) stated that, "Internal validity concerns whether the apparent effects or lack of effects shown by the experimental results are due to the factor being studied, rather than some alternative factor. External validity concerns the extent to which the results of an experiment can be generalized to some wider population of interest." Internal validity is likely to be highest when all factors are under the experimenter's control. External validity will be highest when the experimental system corresponds to the natural system under investigation in all aspects (Naeem

2001) and, following Manly's definition, when the experimental plot is appropriately temporally and spatially scaled for extrapolation. Given the huge number of studies conducted using model ecosystems, and the fact that they all trade-off external for internal validity to some extent, it is appropriate to ask: What understanding we can extrapolate from MEEs to facilitate temporally and/or spatially predictive science? Maybe this very question is why some of the most well known MEEs, such as those that initially evaluated the effects of elevated CO_2 on ecosystem response (e.g., Billings et al. 1982) and the role of biodiversity in ecosystem functioning (e.g., Naeem et al. 1994), have produced much debate and experimentation to challenge/confirm the original findings.

6.4.3 Scaling MEEs – Process Rates and Treatment Effects

Our discussion of the factors that must be considered prior to extrapolating experimental results in space or time from intact systems apply equally to MEEs (see Gardner et al. 2001 for new perspectives and review of this topic). Given these obstacles, and the question of validity, when can we extrapolate process rates and absolute/relative treatment effects from MEEs? The answer is probably "almost never" (if we decide that we can, then uncertainty analysis will be an essential component; see Li and Wu, Chapter 3).

To support our contention that process rates and absolute/relative treatment effects from MEEs can almost never be extrapolated we provide the following hypothetical example. A type (1) MEE (intact system, constructed climate) is used by Dr. Climatron to investigate how soil CO_2 efflux is affected by carbon addition. Prior fieldwork by his group has demonstrated that soil moisture and temperature are regulators of CO_2 efflux and so they collect a soil monolith suitably sized to capture heterogeneity in moisture at the landscape level. Glucose solution amendments are imposed as the treatment and water amendments as the control. Carbon dioxide efflux is then measured across a range of soil moistures at different temperatures. Using field measurements of temperature, moisture and efflux, from carbon amendment and control plots, Dr. Climatron's group cross-calibrate the relationship generated in the laboratory with that in the field. They find that temperature and moisture explain most of the field variation in efflux and, as a result, can quite accurately predict the measured CO_2 efflux of the monoliths maintained in the laboratory from the imposed soil moistures and temperatures. That is, their laboratory measurements have high external validity. They then use their MEE to determine the efflux at temperatures and moistures higher and lower than those observed during their field observations, and at different levels of carbon amendment. Thus, they are quantifying relationships at values beyond those calibrated against field measurements.

This property of MEEs is a strength of the MEE approach, but how confident can we be in their extrapolation? Given the high external validity of their original laboratory measurements (given the good fit with field data) our confidence is high but we recognize that they are assessing *potential* and not *actual* process rates. Hence, whether we can scale their results to the field becomes questionable. Had they first sieved the soil and then reconstructed the monolith, before placing it in the

field (type [2] MEE) or laboratory (type [3] MEE), they would still be measuring potential effects but our confidence in whether the measured process rates reflect those of an intact system would be much less than when they used a type (1) MEE. For example, they may have altered characteristics of the system, such as drainage, through sieving which might influence efflux. Nevertheless, carbon amendment would probably still stimulate heterotrophic respiration, so we might be confident to extrapolate the sign of the treatment effect, which would be positive, to the field.

We suggest that process rates and absolute/relative treatment effects measured in MEEs should not be extrapolated to intact systems. Further, through controlling for regulators and isolating the system in the case of lab MEEs, we remove mechanisms that may contribute to feedbacks, synergisms and thresholds. Thus, the longer we run a study, or the simpler a system is, then the more we must question whether the sign of the treatment effect can be extrapolated. It is worth noting that the sign of treatment effects in relatively long-term and complex MEEs are the same when both field and laboratory MEEs are used (Naeem et al. 1994, Hector et al. 1999), despite marked differences in validity. In contrast, using simpler systems, Navas et al. (1999) demonstrated that type (3) MEEs with isolated plants could not be used to predict the behavior of species mixtures under elevated CO_2 and a nitrogen gradient, whereas monocultures could. Their work suggests that there may be critical levels of validity across which we cannot scale. It is imperative that we determine if such critical thresholds of validity exist given the very large number of MEEs conducted using single plants.

Naeem (2001) provides a more in-depth evaluation of how validity relates to extrapolation and, interestingly, concludes that the current biodiversity-productivity debate is premature given that appropriately valid/scaled experiments have not been conducted to address the question. His conclusion may be applicable to many scientific debates because of the lack of "scale awareness" in the experiments conducted by all sides to support their arguments.

6.4.4 Scaling MEEs – Mechanisms

So when are MEEs useful? We believe MEEs are best suited to identifying causation/mechanisms and become most powerful in this respect as internal validity is increased. For example, demonstration of semi-conservative replication of DNA would not have been possible without a highly simplified MEE (Meselson and Stahl 1958). Similarly, Goddard and Bradford (2003) used the MEE approach to demonstrate the potential for a population of fungi found widely in terrestrial ecosystems to evolutionarily adapt, within relatively few generations, to altered carbon and nitrogen availability in the environment. Further, the role of soil fertility in determining plant responses to elevated CO_2 was worked out using MEEs (Bazzaz and Catovsky 2002). However, at the same time internal validity increases such that our confidence in extrapolating this causation to natural systems decreases due to loss of external validity. In a thought provoking article Lawton (1995) challenges us to consider this loss as a research question as opposed to a limitation. We agree. For the purposes of extrapolation of mechanisms from MEEs to the field, we suggest that short-term, complex MEEs are best suited. If we wish to know how

important a mechanism may be in an intact system then, given that as complexity increases the signal-to-noise ratio decreases, one approach may be to identify a potential mechanism using a highly simplified system. Then, while maintaining the same temporal extent, increase system complexity and measure whether the mechanism of causation is still detectable. Unfortunately there is still a catch – weak interactions, which are hard to detect above non-controlled variation, are increasingly being shown to be important in ecological dynamics (e.g., Berlow 1999).

6.5 CONCLUSIONS

The challenges of global changes and Earth System Science are great. To extrapolate biogeochemical process rates, and treatment effects on them, across time and space greater than the experimental extent requires an understanding of the factors that regulate the process, how these factors are temporally and spatially distributed, and what feedbacks, thresholds and synergies may manifest. We cannot expect the shape of the relationship between process rates, or treatment effects, and regulators to be linear; multiple levels of treatments must be imposed across varying values of regulatory factors. We need to determine if the process being measured is integrative or active within the specific ecosystem under investigation and whether the treatment responses observed are artifacts of the way treatment is imposed (e.g., step vs. gradual).

Spatial and temporal extrapolation differs in a key regard. The latter attempts to predict the future behavior of ecosystems and is inherently more complex due to the greater number of uncertainties and interactions that act across temporal as opposed to spatial scales. Research intent on extrapolation might then be best focused on spatial issues but, given the societal need for temporally predictive science, well-resourced multi-disciplinary studies that critically evaluate their own limitations will be required to tackle temporal scaling. Model ecosystem experiments will contribute to this effort and their use will be best targeted at unravelling the mechanisms behind causation of treatment effects observed in intact systems. Neither intact systems nor model systems alone will provide the necessary understanding required to scale experimental results across time and space, and discussion (Carpenter 1996, Carpenter 1999, Drenner and Mazumder 1999) of which is superior is constructive only if to highlight this point. There is a need for much greater "scale awareness" in ecology and this is reflected both in this and related volumes (e.g., Gardner et al. 2001).

ACKNOWLEDGEMENTS

This research was supported by NSF grant DEB-02-12123, USDA grant #58-1270-3-070, and is a contribution of the Duke Phytotron (NSF-IBN-99-85841).

REFERENCES

Aerts, R. 1997. Climate, leaf litter chemistry and leaf litter decomposition in terrestrial ecosystems: a triangular relationship. Oikos 79:439-449.

Aguiar, M. R., and O. E. Sala. 1999. Patch structure, dynamics and implications for the functioning of arid ecosystems. Trends in Ecology and Evolution 14:273-277.

Alvarez, M. E., and J. H. Cushman. 2002. Community-level consequences of a plant invasion: effects on three habitats in coastal California. Ecological Applications 12:1434-1444.

Barrett, J. E., and I. C. Burke. 2000. Potential nitrogen immobilization in grassland soils across a soil organic matter gradient. Soil Biology & Biochemistry 32:1707-1716.

Bazzaz, F. A., and S. Catovsky. 2002. Plants – from cells to ecosystems: impacts of global environmental change. Pages 94-111 *in* H. A. Mooney and J. G. Canadell, editors. Encyclopaedia of Global Environmental Change. Wiley, Chichester, UK.

Bender, E. A., T. J. Case, and M. E. Gilpin. 1984. Perturbation experiments in community ecology: theory and practice. Ecology 65:1-13.

Berlow, E. L. 1999. Strong effects of weak interactions in ecological communities. Nature 398:330-334.

Billings, W. D., J. O. Luken, D. A. Mortensen, and K. M. Peterson. 1982. Arctic tundra: a source or sink for atmospheric carbon dioxide in a changing environment. Oecologia 53:7-11.

Bradford, M. A., T. H. Jones, R. D. Bardgett, H. I. J. Black, B. Boag, M. Bonkowski, R. Cook, T. Eggers, A. Gange, S. J. Grayston, E. Kandeler, A. E. McCaig, J. N. Newington, J. I. Prosser, H. Setälä, P. L. Staddon, G. M. Tordoff, D. Tscherko, and J. H. Lawton. 2002. Impacts of soil faunal community composition on model grassland ecosystems. Science 298:615-618.

Bradford, M. A., P. A. Wookey, P. Ineson, and H. M. Lappin-Scott. 2001. Controlling factors and effects of chronic nitrogen and sulphur deposition on methane oxidation in a temperate forest soil. Soil Biology & Biochemistry 33:93-102.

Breshears, D. D., and F. J. Barnes. 1999. Interrelationships between plant functional types and soil moisture heterogeneity for semiarid landscapes within the grassland/forest continuum: a unified conceptual model. Landscape Ecology 14:465-478.

Brown, J. R., J. Herrick, and D. Price. 1999. Managing low-output agroecosystems sustainably: the importance of ecological thresholds. Canadian Journal of Forest Research 29:1112-1119.

Bubier, J., P. Crill, and A. Mosedale. 2002. Net ecosystem CO_2 exchange measured by autochambers during the snow-covered season at a temperate peatland. Hydrological Processes 16:3667-3682.

Carpenter, S. R. 1996. Microcosm experiments have limited relevance for community and ecosystem ecology. Ecology 77:677-680.

Carpenter, S. R. 1999. Microcosm experiments have limited relevance for community and ecosystem ecology: reply. Ecology 80:1085-1088.

Cheng, W. X. 1999. Rhizosphere feedbacks in elevated CO_2. Tree Physiology 19:313-320.

Clark, J. S., S. Carpenter, M. Barber, S. Collins, A. Dobson, J. A. Foley, D. M. Lodge, M. Pascual, R. Pielke, Jr., W. Pizer, C. Pringle, W. V. Reid, K. A. Rose, O. Sala, W. H. Schlesinger, D. H. Wall, and D. Wear. 2001. Ecological forecasts: an emerging imperative. Science 293:657-660.

Committee on Global Change. 1988. Toward an understanding of global change. National Academy Press, Washington, D.C.

Cronan, C. S., and D. F. Grigal. 1995. Use of calcium/aluminum ratios as indicators of stress in forest ecosystems. Journal of Environmental Quality 24:209-226.

De Deyn, G. B., C. E. Raaijmakers, H. R. Zoomer, M. P. Berg, P. C. de Ruiter, H. A. Verhoef, T. M. Bezemer, and W. H. van der Putten. 2003. Soil invertebrate fauna enhances grassland succession and diversity. Nature 422:711-713.

DeLucia, E. H., J. G. Hamilton, S. L. Naidu, R. B. Thomas, J. A. Andrews, A. Finzi, M. Lavine, R. Matamala, J. E. Mohan, G. R. Hendrey, and W. H. Schlesinger. 1999. Net primary production of a forest ecosystem with experimental CO_2 enrichment. Science 284:1177-1179.

Dise, N. B., E. Matzner, M. Armbruster, and J. MacDonald. 2001. Aluminum output fluxes from forest ecosystems in Europe: a regional assessment. Journal of Environmental Quality 30:1747-1756.

Drenner, R. W., and A. Mazumder. 1999. Microcosm experiments have limited relevance for community and ecosystem ecology: comment. Ecology 80:1081-1085.

Fang, C., and J. B. Moncrieff. 2001. The dependence of soil CO_2 efflux on temperature. Soil Biology & Biochemistry 33:155-165.

Fernández, R. J., and J. F. Reynolds. 2000. Potential growth and drought tolerance of eight desert grasses: lack of a trade-off? Oecologia 123:90-98.

Field, C. B., M. R. Shaw, H. A. Mooney, E. S. Zavaleta, N. R. Chiariello, and E. E. Cleland. 2003. Assessing environmental changes in grasslands. Science 299:1844-1845.

Frank, D. A., M. M. Kuns, and D. R. Guido. 2002. Consumer control of grassland plant production. Ecology 83:602-606.

Gardner, R. H., W. M. Kemp, V. S. Kennedy, and J. E. Petersen, editors. 2001. Scaling Relations in Experimental Ecology. Columbia University Press, New York.

Gill, R. A., H. W. Polley, H. B. Johnson, L. J. Anderson, H. Maherali, and R. B. Jackson. 2002. Nonlinear grassland responses to past and future atmospheric CO_2. Nature 417:279-282.

Goddard, M. R., and M. A. Bradford. 2003. The adaptive response of a natural microbial population to carbon- and nitrogen-limitation. Ecology Letters 6:594-598.

González, G., and T. R. Seastedt. 2001. Soil fauna and plant litter decomposition in tropical and subalpine forests. Ecology 82:955-964.

Goudriaan, J., H. H. Shugart, H. Bugmann, W. Cramer, A. Bondeau, R. H. Gardner, L. A. Hunt, W. K. Lauenroth, J. J. Landsberg, S. Linder, I. R. Noble, W. J. Parton, L. F. Pitelka, M. Stafford Smith, R. W. Sutherst, C. Valentin, and F. I. Woodward. 1999. Use of models in global change studies. Pages 106-140 *in* B. H. Walker, W. L. Steffen, J. Canadell, and J. S. I. Ingram, editors. Global Change and the Terrestrial Biosphere: Implications for Natural and Managed Ecosystems. A Synthesis of GCTE and Related Research, IGBP Book Series No. 4. Cambridge University, Cambridge.

Granados, J., and C. Körner. 2002. In deep shade, elevated CO_2 increases the vigor of tropical climbing plants. Global Change Biology 8:1109-1117.

Harte, J. 1996. Feedbacks, thresholds and synergies in global change: population as a dynamic factor. Biodiversity and Conservation 5:1069-1083.

Hector, A., B. Schmid, C. Beierkuhnlein, M. C. Caldeira, M. Diemer, P. G. Dimitrakopoulos, J. A. Finn, F. H, P. S. Giller, J. Good, R. Harris, P. Högberg, K. Huss-Danell, J. Joshi, A. Jumpponen, C. Körner, P. W. Leadley, M. Loreau, A. Minns, C. P. H. Mulder, G. O'Donovan, S. J. Otway, J. S. Pereira, A. Prinz, D. J. Read, M. Scherer-Lorenzen, E.-D. Schulze, A.-S. D. Siamantziouras, E. M. Spehn, A. C. Terry, A. Y. Troumbis, F. I. Woodward, S. Yachi, and J. H. Lawton. 1999. Plant diversity and productivity experiments in European grasslands. Science 286:1123-1127.

Holling, C. S., L. H. Gunderson, and D. Ludwig. 2002. In quest of a theory of adaptive change. Pages 3-22 *in* L. H. Gunderson and C. S. Holling, editors. Panarchy: Understanding Transformations in Human and Natural Systems. Island Press, Washington D.C.

Hooper, D. U., and P. M. Vitousek. 1997. The effects of plant composition and diversity on ecosystem processes. Science 277:1302-1305.

Houghton, J. T., Y. Ding, D. J. Griggs, M. Noguer, P. J. van der Linden, X. Dai, K. Maskell, and C. A. Johnson, editors. 2001. Climate Change 2001: The Scientific Basis. Cambridge University Press, Cambridge.

Hütsch, B. W., C. P. Webster, and D. S. Powlson. 1993. Long-term effects of nitrogen-fertilization on methane oxidation in soil of the Broadbalk Wheat Experiment. Soil Biology & Biochemistry 25:1307-1315.

Kampichler, C., A. Bruckner, and E. Kandeler. 2001. The use of enclosed model ecosystems in soil ecology: a bias towards laboratory research. Soil Biology & Biochemistry 33:269-275.

Kemp, W. M., J. E. Petersen, and R. H. Gardner. 2001. Scale-dependence and the problem of extrapolation. Pages 3-57 *in* R. H. Gardner, W. M. Kemp, V. S. Kennedy, and J. E. Petersen, editors. Scaling Relations in Experimental Ecology. Columbia University Press, New York.

Lawton, J. H. 1995. Ecological experiments with model systems. Science 269:328-331.

Lawton, J. H. 1996. The Ecotron facility at Silwood Park: the value of "big bottle" experiments. Ecology 77:665-669.

Lawton, J. H. 2001. Earth system science. Science 292:1965.

Levizou, E., and Y. Manetas. 2001. Combined effects of enhanced UV-B radiation and additional nutrients on growth of two Mediterranean plant species. Plant Ecology 154:181-186.

Loreau, M., S. Naeem, P. Inchausti, J. Bengtsson, J. P. Grime, A. Hector, D. U. Hooper, M. A. Huston, D. Raffaelli, B. Schmid, D. Tilman, and D. A. Wardle. 2001. Biodiversity and ecosystem functioning: current knowledge and future challenges. Science 294:804-808.

Luo, Y. 2001. Transient ecosystem responses to free-air CO_2 enrichment (FACE): experimental evidence and methods of analysis. New Phytologist 152:1-8.

Luo, Y., and J. F. Reynolds. 1999. Validity of extrapolating field CO_2 experiments to predict carbon sequestration in natural ecosystems. Ecology 80:1568-1583.

Manly, B. F. J. 1992. The Design and Analysis of Research Studies. Cambridge University Press, Cambridge.

McGrady-Steed, J., P. M. Harrison, and P. J. Morin. 1997. Biodiversity regulates ecosystem predictability. Nature 390:162-165.

Meselson, M., and F. W. Stahl. 1958. The replication of DNA in *Escherichia coli*. Proceedings of the National Academy of Sciences of the United States of America 44:671-682.

Mitchell, C. E., and P. B. Reich. 2003. Assessing environmental changes in grasslands. Science 299:1844.

Mittelbach, G. G., C. F. Steiner, S. M. Scheiner, K. L. Gross, H. L. Reynolds, R. B. Waide, M. R. Willig, S. I. Dodson, and L. Gough. 2001. What is the observed relationship between species richness and productivity? Ecology 82:2381-2396.

Naeem, S. 2001. Experimental validity and ecological scale as criteria for evaluating research programs. Pages 223-250 *in* R. H. Gardner, W. M. Kemp, V. S. Kennedy, and J. E. Petersen, editors. Scaling Relations in Experimental Ecology. Columbia University Press, New York.

Naeem, S., L. J. Thompson, S. P. Lawler, J. H. Lawton, and R. M. Woodfin. 1994. Declining biodiversity can alter the performance of ecosystems. Nature 368:734-737.

Navas, M. L., E. Garnier, M. P. Austin, and R. M. Gifford. 1999. Effect of competition on the responses of grasses and legumes to elevated atmospheric CO_2 along a nitrogen gradient: differences between isolated plants, monocultures and multi-species mixtures. New Phytologist 143:323-331.

Neff, J. C., A. R. Townsend, G. Gleixner, S. J. Lehman, J. Turnbull, and W. D. Bowman. 2002. Variable effects of nitrogen additions on the stability and turnover of soil carbon. Nature 419:915-917.

Ollinger, S. V., M. L. Smith, M. E. Martin, R. A. Hallett, C. L. Goodale, and J. Aber. 2002. Regional variation in foliar chemistry and N cycling among forests of diverse history and composition. Ecology 83:339-355.

Oren, R., D. S. Ellsworth, K. H. Johnsen, N. Phillips, B. E. Ewers, C. Maier, K. V. R. Schäfer, H. McCarthy, G. Hendrey, S. G. McNulty, and G. G. Katul. 2001. Soil fertility limits carbon sequestration by forest ecosystems in a CO_2-enriched atmosphere. Nature 411:469-472.

Pfisterer, A. B., and B. Schmid. 2002. Diversity-dependent production can decrease the stability of ecosystem functioning. Nature 416:84-86.

Pielke, R. A., Sr. 2002. Overlooked issues in the U.S. national climate and IPCC assessments. Climatic Change 52:1-11.

Pinard, M. A., and J. Huffman. 1997. Fire resistance and bark properties of trees in a seasonally dry forest in eastern Bolivia. Journal of Tropical Ecology 13:727-740.

Prinn, R., H. Jacoby, A. Sokolov, C. Wang, X. Xiao, Z. Yang, R. Eckhaus, P. Stone, D. Ellerman, J. Melillo, J. Fitzmaurice, D. Kicklighter, G. Holian, and Y. Liu. 1999. Integrated global system model for climate policy assessment: feedbacks and sensitivity studies. Climatic Change 41:469-546.

Reich, P. B., J. Knops, D. Tilman, J. Craine, D. Ellsworth, M. Tjoelker, T. Lee, D. Wedin, S. Naeem, D. Bahauddin, G. Hendrey, S. Jose, K. Wrage, J. Goth, and W. Bengston. 2001. Plant diversity enhances ecosystem responses to elevated CO_2 and nitrogen deposition. Nature 410:809-812.

Reynolds, J. F. 2001. The role of controlled environment facilities in global change research. Pages 228-231 *in* H. A. Mooney and J. Canadell, editors. The Earth System: Biological and Ecological Dimensions of Global Environmental Change. Wiley, New York.

Ritchie, M. E., D. Tilman, and J. M. H. Knops. 1998. Herbivore effects on plant and nitrogen dynamics in oak savanna. Ecology 79:165-177.

Schellnhuber, H. J. 1999. 'Earth system' analysis and the second Copernican revolution. Nature 402 (Supp):C19-C23.

Schlesinger, W. H., J. A. Raikes, A. E. Hartley, and A. E. Cross. 1996. On the spatial pattern of soil nutrients in desert ecosystems. Ecology 77:364-374.

Shaver, G. R., L. C. Johnson, D. H. Cades, G. Murray, J. A. Laundre, E. B. Rastetter, K. J. Nadelhoffer, and A. E. Giblin. 1998. Biomass and CO_2 flux in wet sedge tundras: responses to nutrients, temperature, and light. Ecological Monographs 68:75-97.

Shaw, M. R., E. S. Zavaleta, N. R. Chiariello, E. E. Cleland, H. A. Mooney, and C. B. Field. 2002. Grassland responses to global environmental changes are suppressed by elevated CO_2. Science 298:1987-1990.

Siegert, F., G. Rücker, A. Hinrichs, and A. A. Hoffmann. 2001. Increased damage from fires in logged forests during droughts caused by El Niño. Nature 414:437-440.

Smith, S. D., T. E. Huxman, S. F. Zitzer, T. N. Charlet, D. C. Housman, J. S. Coleman, L. K. Fenstermaker, J. R. Seeman, and R. S. Nowak. 2000. Elevated CO_2 increases productivity and invasive species success in an arid ecosystem. Nature 408:79-82.

Smith, V. C., and M. A. Bradford. 2003. Do non-additive effects on decomposition in litter-mix experiments result from differences in resource quality between litters? Oikos 102:235-242.

Taylor, B. R., D. Parkinson, and W. F. J. Parsons. 1989. Nitrogen and lignin content as predictors of litter decay rates: a microcosm test. Ecology 70:97-104.

Tilman, D. 1989. Ecological experimentation: strengths and conceptual problems. Pages 136-157 *in* G. E. Likens, editor. Long-Term Studies in Ecology: Approaches and Alternatives. Springer-Verlag, New York.

Tilman, D., D. Wedin, and J. Knops. 1996. Productivity and sustainability influenced by biodiversity in grassland ecosystems. Nature 379:718-720.

Volk, M., P. A. Niklaus, and C. Körner. 2000. Soil moisture effects determine CO_2 responses of grassland species. Oecologia 125:380-388.

Volney, W. J. A., and R. A. Fleming. 2000. Climate change and impacts of boreal forest insects. Agriculture, Ecosystems & Environment 82:283-294.

Walker, B. H., and W. L. Steffen, editors. 1996. Global Change and Terrestrial Ecosystems. Cambridge University Press, New York.

Wardle, D. A., and K. S. Nicholson. 1996. Synergistic effects of grassland plant species on soil microbial biomass and activity: implications for ecosystem-level effects of enriched plant diversity. Functional Ecology 10:410-416.

Wardle, D. A., G. M. Barker, G. W. Yeates, K. I. Bonner, and A. Ghani. 2001. Introduced browsing mammals in New Zealand natural forests: aboveground and belowground consequences. Ecological Monographs 71:587-614.

Williams, D. L., P. Ineson, and P. A. Coward. 1999. Temporal variations in nitrous oxide fluxes from urine-affected grassland. Soil Biology & Biochemistry 31:779-788.

Zak, D. R., K. S. Pregitzer, P. S. Curtis, J. A. Teeri, R. Foge, and D. L. Randlett. 1993. Elevated atmospheric CO_2 and feedback between carbon and nitrogen cycles. Plant and Soil 151:105-117.

Zutter, B. R., G. R. Glover, R. J. Mitchell, and D. H. Gjerstad. 1997. Response of loblolly pine and sweetgum to intra- and inter-specific competition and influence of soil organic matter. Canadian Journal of Forest Research 27:2079-2087.

CHAPTER 7

A FRAMEWORK AND METHODS FOR SIMPLIFYING COMPLEX LANDSCAPES TO REDUCE UNCERTAINTY IN PREDICTIONS

DEBRA P.C. PETERS, JIN YAO, LAURA F. HUENNEKE, ROBERT P. GIBBENS, KRIS M. HAVSTAD, JEFFREY E. HERRICK, ALBERT RANGO, AND WILLIAM H. SCHLESINGER

7.1. INTRODUCTION

Many of our most pressing ecological problems, such as the conservation of biodiversity, spread of invasive species, patterns in carbon sequestration, and impacts of disturbances (e.g., fire) must be addressed at the landscape scale (see Law et al., Chapter 9, Groffman et al., Chapter 10, Urban et al., Chapter 13). However, much of our information about these problems comes from plot-scale studies that must be extrapolated to the landscape. Because landscapes are complex, this extrapolation is not always straightforward or easy to accomplish (Turner et al. 1989a, Wu and Li, Chapter 2, Braford and Reynolds, Chapter 6). Landscape complexity results from the processes, factors, and their interactions that occur across a range of spatial and temporal scales. The problem is further complicated by the presence of contagious or neighborhood processes that connect different parts of a landscape. Dispersal of seeds by wind or animals, fire, and erosion and deposition of soil and nutrients by wind and water are examples of spatial or contagious processes that influence ecosystem dynamics. Landscape complexity makes it difficult to understand and predict ecosystem dynamics across spatial scales with high levels of confidence or certainty. Our goal is to develop a conceptual framework and operational approach to simplifying complex landscapes in order to minimize both prediction errors and costs associated with measurement, analysis, and prediction.

A number of methods are available to extrapolate information that differ in the key processes involved (King 1991, Jarvis 1995). There are three main classes of extrapolation methods: (1) nonspatial, (2) spatially implicit, and (3) spatially explicit (Peters et al. 2004). These methods differ in the amount of spatial information

J. Wu, K.B. Jones, H. Li, and O.L. Loucks (eds.),
Scaling and Uncertainty Analysis in Ecology: Methods and Applications, 131–146.

required to carry out the analysis. In most cases, the objective of the extrapolation is to obtain a single estimate for an entire landscape.

Nonspatial methods are the simplest and contain the fewest parameters. These methods include linear extrapolation where fine-scale information is extrapolated to broad-scales using weighted averages based on the area covered by each type of landscape unit. The classic example is the extrapolation of net primary production from sampled plots to biomes (Leith and Whittaker 1975). Other extrapolation techniques are possible (King 1991). In each case, it is assumed that spatial location on a map and quantification of contagious processes are not needed for the extrapolation.

Spatially implicit methods include the importance of spatial location in both the input and response variables. For example, gap models that simulate grassland or forest successional dynamics (Peters 2002, Keane et al. 2001, Symstad et al. 2003), nutrient cycling models (e.g., Burke et al. 1991, 1997), and most biogeographic models currently used to predict vegetation types at regional to global scales (e.g., Neilson and Running 1996, Melillo et al. 1995) are spatially implicit methods. These models typically simulate grid cells that differ in properties such as soil texture, precipitation, and temperature. Simulations are conducted for each grid cell containing a unique combination of parameters. Spatial location is important to the extrapolation because location is used to determine the value of some parameters, but it is assumed that the important processes occur within a grid cell; thus connections among grid cells are assumed to be negligible.

Spatially explicit or interactive methods are the most complex in that they require information on spatial location as well as on neighborhood processes. Familiar examples of spatially explicit models include cellular automata (Hogeweg 1988), dispersal models that compute dispersal likelihood in terms of the distance between the target and source sites (Coffin and Lauenroth 1989, Clark et al. 1998, Rastetter et al. 2003), and models of contagious disturbances such as fire and disease (Turner et al. 1989b, Miller and Urban 1999). In each case, simulations are conducted for grid cells that differ in properties such as soil texture and climate. Furthermore, both processes within and among grid cells are important to ecosystem dynamics. Parameter values of a grid cell may depend upon either the identity of its neighboring cells, or specific exchanges of material or individuals among neighboring cells may be modeled explicitly.

Each of the three classes of scaling methods has tradeoffs in errors associated with uncertainty. Studies of model error have shown that simple models are often optimal when information is imprecise (O'Neill 1979, Reynolds and Acock 1985). However, more complicated models may be better when dynamics are complex and extensive data are available; yet these data may be expensive to collect and contain a number of small errors that accumulate to produce disproportionately large uncertainties in predictions (Gardner et al. 1980, Li and Wu, Chapter 3). Thus, there are relative trade-offs between errors of omission (high in simple models, low in complex models) and errors of commission (high in complex models, low in simple models) for each method. In general, one should select the simplest method possible that represents the key processes influencing system dynamics in order to minimize both types of error (Peters et al. 2004).

Typically, researchers select one method for an entire landscape that depends on the question being addressed. However, the use of one method likely results in high errors of omission for some parts of a landscape, and high errors of commission for other parts. For example, a nonspatial extrapolation will result in high errors of omission for the areas on the landscape where contagious processes are particularly important, such as topographic lows where water accumulates and production is higher than the landscape average. Similarly, using a spatially explicit method for an entire landscape will result in high errors of commission associated with including unnecessary and poorly estimated parameters for those areas where spatial location and contagious processes are relatively unimportant, such as level uplands where dynamics are best explained by precipitation and soil texture. Because each parameter has an associated uncertainty in its estimate, including unnecessary parameters increases the overall uncertainty of the prediction (Peters et al. 2004).

Because landscapes consist of a mosaic of sites differing in spatial heterogeneity and degree of connectedness, we expect that a combination of scaling methods is needed to simplify complex landscapes in order to minimize errors of prediction. This general approach is similar to hierarchical scaling strategies (Wu 1999). Linear extrapolations may be most appropriate for the parts of a landscape that are relatively homogeneous. Spatially implicit or explicit approaches are expected to be necessary for those parts with high spatial heterogeneity or connectedness with neighboring sites (Peters et al. 2004).

We focus on the important and timely problem of scaling patterns in carbon sequestration and dynamics across semiarid and arid ecosystems to illustrate our approach of combining these methods to simplify landscapes. Recent estimates suggest that the carbon sink in grasslands and shrublands in the coterminous U.S. from 1980-1990 may be similar to that in forests (Pacala et al. 2001). In particular, shrub-dominated ecosystems are important contributors to carbon sinks due both to their extensive area (44% of the total land area of the U.S.) and to their high potential sequestration rates (Hibbard et al. 2001). The area dominated by shrubs and other woody plants has increased worldwide over the past century because of complex interactions among a number of factors, including effects of large and small animals, drought, fire, climate change, and changes in soil properties (Humphrey 1958, Schlesinger at al. 1990, Allred 1996, Van Auken 2000). Increases in above- and belowground carbon storage as well as increases in emissions of NO_x and non-methane hydrocarbons (e.g., terpenes, isoprene, and other aromatics) have resulted from the replacement of grasses by shrubs (Archer et al. 2001, Hartley and Schlesinger 2001, Jackson et al. 2002).

Estimates for carbon sinks and losses in areas encroached upon by woody plants have a high degree of uncertainty because of landscape-scale variation in edaphic and topographic factors (Pacala et al. 2001, Hurtt et al. 2002). Furthermore, spatial patterns in carbon and other soil nutrients may be complex because of processes such as wind and water erosion, and animal redistribution of plant material and nutrients (Schlesinger and Pilmanis 1998). Our specific objectives were: (1) to illustrate the use of each of the three scaling methods for extrapolating estimates of carbon dynamics based on aboveground net primary production (ANPP) for arid and semiarid landscapes, (2) to examine the key processes and factors leading to

heterogeneity in carbon dynamics at the landscape scale, and (3) to develop a framework to identify the landscape locations where each scaling method is most appropriate.

7.2 SYSTEM DESCRIPTION

The study was conducted using data collected from the Jornada Basin Long Term Ecological Research site (JRN) located in southern New Mexico (32.5°N, 106.8°W). The Jornada consists of the Jornada Experimental Range, a 78,266-ha area administered by the USDA Agricultural Research Service, and the adjacent Chihuahuan Desert Rangeland Research Center, a 25,900-ha area administered by New Mexico State University. The JRN is characteristic of the northern Chihuahuan Desert with long term (80 y) mean annual precipitation of 248 mm/y (SD = 87) and mean monthly temperatures ranging from 3.8°C in January to 26.1°C in July. Elevation ranges from 1200m in the basin to >2500 m in the mountains.

Similar to many other arid and semiarid ecosystems, a key characteristic of the JRN is that much of the area has changed from perennial grasslands to shrublands within the past 100 years (Buffington and Herbel 1965, Gibbens and Beck 1988, Fredrickson et al. 1998). In many areas within the JRN basin, black grama (*Bouteloua eriopoda*) dominated grasslands have been replaced by one of three shrub species: honey mesquite (*Prosopis glandulosa*), creosote bush (*Larrea tridentata*), and tarbush (*Fluorensia cernua*). Grasslands dominated by tobosa (*Hilaria mutica*) commonly occur in low-lying areas. Currently at the JRN, communities dominated by these five species occur on >90% of the study area. Subdominant plants include annuals and other perennial grasses, forbs, subshrubs, and other shrubs.

7.3 EXTRAPOLATION OF ANPP FROM PLOTS TO A LANDSCAPE

We used aboveground net primary productivity (ANPP) sampled seasonally in three exclosures in each of the five major vegetation types as our plot-level estimates of changes in carbon storage (Huenneke et al. 2001, 2002). Within each vegetation type, exclosures were selected to represent the range of variability in production of that type rather than as replicates of average conditions. Each exclosure was sampled using 49 1-m^2 quadrats. Methods of sampling are described in detail in Huenneke et al. (2001, 2002). Annual values of ANPP from 1990-1998 were averaged across exclosures and years to obtain a long term estimate for each vegetation type (http://jornada-www.nmsu.edu). We then used one of three methods (nonspatial, spatially implicit, spatially explicit; Peters et al. 2004) to illustrate how to extrapolate these plot-level estimates to the entire Jornada landscape.

7.3.1 Nonspatial

The average plot-based estimate of ANPP for each of the five vegetation types was extrapolated nonspatially to the landscape scale using a weighted-averaging method

(Table 7.1). Average ANPP from plot-scale estimates ranged from 96 g m^{-2} y^{-1} in a shrub-dominated area (tarbush) to 229 g m^{-2} y^{-1} in an upland perennial grassland (black grama). These ANPP values for each vegetation type were weighted by the area associated with that type using an eight-ha resolution map generated from field surveys conducted in 1998. Most of the area (90%) is dominated by one of two shrubs (creosote bush, honey mesquite). Only 2% of the area is currently dominated by the perennial grass, black grama, a species that historically dominated much of the area. Using this weighted averaging method, the average ANPP for the JRN is 143 g m^{-2} y^{-1} during the period of sampling (1990 to 1998).

Table 7.1. *Vegetation type by ANPP (g m^{-2} from Huenneke et al. 2002) and area to obtain weighted average for the JRN. ANPP is long-term average and standard deviation of all dates and years. Area extent (ha) obtained from Gibbens et al. (in prep). Areas not dominated by these vegetation types (others) were excluded from the analysis.*

Vegetation type	*Mean ANPP (g m^{-2} y^{-1}) [standard deviation]*	*Areal extent (ha)*
Black grama grasslands	229 [114]	699
Creosotebush shrublands	139 [51]	14,485
Honey mesquite shrublands	140 [60]	34,387
Tarbush shrublands	96 [20]	3,826
Tobosa playa grasslands	194 [214]	844
Weighted average	143	

Nonspatial extrapolation of ecosystem variables from plots to larger areas is useful for coarse-scale comparisons where heterogeneity within landscape units is less important than large-scale patterns. For example, comparisons of biomes often use nonspatial extrapolations (Lieth and Whittaker 1975, Webb et al. 1978, Knapp and Smith 2001). Tabular estimates of ANPP for each vegetation type also allow comparisons with similar types of vegetation within the region as well as with other types of grasslands and shrublands (Lauenroth 1979, Le Houérou et al. 1988). However, nonspatial methods have limited utility when dealing with specific parts of a heterogeneous landscape where variation in ANPP is high (Huenneke et al. 2001, 2002). For example, grazing management that assumes a constant, uniform estimate of ANPP for an entire landscape will result in over-use in areas with below-average ANPP and under-use in areas with high ANPP.

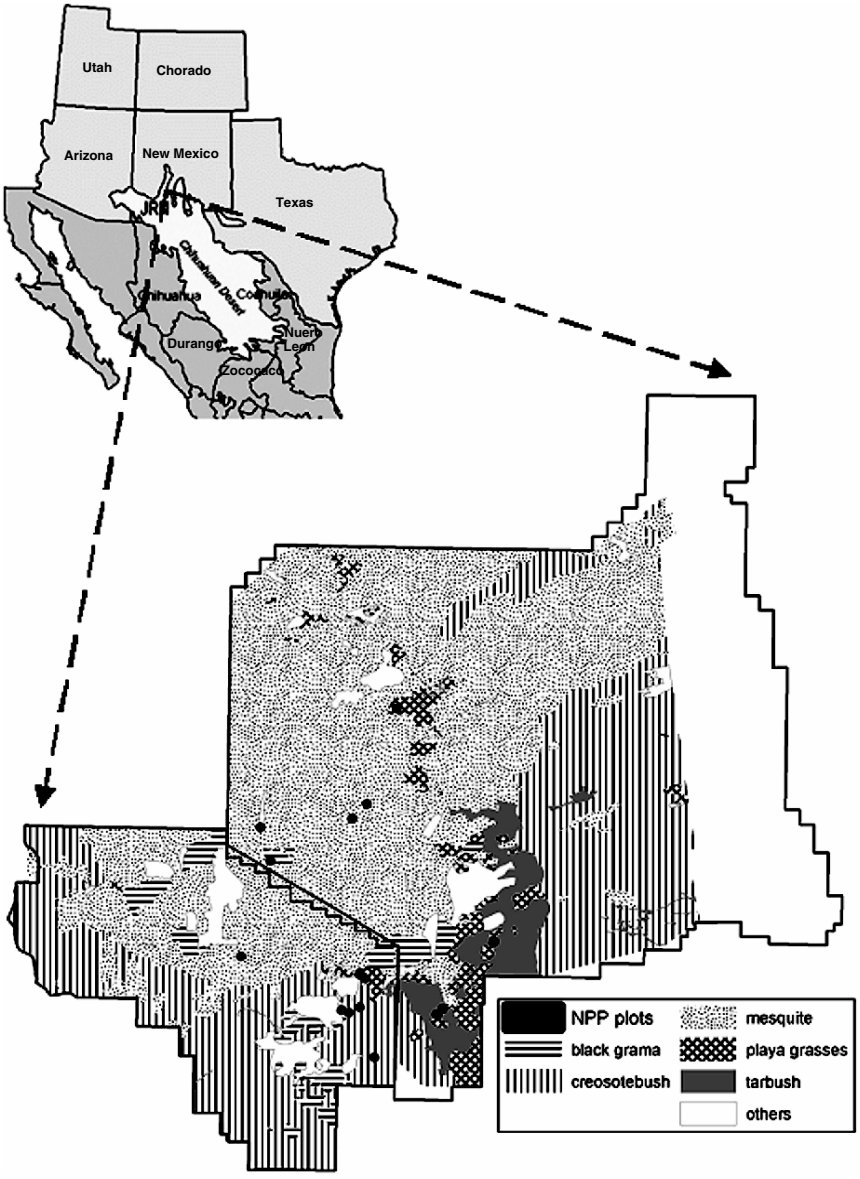


Figure 7.1. *Site map showing location of Jornada LTER within New Mexico and the U.S. Study site insert shows location of the 15 areas sampled for ANPP.*

7.3.2 Spatially Implicit

Spatially implicit methods combine plot-scale estimates of carbon with spatial databases in a geographic information system (GIS). For the JRN, we extrapolated the ANPP estimate for each vegetation type (Table 7.1) using the vegetation map of 1998 (Figure 7.1). This spatially implicit map shows the spatial distribution of ANPP across the JRN (Figure 7.2). Although total ANPP is the same as for the nonspatial approach, large-scale patterns are evident that cannot be discerned from a tabular format (Table 7.1). For example, remnant grassland areas dominated by black grama are located primarily in the west and south whereas tobosa grasslands are located along a previous channel of the Rio Grande that went through the center of the JRN from northwest to southeast ca. 1.6 million years ago (Mack et al. 1996). Low productivity tarbush sites are mostly located in the southeast.

Another spatially implicit method could be used that includes spatial variation in environmental factors in the extrapolation. Plot-scale measures of soil texture, elevation, and precipitation could be used with ANPP estimates to develop a regression equation for each vegetation type. Spatial maps of these same variables (soil texture, elevation, and precipitation) could then be used with the regression equations to extrapolate ANPP across the landscape (Figure 7.3). Although we have not conducted this analysis, this spatially implicit approach would provide a more spatially resolved map than the previous example (Figure 7.2), and would account for potential variation in ANPP as related to variation in environmental factors.

Spatially implicit methods have been used frequently in arid and semiarid landscapes where environmental heterogeneity is often recognized as important. Spatial variation in ANPP has been documented because of variation in elevation and soil properties that likely affect water availability, although the redistribution of water was not actually measured (Ludwig 1986, 1987). Patterns in other properties of vegetation have been found associated with landforms, microtopography, and soils (Stein and Ludwig 1979, Wieranga et al. 1987, Wondzell et al. 1990, 1996, Wondzell and Ludwig 1995). Soil properties, including carbon, are often extrapolated from soil pits and field surveys selected to represent characteristic locations on a landscape (Gile et al. 1981). Although many of these earlier efforts did not publish maps, all of the sampling methods were spatially implicit in that the design was stratified by the environmental variation and the results were extrapolated to similar locations on the landscape.

Spatially implicit methods of extrapolation are increasingly used as the availability of spatial databases and GIS analyses increases. Recent examples include the extrapolation of above- and belowground carbon pools across the JRN landscape from 1858 to present using maps of soils and precipitation as inputs to the CENTURY simulation model (Mitchell et al. 2002). Spatial variation in field estimates of carbon pools have also been extrapolated to the landscape scale using maps of soils and landforms (Monger et al. 2006). Spatial variation in shrub invasion and loss of perennial grasses with implications for changes in biomass quantity and vertical distribution through time have also been related to maps of soil texture, precipitation, elevation, and other factors (Yao et al. 2002). Identification of

landscape locations where shrub invasion has occurred most rapidly allows management efforts to focus on these sensitive areas.

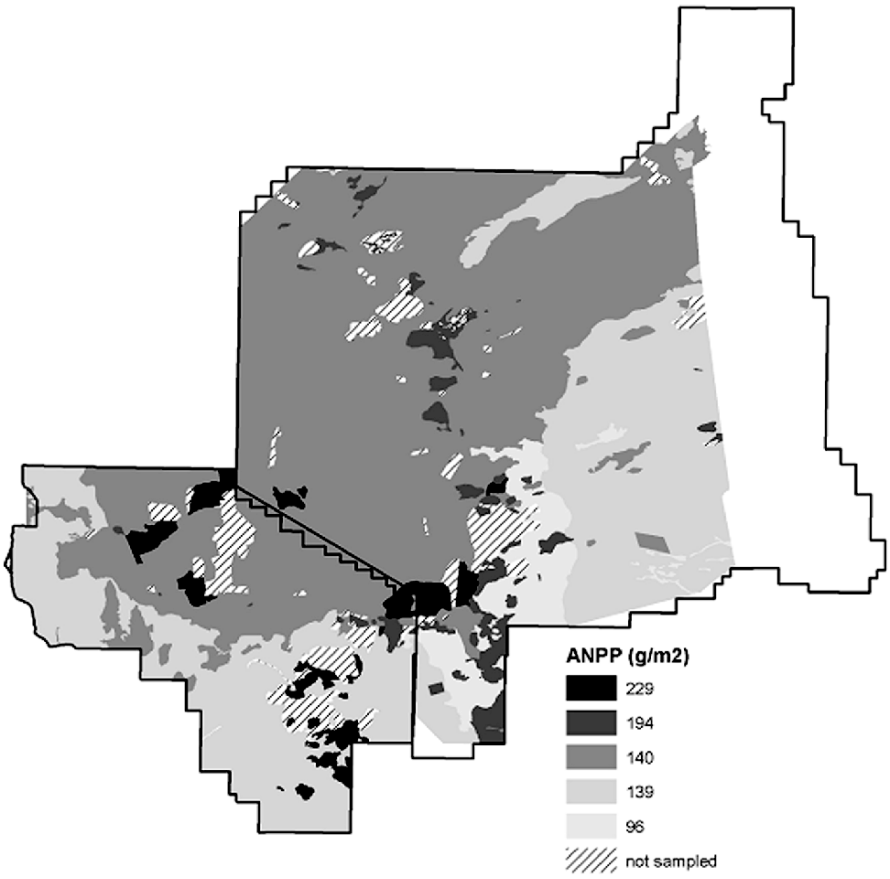


Figure 7.2. *Spatially implicit extrapolation of ANPP to the JRN landscape using the vegetation map and plot-based estimates. Both creosote bush and mesquite have the same average ANPP (139 g* m^{-2}*) and are shown in the same color.*

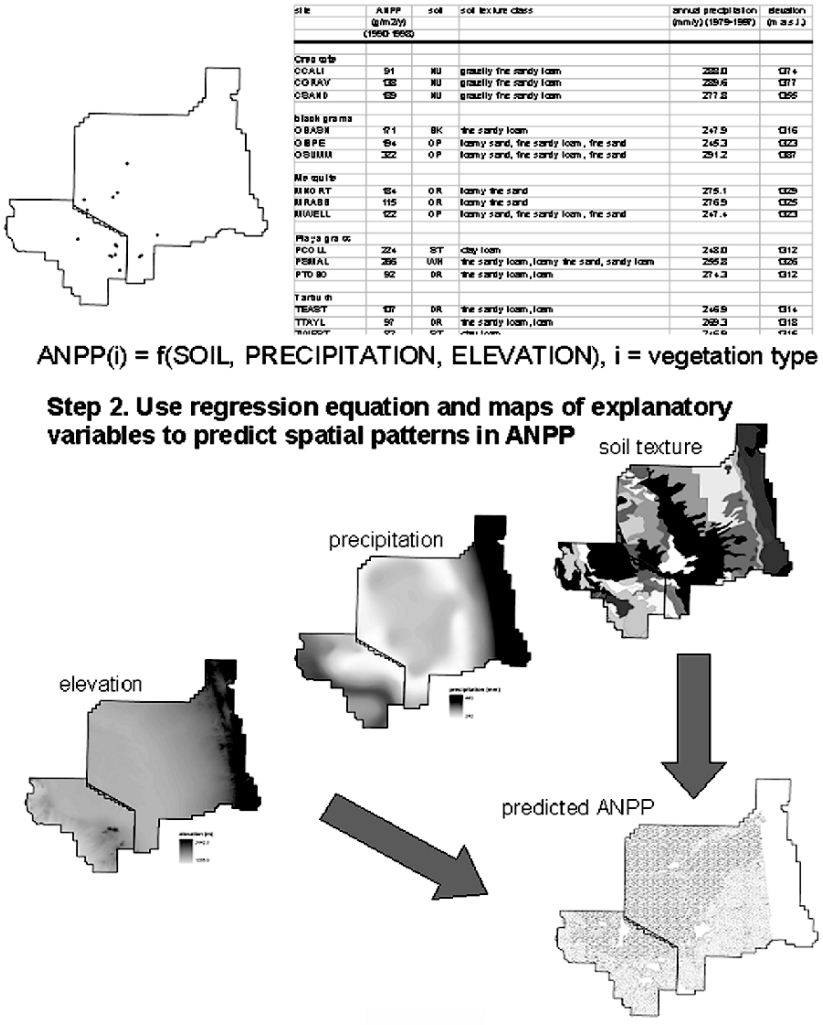


Figure 7.3. *Spatially implicit extrapolation of ANPP to the JRN landscape using a regression equation between plot-based measures of ANPP, soil texture, average precipitation, and elevation. The regression equation is then used to predict ANPP for the entire landscape using maps of the explanatory variables.*

7.3.3 Spatially Explicit

Spatially explicit approaches include landscape location as well as neighborhood or contagious processes, such as seed dispersal or wind and water redistribution of soil particles. These approaches require information on the movement or transfer of

materials, energy or information within and among spatial units on a landscape. Because these transfers are difficult and costly to measure for large areas, field experiments typically focus on specific areas of interest rather than attempting to instrument an entire landscape. For example, Schlesinger and Jones (1984) related patterns in plant biomass to localized runoff and run-on areas in the Mojave Desert. Recent experiments at the JRN have also documented the importance of water redistribution to patterns in vegetation (Wainwright et al. 2000, 2002).

Alternatively, spatially explicit simulation models can be used to represent large areas if sufficient information is known for model parameterization and validation. Peters and Herrick (1999) used a spatially explicit simulation model to examine the importance of seed dispersal to the recovery of perennial grasses following shrub invasion on sites with different vegetation and soil properties. Plot-scale parameters were combined with spatial maps of soil texture and precipitation as well as the movement of seeds among plots to extrapolate model results to a landscape. Spatially explicit simulation models can also be combined with regression models to simulate ecosystem dynamics.

In general, spatially explicit models are becoming increasing popular in ecology as computer limitations decrease and the quality and quantity of spatial information increases (e.g., Dunning et al. 1995, Schimel et al. 1997, He and Mladenoff 1999). However, complex models that require a large number of parameters that are difficult to estimate, and thus, have greater errors associated with them and are more difficult to validate than simple models (Oreskes et al. 1994, Rykiel 1996).

7.4 WHAT MAKES A LANDSCAPE COMPLEX?

Landscapes are complex because of interactions among contagious or neighboring processes and spatial variation in the physical template and disturbance regime. Contagious processes are related to three main vectors of dispersal (water, wind, animals) that redistribute seeds, nutrients, soil particles, and water. The physical template includes factors such as soil properties (texture, depth), precipitation, temperature, and elevation. The disturbance regime includes both natural (e.g., fire) and management-related disturbances (cultivation, roads, herbicide). Landscapes are complex because the importance of these processes, environmental factors, and disturbances varies for different sites. On some sites, spatial heterogeneity in soil texture can be the most important factor for explaining ecosystem dynamics whereas the redistribution of water may be more important on other sites. Furthermore, more than one contagious process or environmental factor may be important for some sites such that complex interactions among processes determine ecosystem dynamics.

In arid and semiarid ecosystems, patterns in ecosystem dynamics and ANPP are complex at the landscape scale because of a number of processes that are also variable at the landscape scale. Spatial variation in water redistribution at the landscape scale is most important along elevational gradients where channels and arroyos move water from upslope to downslope (Figure 7.4). Spatial variability in soil properties interacts with variability in water redistribution to generate complex patterns in ANPP. The result is that upper alluvial fans or bajadas dominated by

creosote bush with thin, rocky soils have higher runoff, less available water, and lower ANPP than downslope positions. The extreme situation occurs when water accumulates in playas on soils with high water holding capacity and high plant production (Huenneke et al. 2002). Because water is an important dispersal agent for seeds as well as soil particles, plant litter, and nutrients, these materials are also expected to be heterogeneously distributed across the landscape.

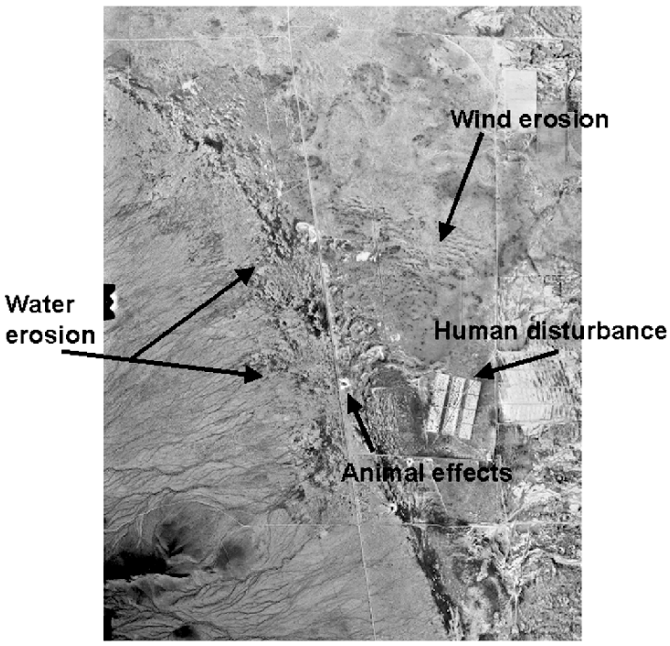


Figure 7.4. *Aerial photo showing complex landscapes at the JRN. Some areas near the mountains are dominated by water erosion as shown by the arroyos. Other areas on sandy soils are dominated by wind erosion as shown by dunes. Effects of animals on low plant production occur near water sources (wells). Human-caused disturbances are also non-uniformly distributed across the landscape. Areas with homogeneous vegetation and soils also occur.*

Wind redistribution of particles is also unevenly distributed at the landscape scale with important consequences for patterns in ANPP (Figure 7.4). Sandy soils dominated by honey mesquite are more susceptible to wind erosion than other soil-plant community combinations (Gibbens et al. 1983, Gillette and Chen 2001). Soil particles eroded from interdune areas in mesquite-dominated systems are deposited both locally in dunes and at large distances in vegetation types located downwind.

The effects of animals on ecosystem dynamics and ANPP are also non-uniformly distributed across a landscape (Figure 7.4). Both small and large animals are effective dispersal agents of seeds and nutrients. Densities of small animals are often related to soil properties and vegetation type (Kerley and Whitford 2000).

Furthermore, small animals, such as ants and rodents, are selective in the seeds collected with differential effects on seed availability and plant species dynamics (Inouye et al. 1980). Large animals, such as cattle, are effective dispersal agents of mesquite seeds, and likely played an important role in the expansion of this shrub into perennial grasslands. Spatial heterogeneity in ANPP also results from non-uniform grazing patterns by cattle (Paulsen and Ares 1962). Areas currently excluded from cattle often have higher plant production than adjacent grazed areas. Low plant production also occurs near watering holes where grazing and trampling are intense (Fusco et al. 1995, de Soyza et al. 1997, Nash et al. 1999).

Spatial heterogeneity in ANPP also occurs as a result of spatial variation in disturbances. Fire occurs most frequently in grassland systems with high production and continuous fuel load. Although low production occurs immediately following a fire, high production is possible later in the season if rainfall is high (Drewa et al. 2001). By contrast, low production can be maintained if fire occurs during or before a drought. Other types of disturbances, such as herbicide treatments, cultivation, road and building construction are also heterogeneously distributed across a landscape (Figure 7.4). These disturbances affect patterns in ANPP both locally on the disturbed area as well as in adjacent areas through modifications to wind and water erosion.

7.5 SIMPLIFYING COMPLEX LANDSCAPES: A NEW CONCEPTUAL FRAMEWORK

In our conceptual framework, landscapes consist of a mosaic of sites where spatial variation in the environment and contagious processes may or may not be important in understanding and predicting ecosystem dynamics. Our approach to simplifying complex landscapes is to determine the locations on a landscape where spatial information and contagious processes must be known in order for predictions to be accurate. Predictions for the remainder of the landscape can be obtained using estimates from representative sites that are extrapolated to similar areas using either nonspatial or spatially implicit methods (Peters et al. 2004). One approach to identifying these locations and to determining the important spatial processes is to combine remotely sensed images with field data and spatial databases residing in GIS.

For example, remotely sensed images combined with field estimates of ANPP can be used both to determine the most appropriate vegetation spectral index for each of the major ecosystem types, and to identify the locations (“hot spots”) with extremely low and high ANPP values for each ecosystem type compared with the rest of the landscape. Spatial databases can be used to provide insight into the key processes operating to generate these extremely low or high values. For example, digital elevation models can be used to determine the locations where water is expected to runoff as a result of steep slopes or where water accumulates when the slopes are shallower than the surrounding locations. Spatial databases of animal distribution can be generated using average stocking densities combined with the location of exclosures (no animal activity) and water sources (locations of intense animal activity). Correlating vegetation indices from the remotely sensed images

with the spatial databases can confirm the greater importance of spatial processes at these key locations compared with the rest of the landscape. Uncertainty analyses can then be conducted to identify the major sources of uncertainty, and to explore the effects of reducing the uncertainty in predictions by including spatial databases. Identifying these locations and key processes is the first step in simplifying complex landscapes in order to prioritize management decisions and to guide research questions and experimental designs.

7.6 SUMMARY AND CONCLUSIONS

Complex landscapes pose a critical challenge to ecologists. Addressing problems at the landscape scale requires the extrapolation of information from plot-scale studies. Three general classes of extrapolation methods exist that differ in the amount of spatial information required. For a given problem, ecologists typically use one method for all parts of a landscape. However, the use of one method likely results in high errors of omission for some parts of a landscape, and high errors of commission for other parts, thus resulting in high uncertainty in predictions for the entire landscape. An alternative approach was developed that simplifies complex landscapes into different parts where each extrapolation method is most appropriately used. This approach reduces the uncertainty in predictions at the landscape scale, and provides guidance to ecologists and land managers interested in the key parts of the landscape where spatial variation and contagious processes have the greatest impact on ecosystem dynamics. This approach is expected to also be useful for other ecosystems where complexity in landscape structure and spatial processes are important.

REFERENCES

Allred, K. W. 1996. Vegetative changes in New Mexico rangelands. Pages 168-231 *in* E. A. Herrera and L. F. Huenneke, editors. New Mexico's Natural Heritage: Biological Diversity in the Land of Enchantment. New Mexico Journal of Science 36. New Mexico Academy of Science, Albuquerque, NM.

Archer, S, T. W. Boutton, and K. A. Hibbard. 2001. Trees in grasslands: biogeochemical consequences of woody plant expansion. Pages 115-137 *in* E. D. Schulze, S. P. Harrison, M. Heimann, E. A. Holland, J. Lloyd, I. C. Prentice, and D. Schimel, editors. Global Biogeochemical Cycles in the Climate System. Academic Press, San Diego, CA.

Buffington, L. C., and C. H. Herbel. 1965. Vegetational change on a semidesert grassland range from 1858 to 1963. Ecological Monographs 35:139-164.

Burke, I. C., T. G. F. Kittel, W. K. Lauenroth, P. Snook, C. M. Yonker, and W. J. Parton. 1991. Regional analysis of the Central Great Plains. BioScience 41:685-692.

Burke, I. C., W. K. Lauenroth, and W. J. Parton. 1997. Regional and temporal variation in net primary production and nitrogen mineralization in grasslands. Ecology 78:1330-1340.

Clark, J. S., E. Macklin, and L. Wood. 1998. Stages and spatial scales of recruitment limitation in southern Appalachian forests. Ecological Monographs 68:213-235.

Coffin, D. P., and W. K. Lauenroth. 1989. Disturbances and gap dynamics in a semiarid grassland: a landscape-level approach. Landscape Ecology 3:19-27.

de Soyza, A. G., W. G. Whitford, and J. E. Herrick. 1997. Sensitivity testing of indicators of ecosystem health. Ecosystem Health 3:44-53.

Drewa, P. B., D. P. C. Peters, and K.M. Havstad. 2001. Fire, grazing, and shrub invasion in the Chihuahuan Desert. Pages 31-39 *in* T. P. Wilson and K. E. M. Galley, editors. Proceedings of the

Invasive Species Workshop: The Role of Fire in the Control and Spread of Invasive Species. Fire Conference 2000: the National Congress on Fire Ecology, Prevention, and Management. Misc. Pub. No. 11. Tall Timbers Research Station, Tallahassee, FL.

Dunning, J. B., D. J. Stewart, B. J. Danielson, B. R. Noon, T. L. Root, R. H. Lamberson, and E. E. Stevens. 1995. Spatially explicit population models: current forms and future uses. Ecological Applications 5:3-11.

Fredrickson, E., K. M. Havstad, R. Estell, P. Hyder. 1998. Perspectives on desertification: south-western United States. Journal Arid Environments 39:191-207.

Fusco, M., J. Holechek, A. Tembo, A. Daniel, and M. Cardenas. 1995. Grazing influences on watering point vegetation in the Chihuahuan Desert. Journal Range Management 48:32-38.

Gardner, R. H., R. V. O'Neill, J. B. Mankin, and D. Kumar. 1980. Comparative error analysis of six predator-prey models. Ecology 61:323-332.

Gibbens, R. P., and R. F. Beck. 1988. Changes in grass basal area and forb densities over a 64-year period on grassland types of the Jornada Experimental Range. Journal Range Management 41:186-192.

Gibbens, R. P., J. M. Tromble, J. T. Hennessy, and M. Cardenas. 1983. Soil movement in mesquite dunelands and former grasslands of southern New Mexico from 1933 to 1980. Journal Range Management 36:145-148.

Gile, L. H., J. W. Hawley, and R. B. Grossman. 1981. Soils and geomorphology in the basin and range area of southern New Mexico: guidebook to the Desert Project. New Mexico Bureau of Mines and Mineral Resources, Socorro, NM.

Gillette, D. A., and W. Chen. 2001. Particle production and aeolian transport from a "supply-limited" source area in the Chihuahuan desert, New Mexico, United States. Journal of Geophysical Research 106:5267-5278.

Hartley, A. E., and W. H. Schlesinger. 2000. Environmental controls on nitric oxide emission from northern Chihuahuan desert soils. Biogeochemistry 50:279-300.

He, H. S., and D. J. Mladenoff. 1999. Spatially explicit and stochastic simulation of forest-landscape fire disturbance and succession. Ecology 80:81-99.

Hibbard, K. A., S. Archer, D. S. Schimel, and D. W. Valentine. 2001. Biogeochemical changes accompanying woody plant encroachment in a subtropical savanna. Ecology 82:1999-2011.

Hogeweg, P. 1988. Cellular automata as a paradigm for ecological modeling. Applications in Mathematics and Computation 27:81-100.

Huenneke, L. F., J. P. Anderson, M. Remmenga, and W. H. Schlesinger. 2002. Desertification alters patterns of aboveground net primary production in Chihuahuan ecosystems. Global Change Biology 8:247-264.

Huenneke, L. F., D. Clason, and E. Muldavin. 2001. Spatial heterogeneity in Chihuahuan Desert vegetation: implications for sampling methods in semi-arid ecosystems. Journal of Arid Environments 47:257-70.

Humphrey, R. R. 1958. The desert grassland: a history of vegetational change and an analysis of causes. Botanical Review 24:193-252.

Hurtt, G. C., S. W. Pacala, P. R. Moorcroft, J. Casperson, E. Shevliakova, R. A. Houghton, and B. Moore. 2002. Projecting the future of the U.S. carbon sink. Proceedings National Academy of Sciences 99:1389-1394.

Inouye, R. S., G. S. Byers, and J. H. Brown. 1980. Effects of predation and competition on survivorship, fecundity, and community structure of desert annuals. Ecology 6:1344-1351.

Jackson, R. B., J. L. Banner, E. G. Jobbagy, W. T. Pockman, and D. H. Wall. 2002. Ecosystem carbon loss with woody plant invasion of grasslands. Nature 418:623-626.

Jarvis, P. G. 1995. Scaling processes and problems. Plant, Cell and Environment 18:1079-1089.

Keane, R. E., M. Austin, R. Dahlman, C. Field, A. Huth, M. Lexer, D. Peters, A. Solomon, and P. Wyckoff. 2001. Plant mortality in gap models. Journal of Climate Change 51:509-540.

Kerley, G. H., and W. Whitford. 2000. Impact of grazing and desertification in the Chihuahuan Desert: plant communities, granivores, and granivory. American Midland Naturalist 144:78-91.

King, A. W. 1991. Translating models across scales in the landscape. Pages 479-517 *in* M. G. Turner and R. H. Gardner, editors. Quantitative Methods in Landscape Ecology. Springer-Verlag, New York.

Knapp, A. K., and M. D. Smith. 2001. Variation among biomes in temporal dynamics of above-ground primary production. Science 291:481-484.

Lauenroth, W. K. 1979. Grassland primary production: North American grasslands in perspective. Pages 3-24 *in* N. R. French, ed. Perspectives in grassland ecology. Springer-Verlag, New York.

Le Houérou, H. N., R. L. Bingham, and W. Skerbek. 1988. Relationship between the variability of primary production and the variability of annual precipitation in world arid lands. Journal Arid Environments 15:1-18.

Lieth, H., and R. H. Whittaker, editors. 1975. The Primary Productivity of the Biosphere. Springer, New York.

Ludwig, J. A. 1986. Primary production variability in desert ecosystems. Pages 5-17 *in* W. G. Whitford, editor. Patterns and Process in Desert Ecosystems. University of New Mexico Press, Albuquerque, NM.

Ludwig, J. A. 1987. Primary productivity in arid lands: myth and realities. Journal Arid Environments 13:1-7.

Mack, G. H., W. C. McIntosh, M. R. Leeder, and H. C. Monger. 1996. Plio-Pleistocene pumice floods in the ancestral Rio Grande, southern Rio Grande rift, USA. Sedimentary Geology 103:1-8.

Melillo, J. M., J. Borchers, J. Chaney, H. Fisher, S. Fox, A. Haxeltine, A. Janetos, D. W. Kicklighter, T. G. F. Kittel, A. D. McGuire, R. McKeown, R. Neilson, R. Nemani, D. S. Ojima, T. Painter, Y. Pan, W. J. Parton, L. Pierce, L. Pitelka, C. Prentice, B. Rizzo, N. A. Rosenbloom, S. Running, D. S. Schimel, S. Sitch, T. Smith, and I. Woodward. 1995. Vegetation Ecosystem Modeling and Analysis Project – Comparing biogeography and biogeochemistry models in a continental-scale study of terrestrial ecosystem responses to climate-change and CO_2 doubling. Global Biogeochemical Cycles 9:407-437.

Miller, C., and D. L. Urban. 1999. Interactions between forest heterogeneity and surface fire regimes in the southern Sierra Nevada. Canadian Journal of Forest Research 29:202-212.

Mitchell, K. A., D. P. C. Peters, J. E. Herrick, and H. C. Monger. 2002. Long term changes in carbon pools accompanying shrub invasion of a desert grassland: the Jornada Experimental Range 1850 - today. Bulletin Ecological Society America 83:215.

Monger, H.C. 2006. Soil development in the Jornada Basin. Pages 50-66 *in* K. M. Havstad, L. F. Huenneke, and W. H. Schlesinger, editors. Structure and Function of a Chihuahuan Desert Ecosystem: The ornada Basin Long Term Ecological Research Site. Oxford University Press, New York, NY.

Nash, M. S., W. G. Whitford, A. G. De Soyza, J. W. Van Zee, and K. M. Havstad. 1999. Livestock activity and Chihuahuan Desert annual-plant communities: boundary analysis of disturbance gradients. Ecological Applications 9:814-823.

Neilson, R. P., and S. W. Running. 1996. Global dynamic vegetation modelling: coupling biogeochemistry and biogeography models. Pages 451-465 *in* B. H. Walker and W. L. Steffen, editors. Global Change and Terrestrial Ecosystems. Cambridge University Press, Cambridge, UK.

O'Neill, R. V. 1979. Natural variability as a source of error in model predictions. Pages 23-32 *in* G. S. Innis and R. V. O'Neill, editors. Systems Analysis of Ecosystems. International Co-operative Publishing House, Fairland, MD.

Oreskes, N., K. Shrader-Frechette, and K. Belitz. 1994. Verification, validation, and confirmation of numerical models in the earth sciences. Science 263:641-646.

Pacala, S. W., G. C. Hurtt, D. Baker, P. Peylin, R. A. Houghton, R. A. Birdsey, L. Heath, E. T. Sundquist, R. F. Stallard, P. Ciais, P. Moorcroft, J. P. Caspersen, E. Shevliakova, B. Moore, G. Kohlmaier, E. Holland, M. Gloor, M. E. Harmon, S.-M. Fan, J. L. Sarmiento, C. L. Goodale, D. Schimel, and C. B. Field. 2001. Consistent land- and atmosphere-based U.S. carbon sink estimates. Science 292:2316-2320.

Paulsen, H. A., and F. N. Ares 1962. Grazing values and management of black grama and tobosa grasslands and associated shrub ranges of the Southwest. USDA Forest Service Bulletin 1270.

Peters, D.P.C. 2002. Plant species dominance at a grassland-shrubland ecotone: an individual-based gap dynamics model of herbaceous and woody species. Ecological Modelling 152:5-32.

Peters, D. P. C., and J. E. Herrick. 1999. Vegetation-soil feedbacks and sensitivity of Chihuahuan desert ecosystems to climate change. Bulletin of the Ecological Society of America 80:69.

Peters, D. P. C., J. E. Herrick, D. L. Urban, R. H. Gardner, and D. D. Breshears. 2004. Strategies for ecological extrapolation. Oikos 106:627-636.

Rastetter, E. B., J. D. Aber, D. P. C. Peters, D. S. Ojima, and I. C. Burke. 2003. Using mechanistic models to scale ecological processes across space and time. BioScience 53:1-9.

Reynolds, J. F., and B. Acock. 1985. Predicting the response of plants to increasing carbon dioxide: a critique of plant growth models. Ecological Modelling 29:107-129.

Rykiel, E. J. 1996. Testing ecological models: the meaning of validation. Ecological Modelling 90:229-244.

Schimel, D. S., W. R. Emanuel, B. Rizzo, T. M. Smith, F. I. Woodward, H. Fisher, G. F. Kittel, R. McKeown, T. Painter, N. Rosenbloom, D. S. Ojima, W. J. Parton, D. W. Kicklighter, A. D. McGuire, J. M. Melillo, Y. Pan, A. Haxeltine, C. Prentice, S. Sitch, K. Hibbard, R. Nemani, L. Pierce, S. W. Running, J. Borchers, J. Chaney, R. Neilson, and B. B.H. 1997. Continental scale variability in ecosystem processes: models, data, and the role of disturbance. Ecological Monographs 67:251-271.

Schlesinger, W. H., and C. S. Jones. 1984. The comparative importance of overland runoff and mean annual rainfall to shrub communities of the Mojave Desert. Botanical Gazette 145:116-124.

Schlesinger, W. H., and A. M. Pilmanis. 1998. Plant-soil interactions in deserts. Biogeochemistry 42:169-187.

Schlesinger, W. H., J. F. Reynolds, G. L. Cunningham, L. Huenneke, W. M. Jarrell, R. A. Virginia, and W. G. Whitford. 1990. Biological feedbacks in global desertification. Science 247:1043-1048.

Stein, R. A., and J. A. Ludwig. 1979. Vegetation and soil patterns on a Chihuahuan Desert bajada. The American Midland Naturalist 101:28-37.

Symstad, A. J., F. S. Chapin, D. H. Wall, K. L. Gross, L. F. Huenneke, G. G. Mittelbach, D. P. C. Peters, and D. Tilman. 2003. Long-term and large-scale perspectives on the relationship between biodiversity and ecosystem functioning. Bioscience 53:89-98.

Turner, M. G., V. H. Dale, and R. H. Gardner. 1989a. Predicting across scales: theory development and testing. Landscape Ecology 3:245-252.

Turner, M. G., R. H. Gardner, V. H. Dale, and R. V. O'Neill. 1989b. Predicting the spread of disturbance across heterogeneous landscapes. Oikos 55:121-129.

Van Auken, O. W. 2000. Shrub invasions of North American semiarid grasslands. Annual Review Ecology Systematics 31:197-215.

Wainwright, J., A. J. Parson, and A. D. Abrahams. 2000. Plot-scale studies of vegetation, overland flow and erosion interactions: case studies from Arizona and New Mexico. Hydrological Processes 14:2921-2943.

Wainwright, J., A. J. Parson, W. H. Schlesinger, and A. D. Abrahams. 2002. Hydrology-vegetation interactions in areas with discontinuous flow on a semi-arid bajada, southern New Mexico. Journal Arid Environments 51:219-258.

Webb, W. L., S. R. Szarek, W. K. Lauenroth, R. S. Kinerson, and M. H. Smith. 1978. Primary productivity and water-use in native forest, grassland and desert ecosystems. Ecology 59:1239-1247.

Wieranga, P. J., J. M. H. Hendrick, M. H. Nash, J. A. Ludwig, and L. A. Daugherty. 1987. Variation of soil and vegetation with distance along a transect in the Chihuahuan Desert. Journal Arid Environments 13:53-63.

Wondzell, S. M., and J. A. Ludwig. 1995. Community dynamics of desert grasslands: influences of climate, landforms, and soils. Journal of Vegetation Science 6:377-390.

Wondzell, S. M., J. M. Cornelius, and G. L. Cunningham. 1990. Vegetation patterns, microtopography, and soils on a Chihuahuan Desert playa. Journal Vegetation Science 1:403-410.

Wondzell, S. M., G. L. Cunningham, and D. Bachelet. 1996. Relationships between landforms, geomorphic processes, and plant communities on a watershed in the northern Chihuahuan Desert. Landscape Ecology 11:351-362.

Wu, J. 1999. Hierarchy and scaling: extrapolating information along a scaling ladder. Canadian Journal Remote Sensing 25:367-380.

Yao, J., D. P. C. Peters, R. P. Gibbens, and K. M. Havstad. 2002. Response of perennial grasses to precipitation in the northern Chihuahuan Desert: implications for grassland restoration. Bulletin Ecological Society America 83:302.

CHAPTER 8

BUILDING UP WITH A TOP-DOWN APPROACH:

The Role of Remote Sensing in Deciphering Functional and Structural Diversity

CAROL A. WESSMAN AND C. ANN BATESON

8.1 INTRODUCTION

> *". . . . particularity and contingency, which characterize the ecological sciences, and generality and simplicity, which characterize the physical sciences, are miscible, and indeed necessary, ingredients in the quest to understand humankind's home in the universe."*
>
> *~ John Harte, 2002*

Historically, ecologists have dealt with the complexity of ecological systems through small, place-centered studies. This approach, in controlling for environmental variation, has led to important advances in our understanding of mechanisms behind ecological phenomena, and has moved us more directly from pattern to causal processes (Harte 2002). However, the need to develop broader "universal laws" in ecology is important in the context of global-wide environmental change. While ecological "laws" may lack the exactness and universality of physical laws, they will improve our power to predict the consequences of change due to human activities and climate variation, "signposting routes to a sustainable future" (Lawton 2001).

Ecological scaling, and hence any efforts to define universalities, is challenged by inherent nonlinear synergies and heterogeneity, cross-scale processes, thresholds, and emergent properties of ecosystems (e.g., Wu 1999, Peterson 2000, Wu and Li, Chapter 2). Techniques are required that are able to translate place-centered, mechanistic understanding (the "peculiarities and contingencies" *sensu* Harte 2002) across a range of spatial and temporal scales. Remote sensing of the Earth's surface, while limited in its ability to fully address all the challenges, helps constrain the scaling problem through its synoptic view of biophysical and biochemical structure across different scales (Wessman 1992, Wessman and Asner 1998). The structure of landscapes and regions (i.e., the properties of cover types and their distribution) are

J. Wu, K.B. Jones, H. Li, and O.L. Loucks (eds.),
Scaling and Uncertainty Analysis in Ecology: Methods and Applications, 147–163.

of great importance and interest to the ecological scaling effort for two reasons. First, structure superimposes constraints on the functioning of ecological systems at broad to finer scales. Second, the structure itself is an expression of the functional properties that emerge from interactions among biological, physical and geochemical processes. Scaling in remote sensing, by its attention to surface heterogeneity and the derivation of surface parameter algorithms (Chen 1999), takes an important role in earth system science, predicated on the assumption that we recognize the appropriate features to be scaled (Wessman 1992, Wessman and Asner 1998). In many respects, remote sensing mandates a generality in our observations of the earth's surface that, welcome or not, forces a new perspective of ecological properties. Through interchange between observational-based science and ecological scaling theory, generalities in ecological dynamics will surface.

Scaling in remote sensing is controlled in two fundamental ways: (1) heterogeneity in composition and configuration of the landscape under observation relative to sensor characteristics, and (2) nonlinear functional relationships between surface radiation and ecologically relevant parameters. The objective of this paper is to explore these aspects of the remote sensing scaling problem, and the power and uncertainties they interject into the upscaling of field measurements. First, we introduce the ecological variables and relationships that the community has or is attempting to measure with remote sensing. We then review briefly the radiation and remote sensing properties that are important to upscaling. Finally, we synthesize some of the more important sources of uncertainties and error in algorithms used to scale surface parameters.

8.2 ECOLOGICAL VARIABLES RETRIEVED BY REMOTE SENSING

Extraction of variables from imagery falls into two broad categories. First, classification techniques are used to categorize the landscape into discrete recognizable units with relevance to various mapping aims. Second, continuous variables are retrieved on the basis of spectral-biophysical relationships in each pixel. In the former, within-class variance is ultimately ignored as pixels are assigned discrete values representing a given category. In the latter, pixel values will lie within the range of variability of the retrieved variable contained by the observation. Continuous variables such as leaf area index (LAI), fractional cover and fraction of absorbed photosynthetically active radiation (fAPAR) represent basic components or singular functions of the landscape and will largely scale as a function of the radiative transfer properties of landscape components, their relative dominance in the grid cell, and the linearity of the retrieval algorithm. Scaling of discrete, categorical variables will be influenced by the size and nesting of ground components detectable at diverse resolutions.

The ecological variables that take on significance as scales are increased and that can be estimated remotely are integrative in nature or represent important constraints on processes (Table 8.1). These include structural variables such as LAI, biomass, land cover, and fractional cover of landscape components (e.g., green vegetation, bare soil) that quantify, to greater or lesser degrees, the spatial heterogeneity important to extrapolation or modeling of related processes. fAPAR is one of the

few parameters that provide a direct connection between ecosystem structure and function (Asner and Wessman 1997), and it also provides a means to link to other functional attributes of ecosystems such as nitrogen use, CO_2 assimilation, and water loss (Sellers 1987, Running et al. 1994, Field et al. 1995). Remote sensing of foliar chemistry is of interest due to its role in ecosystem productivity. Estimates of spatial variation in canopy chemistry from hyperspectral imagery, while empirically derived through local to regional calibration, are valuable for landscape-level links to ecosystem processes such as productivity and decomposition (Wessman et al. 1988, Zagolski et al. 1996, Martin and Aber 1997, Smith et al. 2002).

Table 8.1. *Ecological variables commonly derived from remotely sensed imagery (after Milne and Cohen 1999).*

Variable	*Type*	*Units*	*Applications*
Land Cover	Categorical	ha, %	Ecosystem model stratification, land use change, habitat characterization, ecosystem management
LAI	Continuous	$m^2\ m^{-2}$	CO_2 and trace gas exchange models or measurement extrapolation, carbon allocation
NPP, NEP	Continuous	$g\ m^{-2}\ yr^{-1}$	Estimation of ecosystem carbon gain
Fractional cover	Continuous	%	Ecosystem model stratification, land use/land cover change, succession, biophysical land surface modeling
Canopy chemistry	Continuous	%, $g\ m^{-2}$	Ecosystem productivity models, decomposition
Canopy geometry	Continuous	m	Land-atmosphere energy flux, climate models

Each of these variables gains significance with broadened extent in either of two ways. The importance of their magnitude and distribution, alone or in combination with other remotely sensed variables, may indicate configuration and/or connectivity of landscape components undetectable at the field level. Also, certain remotely sensed variables are incorporated into process models to drive or constrain simulations of biogeochemical process, land-atmosphere energy and trace gas flux, and large-scale climate models.

8.3 THE RETRIEVAL OF ECOLOGICAL VARIABLES

Remote sensing is a valuable means of upscaling ecological variables due to multiple scales of observations and stability of satellite platforms in space and time. Repetitive acquisitions at temporal scales of interest (e.g., diurnal, seasonal,

interannual, duration of transient weather systems) enable analyses of change to answer questions about trends and cycles. But, fundamental to the retrieval and interpretation of ecological variables from remote sensing platforms are the reflectance characteristics of the observed surface and its components. Aside from some types of mapping and sophisticated radiative transfer methods, remote sensing does not replace any widely used ground measurement in ecology (Prince 1999). Most parameters derived from remote sensing data require a certain level of calibration with ground conditions (van Leeuwen et al. 1997, Qi et al. 2000). Each technique for the inference of an ecological variable must be understood in terms of the factors contributing to the measurement of reflectance integrated over the pixel.

The reflectance signal measured by the sensor is the integrated outcome of a complex interaction of surface scattering properties, including single and multiple scattering of photons, and solar and sensor viewing geometry. For example, vegetation reflectance is primarily a function of tissue (wood, green, senesced) optical properties, canopy structure and geometry (leaf and stem area and orientation, foliage clumping), soil reflectance, illumination conditions, and viewing geometry (Ross 1981, Myneni et al. 1989). The structural attributes of ecosystems (e.g., species composition, vertical structure, canopy closure) determine the relative contributions of tissue, canopy and landscape factors that drive the variation in a reflectance signal (Asner 1998). Sorting out these factors requires an understanding of the sources of variation at each scale (which is ecosystem dependent) as well as an adequate sampling (spectral, angular, and temporal) of the spectrum.

Spectral algorithms like vegetation indices (VI) that are relatively simple and are composed of few spectral bands are widely used to monitor vegetation dynamics and to infer biophysical properties such as leaf area index (LAI) and fAPAR. The main shortcoming of vegetation indices is the lack of functional relationships with biophysical parameters (van Leeuwen et al. 1997, Qi et al. 2000). Empirical or modeled relationships with variables such as LAI, fAPAR, and biomass can be developed on a site-specific basis, but these relationships are subject to changes in vegetation properties, soil background, atmosphere and the sun-surface geometry. Validation efforts for the MODIS sensor are aimed at testing the accuracy of VI products via multi-scaled analyses in order to gain an understanding of the causes of errors for potential improvement in future algorithms (Tian et al. 2002a).

Radiative transfer models provide a means to assess the canopy radiation regime from a physical and mechanistic basis, and model inversions derive variables that are more directly related to biophysical properties of vegetation. Current models are based on the physics of photon scattering, and range in complexity from one-dimensional (vertical profile) algorithms to complex three-dimensional landscape simulations (Liang and Strahler 1993, Myneni and Asrar 1993, Kuusk 1995, Asner and Wessman 1997). The models that include scaled scattering characteristics of leaves, canopies, and soil can be used to explore the radiation regime in landscapes. For example, understanding the relative influence of structure at different scales (tissue, canopies, landscapes) on the fAPAR of an ecosystem helps determine what variables need to be accounted for and which can be ignored or held at a constant value in future studies (Asner et al. 1998b, Asner et al. 1998c). The anisotropic scatter of radiation by vegetation canopies has been exploited through measurement

and modeling of the bidirectional distribution function (BRDF) to retrieve LAI and canopy geometry (Li and Strahler 1992, Privette et al. 1994, Braswell et al. 1996, Asner et al. 1997, Asner 2000). This continues to be a strong area of research with new operational BRDF instruments (e.g., MISR, POLDER) (Jin et al. 2002, Chopping et al. 2003). Less quantitative, but analogous to these modeling approaches, spectral mixture analysis decomposes the reflectance signal into the fundamental contributing components of the landscape (e.g., soil, green foliage, senesced vegetation) and simplifies the interpretable connection to biophysical parameters (Wessman et al. 1997, Sabol et al. 2002).

In this paper, we focus on the upscaling of biophysical variables through remote sensing. However remote sensing methods in land cover classification and feature mapping are very important aspects of remote sensing scaling and need to be noted here. Classification is a well-established approach to map land-cover types that represent composites such as vegetation and land use, habitat, or ecosystem types. Land cover maps are used widely as one means to quantify landscape heterogeneity and parameterize the biophysical properties of plant canopies in models of climate and biogeochemical processes. Efforts continue to improve accuracies and quantify errors associated with classification algorithms (e.g., Hlavka and Dungan 2002), particularly in the interest of scaling to continental and global scales (e.g., Friedl et al. 2002, Lotsch et al. 2003). The ability to describe spatial patterns and the underlying processes that generate them is largely determined by the relationships between the objects in the scene and the scales at which we observe them. A significant literature exists on scaling issues specific to feature extraction, spatial structures, and the spatial variation in remote sensing imagery (Marceau and Hay 1999). Multiscale approaches to upscaling and feature extraction are being developed to contend with the multi-scaled and spatially distributed objects in a landscape (Hay et al. 1997, 2001).

8.4 QUANTIFYING BIAS AND ERROR

There are several sources of error and bias in analyses based on remote sensing that are caused by the indirect relationship between reflected radiation and surface characteristics. The most obvious source of error is the intervening atmosphere between the sensor and the reflecting surface. Although methods have been developed to remove atmospheric effects from imagery (Gao et al. 1993, Qu et al. 2000), these methods do not duplicate ground based spectra and can themselves delete spectral features critical for detecting the presence of ground components in the scene (Kruse and Dwyer 1993).

Another source of error that has been studied extensively in the literature and is algorithmic in nature involves extrapolating relationships between remotely sensed parameters and surface attributes from the scale for which they were developed to a coarser resolution. For example, a functional relationship between LAI and the NDVI, whose defining constants were found by relating the NDVI values at a 30 m pixel size to ground values adequately sampled on 30 m plots, may not be the correct relationship for predicting LAI from AVHRR or MODIS at 1 km pixels. In order for a quantitative algorithm to be scale invariant with respect to a particular

landscape, either the algorithm must implement a linear function between the parameters or ecological parameter values within the larger pixel must be constant (Hu and Islam 1997). Hence, for quantitative algorithms, heterogeneity and non-linearity are the two factors determining the magnitude of scaling errors and bias. Since most radiance-to-ground relationships are developed on a small scale, methods are needed to measure the potential aggregation error and ideally add a correction factor. We will review several approaches for estimating large-scale quantities using smaller scale, field-acquired measurements. Many of these methods require pixel-level knowledge of fractional coverage by ground components. We investigate spectral mixture analysis (SMA), which inverts a mixture model to retrieve cover fractions, as a tool for measuring heterogeneity that can be incorporated into scaling methods.

Most but not all algorithms bridging remote sensing and surface attributes require adequate ground sampling for calibration and validation. Misregistration of image pixels with their ground locations is common and has an impact relative to the scene heterogeneity. Spatial statistics has provided tools such as the variogram, local variance and kriging to help in designing efficient sampling schemes by detecting spatial correlations and consequently appropriate lags between samples to achieve a set of statistically independent values. We will review some of these methods and issues related to error in their application.

8.4.1 Aggregation Error

We first present a theoretical framework for understanding the issues involved in applying a functional relationship from the scale for which it is correct to an aggregate. Let R be a function between the bands of a remote sensing instrument with spatial resolution L and a surface parameter P. That is,

$$R(b_1^L(i), ..., b_m^L(i)) = P^L(i) \tag{8.1}$$

where $b_j^L(i)$ is reflectance of the i^{th} pixel in band j and $P^L(i)$ is the ground parameter value in pixel i. Suppose we aggregate the pixels of the instrument into superpixels of size nL. We can compute two quantities:

$$\frac{1}{n^2}\sum_{i=1}^{n^2} R(b_1^L(i), ..., b_m^L(i)) \tag{8.2}$$

and

$$R(\frac{1}{n^2}\sum_{i=1}^{n^2} b_1^L(i), ... \frac{1}{n^2}\sum_{i=1}^{n^2} b_m^L(i)) \tag{8.3}$$

An algorithm implementing Equation 8.2 will be described as parameter-aggregated since it is the mean of the ground parameter $P^L(i)$ over the subpixels that comprise the superpixel. Algorithms implementing Equation 8.3, on the other

hand, will be described as band-aggregated since the bands are aggregated into a superpixel before the function R is applied. R is scale invariant if and only if Equation 8.2 and Equation 8.3 are equal and these two quantities are identically equal if and only if R is a linear function of its variables. However, for a particular scene, if objects such as forests or grasslands have a typical size at least as large as the size of the superpixel, then all subpixels will be similar and hence their bands values and means will be approximately equal and scale invariance will persist for any function. Equality of band values over the subpixels, of course, will not happen in a real scene, but band variation over an object often has a negligible effect on the equality between Equation 8.2 and Equation 8.3 above (e.g., within forest variance in LAI produces small scaling errors). Since R was developed for resolution L, Equation 8.2 is the correct value for the ground parameter computed for the superpixel. However, for large-scale imagery such as AVHRR, we usually do not have subpixel information and ground sampling to produce relationships is not feasible. We look at error associated with using Equation 8.3 to estimate LAI from the normalized difference vegetation index (NDVI). But, first, we investigate the nonlinearity of the NDVI itself.

8.4.1.1 NDVI

The NDVI exploits differences in vegetation reflectance response between the red and NIR to detect ground vegetation properties and is computed by

$$NDVI = (NIR - red)/(NIR + red) \tag{8.4}$$

However, since the NDVI is not a linear function of its variables (NIR and red reflectance), the NDVI of a superpixel need not equal the mean NDVI of its subpixels (i.e., the parameter-aggregated NDVI doesn't equal the band-aggregated NDVI). This discrepancy will introduce error in scaling up a functional relationship R between the NDVI and a ground parameter even when R is linear provided that the lower resolution pixels are not homogeneous. Hu and Islam (1997) investigated the effects of landscape heterogeneity on scaling errors with respect to the NDVI and reported the following:

a) The relative difference between the parameter-aggregated and band-aggregated NDVI (i.e., relative scaling error) is a function of within-superpixel variance of red and NIR reflectances and within superpixel covariance between the two bands. If the distribution of red and NIR reflectances among the subpixels is too highly heterogeneous, then the error will be too great to approximate the parameter-aggregated NDVI with the band-aggregated NDVI algorithm.
b) There was significant scaling error in a hypothetical example when the band-aggregated algorithm was used to estimate NDVI.
c) A correction term CT to account for within-superpixel heterogeneity can be computed from information found on the superpixel level and in the hypothetical example band-aggregated NDVI + CT gave a good

approximation to the parameter-aggregated NDVI. The term *CT* is computed through functional relationships parameterizing within-superpixel variance and covariance between NIR and red bands with reflectances acquired by the sensor.

The NDVI error measurement and correction algorithm in Hu and Islam (1997) has not been tested on remotely sensed data from two different instruments at two different scales. Since its assumptions are simplistic (e.g., all vegetation has the same reflectance and likewise for background) and only first and second order correction terms were considered, the result remains preliminary until such testing is performed. Aman et al. (1992) degraded SPOT and TM data collected over tropical sites in West Africa and agricultural fields in France from 20 and 30 m respectively, to resolutions ranging from 200-1000 m, which are more suitable for global vegetation studies. They found a significant linear correlation between parameter-aggregated and band-aggregated NDVI for resolutions below 1000 m with slope and intercept close to 1 and 0 respectively. Hence, they concluded, on the basis of their samples, that the parameter-aggregated NDVI can be estimated from the band-aggregated NDVI with acceptable errors. That is, the errors are less than the uncertainties relating the high resolution NDVI and ground parameters and the errors resulting from radiometric corrections. However, similar experiments with other landscape types need to be performed to determine the domain of acceptability for the band-aggregated algorithm.

8.4.1.2 NDVI and LAI

The infeasibility of collecting LAI values in the field on a scale required for global and regional vegetation studies mandates efforts to compute LAI through functional relationships with vegetation indices derived from remotely sensed imagery or through inversions of radiative transfer models. Chen and Cihlar (1996) found the nonlinear relationship

$$NDVI = 0.5520*LAI^{0.1844} \tag{8.5}$$

between the NDVI computed from TM imagery and LAI values collected in boreal conifer forests in the Boreal Ecosystem-Atmosphere Study (BOREAS) site with a plant canopy analyzer (*LAI-2000*, *Licor*).

The scaling error incurred in upscaling this algorithm from the TM 30 m pixel resolution to 1 km AVHRR pixels was investigated by first degrading the TM imagery to the AVHRR resolution (Chen 1999). Study areas of 990 m × 990 m with mixtures of vegetation and water were selected from the imagery to give a range of water coverage from 0% to 93%. Coexistence of these two highly contrasting surfaces in the same low resolution pixel was expected to produce large scaling errors since the LAI retrieval algorithm is nonlinear. The correct computation of LAI is the parameter-aggregated algorithm which first computes LAI from the TM-derived relationship at each subpixel and then averages LAI over the TM subpixels of the AVHRR pixel. However, typically subpixel values would not be available and

Chen (1999) investigates the error in scaling up the TM-derived relationship Equation 8.5 via the band-aggregated algorithm, which uses the NDVI of the AVHRR pixel as input into Equation 8.5. Note that in the band-aggregated algorithm, non-linearity appears both in the computation of the NDVI from the NIR and red reflectance bands and in the computation of LAI from the NDVI. A theoretical error analysis based on mixed vegetation and water pixels derived a relative error that only depends on knowledge of the water fraction w within each pixel and the scaling exponents relating AVHRR NDVI and TM NDVI to LAI. The latter exponent ($TM\ NDVI = cL^b$) corresponds to 0.1844 in Equation 8.5. Constants of the AVHRR power law ($AVHRR\ NDVI = c_0 L^{b0}$) can be estimated if subpixel water fractions are known since AVHRR NDVI at the BOREAS site has a strong power law relationship with (1-w) and LAI in the mixed pixel is proportional to (1-w), with an unknown constant of proportionality equal to LAI of 100% vegetation, i.e., LAI(100%). It is shown in Chen (1999) that the relative error for a pixel is given by

$$\frac{(\mathrm{LAI}_{(\text{parameter - aggregated})} - \mathrm{LAI}_{(\text{band - aggregated})})}{\mathrm{LAI}(100\%)} = (1-w) + (1-w)^{b_0/b} \qquad (8.6)$$

where w is the fraction of the pixel covered by water. The scaling exponent b for AVHRR pixels over the BOREAS site had a value of 0.68; by differentiating Equation 8.6 with respect to w, the maximum relative error is 0.44 and occurs for $w = 0.384$. Computations of relative errors for the 990m × 990m study sites fell encouragingly close to the theoretical predictions derived from Equation 8.6.

Chen (1999) also revealed problems with linear algorithms. Although changing densities within the pixel do not introduce error into linear algorithms, scaling errors do occur when a linear relationship changes over diverse landscape components (e.g., vegetation versus water) and discontinuities are introduced. For example, the simple ratio (SR) scaled with LAI as $SR = 2.78+0.824*LAI$ at the BOREAS sites and a bias was introduced since SR over water ($LAI = 0$) is approximately 1 and not 2.78. Scaling errors are easier to derive for the nonlinear algorithm, since in the linear case error derivation requires knowledge of LAI (100%) as well as the water fraction.

With both the linear and the nonlinear algorithms, negative biases occur when an algorithm from a finer resolution is used to estimate LAI at coarser resolutions. For a pure pixel (all vegetation), the bias with the nonlinear algorithm in Chen (1999) was less than 2% and there was no bias for the linear algorithm. Hence, errors with pure pixels could be ignored. When water and vegetation were both present, negative biases occurred in Chen (1999) close to 40% for the linear algorithm and exceeded 44% for the nonlinear algorithm. Errors increased with increased heterogeneity.

Another technique for estimating LAI per pixel is based on inversion of a radiative transfer model that produces top-of-canopy reflectance in terms of leaf tissue and soil radiative properties, LAI and leaf angle distribution (LAD). From remotely sensed data, top-of-canopy reflectance is known and other parameters such as LAD and leaf optical properties may be estimated from other sources or constrained to lie within a realistic range of values. Top-of-canopy reflectance

viewed at different, but known, sun sensor geometries gives several different constrained equations which can be inverted to retrieve values of parameters like LAI. More complexity is added to the model when other elements besides vegetation canopies are included in the pixel. In this case of pixel heterogeneity, the one-dimensional (1-D) model is supplanted with a three-dimensional (3-D) model equation, which accounts for horizontal transport of photons potentially interacting with more than one element type within the pixel. Tian et al. (2002b) retrieved LAI values from AVHRR 1 km data aggregated to several coarser resolutions (8, 16, 32, and 64 km). Their imagery was acquired over North America and classified into 6 vegetation biomes and bare soil. A lookup table was used in their model equation to associate with each biome its single scattering albedo, which measures the total scattering of energy per unit volume of the biome canopy. Coarser scales result in more mixed pixels. However, equating reflectance of mixed pixels with subpixel reflectance averages in the model equation neglects the effects of heterogeneous scattering elements within the pixel on the radiative regime and can lead to significant errors in the retrieval of LAI (Tian et al. 2003). Tian et al. (2003) found large LAI errors when forests were minority biomes within non-forest pixels and developed a spatial resolution-dependent radiative transform formulation. In this formulation, the single scattering albedo is adjusted to become a weighted average of the single scattering albedos of the six biomes with weights equal to the fractional cover within the pixel. Hence, again, knowledge of fractional cover is a requirement for estimating LAI.

8.4.2 Spectral Mixture Analysis

During the 1980's, researchers began to examine spectral mixture analysis (SMA) as a means to characterize subpixel heterogeneity by modeling a pixel's reflectance as a linear combination of the reflectance spectra of ground components (e.g., soil, green vegetation, dead vegetation, rock, etc.), called endmembers (Adams and Adams 1984, Adams et al. 1986, Smith et al. 1990). The coefficients in the linear model should lie between 0 and 1 and may be constrained to sum to 1. Physically, they correspond to the fractional coverages of the ground components in the pixel. Consequently, SMA is a promising tool for providing heterogeneity parameters needed to extend algorithms from finer to coarser resolution.

Asner and colleagues (Asner et al. 1997, Asner et al. 1998a) combined SMA with an easily inverted 1-D model to calculate LAI for woody and herbaceous vegetation types in a complex savanna landscape. Inversion of the 1-D model for each of the cover types, in effect, accounted for the spatially heterogeneous landscape and, avoided a computationally intense inversion of a 3-D model. The study was initiated with high resolution spectral mixture analysis (Landsat TM) to compute fractional cover of trees, shade, senescent grass, bare soil and water. A suite of AVHRR images was acquired over the same area at different sun-sensor geometries. For each AVHRR image, the SMA model using the TM fraction covers was inverted to produce NIR and red reflectance values for tree, grass, shade and soil at the 1-km scale of the AVHRR pixel (Asner et al. 1997). Inconsistent shade fractions resulting from different sun-sensor geometries were corrected using a

geometrical-optical model (Li and Strahler 1992) prior to the inversions. These angular reflectances were used with a 1-D radiative transfer model to compute LAI for each vegetation type (tree and grass). Regional canopy LAI was computed by multiplying the type LAIs by their respective fractional covers. SMA in this example was critical not only because of the efficiency of inverting a 1-D radiative transfer model to compute LAI for two vegetation types (tree and grass), but also because inversion of the AVHRR SMA model produced endmember reflectance values for the AVHRR instrument, resolution and sun-sensor geometry.

The promise of SMA for solving scaling problems must be tempered, of course, by recognition of several sources of errors in the mixture model. First of all, multiple scattering of photons between different ground components can invalidate the linearity assumption. However, fraction errors from nonlinear mixing can be minimized by acquiring images from view angles close to the hot spot direction (Villeneuve et al. 1998). SMA requires the knowledge of endmembers, which are reflectance spectra of pure ground components. Selection of endmember spectra is the most difficult task in SMA and the most profound source of cover fraction errors. Perhaps, the most common methods of acquiring endmembers are collecting them from the field or picking pixels from the image that are homogeneously covered by one ground component. However, it is very difficult to align field-collected spectra with image spectra even after (or because of) atmospheric correction or conversion from radiance to reflectance (Kruse and Dwyer 1993). Moreover, remotely sensed images over arid and semi-arid landscapes may not have at their resolution pure pixels of green vegetation and in this situation using image endmembers will distort all cover fractions. A promising solution to the endmember selection problem has been methods that derive endmembers from the variance structure of the data (Boardman 1993, Bateson and Curtiss 1996) using principal component analysis. These derived endmembers do not necessarily coincide with pixel reflectances and may represent pure spectra when there are no pure pixels in the image. Moreover, since they are derived from the image, they have been subjected to all image pre-processing. In recent years, the assumption that each ground component is represented by a unique spectral signature has been questioned and new mixture models (Asner and Lobell 2000, Bateson et al. 2000) have been devised that substitute for a single endmember spectrum a bundle or collection of spectra representing endmember variability. Bundle unmixing produces ranges of possible fraction values (Bateson et al. 2000) or mean and standard deviation (Asner and Lobell 2000).

8.5 CALIBRATION AND VALIDATION

Tracking ground characteristics such as LAI, fAPAR and biomass with indices such as the NDVI derived from spectral reflectance requires ground sampling to establish functional relationships between the ground characteristics and reflectance-based indices, whose values have no direct physical interpretation. Regression modeling and curve fitting are common methods used to determine from image pixel values and corresponding ground samples the best equation to relate the imagery to the landscape. Because typical pixel sizes range from 20m to 1km and the extent of the

imagery is usually very large, efficient sampling strategies adequately representing image and ground variability are needed. Random sampling with a sample size $n = (\sigma t / e)^2$, where σ is the standard deviation, e the desired error and t the Student's t-value for a 95% confidence interval and (n-1) degrees of freedom, can result in unnecessary ground sampling since it does not consider spatial correlations. That is, sampling neighboring pixels or neighboring regions within a pixel with correlated values can introduce costly redundancies. Systematic sampling based on spatial statistics may achieve results at least as good with a fraction of the effort.

An important tool of spatial statistics is the semivariogram, which is a function $\gamma(h)$ measuring the average dissimilarity between parameter values sampled at ground locations or computed for image pixels that are h units apart. The semivariogram at lag h is mathematically defined by

$$\gamma(h) = \frac{1}{2k(h)} \sum_{i=1}^{k(h)} (V(x_i) - V(x_i + h))^2 \tag{8.7}$$

where $V(x_i)$ is the value of the parameter at ground location or pixel x_i, $V(x_i+h)$ is the value at a location or pixel h units away and $k(h)$ is the number of differences at lag h. In many natural scenes, $\gamma(h)$ will increase with h since nearby locations have similar characteristics compared to those at a distance. Since semivariance in the field or in an image is computed for discrete lag values, a continuous mathematical curve C is fitted to the scatter plot of lag versus semivariance in order to analyze spatial patterns. A spherical model (Isaaks and Srivastava 1989) is most commonly used for the fit and has three properties useful for understanding spatial correlations and deriving sampling strategies. When a spherical model is used, the graph of C increases with h until it reaches a plateau P at the lag S. P (referred to as the *sill* of the semivariogram) estimates the true variance of the data and S (called the *range*) is the lag distance at which values become uncorrelated and represents the typical size of objects in the scene. The value $\lim_{h \to 0} \gamma(0) = c0$ is called the nugget and its deviation from 0 may be due to fine scale or subpixel variance, measurement error or fitting with an incorrect model (Isaacs and Srivastava 1989).

Tian et al. (2002b) devised sampling strategies for validating MODIS LAI products by decomposition of semivariograms into hierarchical components (e.g., semivariances of forest, stands and trees) to reveal the spatial pattern of different characteristic scales within the scene. Other applications of the semivariogram to ground validation and calibration can be found in (Curran 1988).

Atkinson et al. (2000) used semivariograms to determine two scales of variations for biophysical properties of mean tree diameter at breast height, mean diameter at first leafing branch and tree density in a tropical forest of Cameroon Africa. Two sampling strategies were investigated in the analysis. One strategy maintained 1 ha subplots, while the other strategy averaged them to obtain a larger nugget area. Large nugget values relative to the sill for subplot semivariograms suggested variation at the 1 ha scale. The sill was reached at 20-25 km for all semivariograms.

From the analysis, the authors concluded that AVHRR 1 km data could capture the large-scale variance.

Atkinson et al. (2000) also evaluated within-pixel sampling strategies based on ordinary block kriging that is a best linear unbiased estimation (BLUE) method that uses the semivariogram to approximate mean values from a set of sample points. Block kriging determines the weights in the weighted average

$$\overline{V}(P) = \sum_{i=1}^{k} w_i V(x_i) \tag{8.8}$$

where $V(x_i)$ is the parameter value at point x_i and each x_i is in the ground pixel P whose mean value $\overline{V}$ is being approximated. To insure an unbiased estimate (i.e., mean error of the estimator is 0), weights w_i are constrained to sum to one. The weights are also constrained to minimize the variance σ_k of the errors. This variance is referred to as the block kriging or estimation variance. Under the BLUE constraints the estimation variance is

$$\sigma_k = \sum wi\bar{\gamma}(xi,P) + \mu - \bar{\gamma}(P,P) \tag{8.9}$$

where μ is the Lagrange parameter, $\bar{\gamma}(x_i,P)$ is the integral semi-variance between the pixel P and x_i, and $\bar{\gamma}(P,P)$ is the within pixel variance (for more details, *see* Atkinson et al. 2000, Isaaks and Srivastava 1989). Note that σ_k does not depend on the particular values of $V(x_i)$, but only on the semivariogram and the spatial pattern of x_i's. Hence, given the semivariogram, the kriging variance can be computed for any sampling strategy. Burgess et al. (1981) used block kriging to show that systematic sampling is more efficient than random sampling. Other researchers have used it to determine sample sizes needed for a specified precision (Webster et al. 1989).

Atkinson et al. (2000) examined two sampling strategies for scaling up 1 ha subplots for comparison with 1km AVHRR pixels. The subplots were arranged in an equilateral triangular pattern with each triangle constituting a plot. Although the estimation variances derived from representing each plot by the average of the three subplots were approximately 3 times less than those calculated as if a single subplot value was used to represent the triangular plot, the regression precisions obtained with the two sampling strategies were very close. They differed by factors of 1.02 (basal area), 1.03 (biomass) and 1.03 (tree density). That is, a sampling strategy based on intensive fieldwork may yield considerably better estimates of ground parameters than a less costly one without substantially improving the precision of the regression model to predict values in unknown locations. Regression modeling and prediction are, of course, the ultimate goals.

In conclusion, sampling and validation strategies based on semivariance and kriging can prove the sufficiency of smaller, more easily obtainable sample sizes.

However, even these sizes can be unnecessarily large when the purpose of the data (i.e., to develop regression equations) is ignored.

8.6 CONCLUSIONS

Biophysical variables (structural and functional) needed to track global environmental change must be collected at large scales that require the use of remotely sensed data. However, field studies are still necessary to relate remotely sensed parameters to their landscape counterparts. Consideration of the scaling problems inherent in extrapolating from the field to the image is critical to the use of remote sensing as a tool. An awareness, at the very least, of the sources of variance within a reflectance observation is important, as the structure of the canopy (leaf area, presence of senescent material, etc.) and landscape (canopy closure, background, etc.) will strongly influence the reflectance signal, and hence the biophysical interpretation.

When remote sensing algorithms are nonlinear (e.g., computing LAI from NDVI, radiative transfer inversions), errors resulting from scaling from high to low resolution are mainly due to increased mixing of ground components in the larger pixel. Maps of the landscape based on spectral mixture analysis, classification or other methods are useful in measuring and correcting for this error. However, classification accuracy is sensitive to pixel size relative to the size of objects in the scene since relative pixel size impacts within-class variance and the level of classes (i.e., tree stands versus forests) that can be mapped. Pixel size also impacts selection of endmembers for SMA when pure image endmembers are being sought, since pixel heterogeneity increases with pixel size. However, methods for constructing endmembers from the variance structure of the data are promising techniques for retrieving endmembers from the image when no pure pixels reside in the imagery. Scaling up parameter values from the ground to the image requires calibration and validation. Spatial statistics provides tools (local variance, semivariogram and kriging) for determining adequate distances between sample locations and testing efficiency versus accuracy trade-offs for various sampling strategies.

Moving from high resolution (e.g., TM) to coarse resolution imagery (e.g., AVHRR, MODIS) has scaling challenges, but it seems that, with adequate measures of surface heterogeneity through such methods as spectral mixture analysis and land cover classification, the problems are not insurmountable. The generality needed to gain perspective of the large-scale properties of ecological phenomena is attainable through remote sensing, yet we must understand the tool well enough to accurately accomplish the scaling operations we need. In concert with this, an active and reciprocal connection between remote sensing, ecological field studies, and scaling theory is important to guide scaling efforts and allow for the “surprises” which deepen our insights into the general behaviors of ecological systems.

ACKNOWLEDGEMENTS

We thank Jianguo Wu for organizing this workshop and book. We also thank NASA for funding our projects and others in order to explore large-scale ecology and ways to get there.

REFERENCES

Adams, J. B., and J. Adams. 1984. Geological mapping using Landsat MSS and TM images: removing vegetation by modeling spectral mixtures. 3rd Thematic Conference: "Remote Sensing for Exploration Geology" ERIM 2:615-622.

Adams, J. B., J. O. Smith, and P. E. Johnson. 1986. Spectral mixture modeling: a new analysis of rock and soil types at the Viking Lander I site. Journal of Geophysical Research 91:8098-8112.

Aman, A., H. P. Randriamanantena, A. Podaire, and R. Frouin. 1992. Upscale integration of normalized difference vegetation index: the problem of heterogeneity. IEEE Transactions in Geoscience and Remote Sensing 30:326-338.

Asner, G. P. 1998. Biophysical and biochemical sources of variability in canopy reflectance. Remote Sensing of Environment 64:234-253.

Asner, G. P. 2000. Contributions of multi-view angle remote sensing to land-surface and biogeochemical research. Remote Sensing Reviews 18:137-162.

Asner, G. P., and D. B. Lobell. 2000. A biophysical approach for automated SWIR unmixing of soils and vegetation. Remote Sensing of Environment 71:99-112.

Asner, G. P., and C. A. Wessman. 1997. Scaling PAR absorption from the leaf to landscape level in spatially heterogeneous ecosystems. Ecological Modelling 103:81-97.

Asner, G. P., C. A. Bateson, J. L. Privette, N. E. Saleous, and C. A. Wessman. 1998a. Estimating vegetation structural effects on carbon uptake using satellite data fusion and inverse modeling. Journal of Geophysical Research 103:28839-28853.

Asner, G. P., C. A. Wessman, and S. Archer. 1998b. Scale dependence of absorption of photosynthetically active radiation in terrestrial ecosystems. Ecological Applications 8:1003-1021.

Asner, G. P., C. A. Wessman, C. A. Bateson, and J. L. Privette. 1998c. Impact of tissue, canopy, and landscape factors on the hyperspectral reflectance variability of arid ecosystems. Remote Sensing of Environment 74:69-84.

Asner, G. P., C. A. Wessman, and J. L. Privette. 1997. Unmixing the directional reflectances of AVHRR sub-pixel landcovers. IEEE Transactions in Geoscience and Remote Sensing 35:868-878.

Atkinson, P. M., G. M. Foody, P. J. Curran, and D. S. Boyd. 2000. Assessing the ground data requirements for regional scale remote sensing of tropical forest biophysical properties. International Journal of Remote Sensing 21:2571-2587.

Bateson, C. A., and B. Curtiss. 1996. A manual endmember selection method and spectral unmixing. Remote Sensing of Environment 55:229-243.

Bateson, C. A., G. P. Asner, and C. A. Wessman. 2000. Endmember bundles: a new approach to incorporating endmember variability into spectral mixture analysis. IEEE Transactions in Geoscience and Remote Sensing 38:1083-1094.

Boardman, J. 1993. Automated spectral unmixing of AVIRIS data using convex geometry concepts. Pages 11-14 *in* Proceedings of the 4th Annual JPL Airborne Geoscience Workshop, Jet Propulsion Laboratory, Pasadena, CA.

Braswell, B. H., D. S. Schimel, J. L. Privette, B. Moore III, W. J. Emery, E. W. Sulzman, and A. T. Hudak. 1996. Extracting ecological and biophysical information from AVHRR optical data: an integrated algorithm based on inverse modeling. Journal of Geophysical Research 101:23335-23348.

Burgess, T.M., Webster, R., and McBratney. 1981. Optimal interpolation and isarithmic mapping of soil properties. IV. Sampling strategy. Journal of Soil Science 32:643-659.

Chen, J. M. 1999. Spatial scaling of a remotely sensed surface parameter by contexture. Remote Sensing of Environment 69:30-42.

Chen, J. M., and J. Cihlar. 1996. Retrieving leaf area index of a boreal conifer forest using Landsat TM images. Remote Sensing of Environment 55:153-162.

Chopping, M. J., A. Rango, K. M. Havstad, F. R. Schiebe, J. C. Ritchie, T. J. Schmugge, A. N. French, L. H. Su, L. McKee, and M. R. Davis. 2003. Canopy attributes of desert grassland and transition communities derived from multiangular airborne imagery. Remote Sensing of Environment 85:339-354.

Curran, P. J. 1988. The semivariogram in remote sensing: an introduction. Remote Sensing of Environment 37:493-507.

Field, C. B., J. T. Randerson, and C. M. Malmstrom. 1995. Global net primary production: combining ecology and remote sensing. Remote Sensing of Environment 51:74-88.

Friedl, M. A., D. K. Melver, J. C. F. Hodges, X. Y. Zhang, D. Muchoney, A. H. Strahler, C. E. Woodcock, G. S., S. A., C. A., A. Baccini, F. Gao, and C. Schaaf. 2002. Global land cover mapping from MODIS: algorithms and early results. Remote Sensing of Environment 83:287-302.

Gao, B. C., A. F. H. Goetz, and K. B. Heidebrecht. 1993. Derivation of scaled surface reflection from AVIRIS data. Remote Sensing of Environment 44:165-178.

Harte, J. 2002. Toward a synthesis of the Newtonian and Darwinian worldviews. Physics Today 10:29.

Hay, G. J., D. J. Marceau, P. Dubê, and A. Bouchard. 2001. A multiscale framework for landscape analysis: object-specific analysis and upscaling. Landscape Ecology 16:471-490.

Hay, G. J., K. O. Niemann, and D. G. Goodenough. 1997. Spatial thresholds, image-objects, and upscaling: a multiscale evaluation. Remote Sensing of Environment 62:1-19.

Hlavka, C. A., and J. L. Dungan. 2002. Areal estimates of fragmented land cover: effects of pixel size and model-based corrections. International Journal of Remote Sensing 23:711-724.

Hu, Z., and S. Islam. 1997. A framework for analyzing and designing scale invariant remote sensing algorithms. IEEE Transactions in Geoscience and Remote Sensing 35:747-754.

Isaaks, E. H., and R. M. Srivastava. 1989. An Introduction to Applied Geostatistics. Oxford University Press, New York.

Jin, Y., F. Gao, C. B. Schaaf, X. Li, A. H. Strahler, C. J. Bruegge, and J. V. Martonchik. 2002. Improving MODIS surface BRDF/albedo retrieval with MISR multiangle observations. IEEE Transactions in Geoscience and Remote Sensing 40:1593-1604.

Kruse, F. A. and J.L. Dwyer. 1993. The effects of AVIRIS atmospheric calibration methodology on identification and quantification mapping of surface mineralogy, Drums Mountain, Utah. Pages 101-104 *in* Proceedings of the 4th Annual JPL Airborne Geoscience Workshop, Jet Propulsion Laboratory, Pasadena, CA.

Kuusk, A. 1995. A fast, invertible canopy reflectance model. Remote Sensing of Environment 51:342-350.

Lawton, J. 2001. Earth system science. Science 292:1965.

Li, X., and A. H. Strahler. 1992. Geometric-optical directional reflectance modeling of the discrete crown vegetation canopy: effect of crown shape and mutual shadowing. IEEE Transactions in Geoscience and Remote Sensing 30:276-292.

Liang, S., and A. H. Strahler. 1993. An analytical BRDF model of canopy radiative transfer and its inversion. IEEE Transactions in Geoscience and Remote Sensing 31:1081-1095.

Lotsch, A., Y. Tian, M. A. Friedl, and R. B. Myneni. 2003. Land cover mapping in support of LAI and FPAR retrievals from EOS-MODIS and MISR: classification methods and sensitivities to errors. International Journal of Remote Sensing 24:1997-2016.

Marceau, D. J., and G. J. Hay. 1999. Remote sensing contributions to the scale issue. Canadian Journal of Remote Sensing 25:357-366.

Martin, M. E., and J. D. Aber. 1997. High spectral resolution remote sensing of forest canopy lignin, nitrogen, and ecosystem processes. Ecological Applications 7:431-443.

Myneni, R. B., and G. Asrar. 1993. Radiative transfer in three-dimensional atmosphere-vegetation media. Journal of Quantitative Spectroscopy and Radiative Transfer 49:585-598.

Myneni, R. B., J. Ross, and G. Asrar. 1989. A review on the theory of photon transport in leaf canopies. Agricultural Forest Meteorology 45:1-153.

Peterson, G. D. 2000. Scaling ecological dynamics: self-organization, hierarchical structure, and ecological resilience. Climatic Change 44:291-309.

Prince, S. D. 1999. What practical information about land-surface function can be determined by remote sensing? Where do we stand? *in* J. D. Tenhunen and P. Kabat, editors. Integrating Hydrology, Ecosystem Dynamics, and Biogeochemistry in Complex Landscapes. Wiley, Chichester, UK.

Privette, J. L., R. B. Myneni, C. J. Tucker, and W. J. Emery. 1994. Invertibility of a 1-D discrete ordinates canopy reflectance model. Remote Sensing of Environment 48:89-105.

Qi, J., Y. H. Kerr, M. S. Moran, M. Weltz, A. R. Huete, S. Sorooshian, and R. Bryant. 2000. Leaf area index estimates using remotely sensed data and BRDF models in a semiarid region. Remote Sensing of Environment 73:18-30.

Qu, Z., A. F. H. Goetz, and K. B. Heidebrecht. 2000. High-accuracy atmospheric correction for hyperspectral data (HATCH). Pages 373-380 *in* Proceedings of the 9th Annual JPL Airborne Geoscience Workshop, Jet Propulsion Laboratory, Pasadena, CA.

Ross, J. K. 1981. The Radiation Regime and Architecture of Plant Stands. Kluwer, Boston, MA.

Running, S. W., T. R. Loveland, and L. L. Pierce. 1994. A vegetation classification logic based on remote sensign for use in global biogeochemical models. Ambio 23:77-89.

Sabol, D. E., A. R. Gillespie, J. B. Adams, M. O. Smith, and C. J. Tucker. 2002. Structural stage in Pacific Northwest forests estimated using simple mixing models of multispectral images. Remote Sensing of Environment 80:1-16.

Sellers, P. J. 1987. Canopy reflectance, photosynthesis, and transpiration. II. The role of biophysics in the linearity of their interdependence. International Journal of Remote Sensing 6:1335-1372.

Smith, M. L., S. V. Ollinger, M. E. Martin, J. D. Aber, R. A. Hallett, and C. L. Goodale. 2002. Direct estimation of aboveground forest productivity through hyperspectral remote sensing of canopy nitrogen. Ecological Applications 12:1286-1302.

Smith, M. O., S. L. Ustin, J. B. Adams, and A. R. Gillespie. 1990. Vegetation in desert I. A regional measure of abundance from multispectral images. Remote Sensing of Environment 31:1-26.

Tian, Y., Y. Wang, Y. Zhang, Y. Knyazikhin, J. Bogaert, and R. B. Myneni. 2003. Radiative transfer based scaling of LAI retrievals from reflectance data of different resolutions. Remote Sensing of Environment 84:143-159.

Tian, Y., C. E. Woodcock, Y. Wang, J. L. Privette, N. V. Shabanov, L. Zhou, Y. Zhang, W. Buermann, J. Dong, B. Veikkanen, T. Hame, K. Andersson, M. Ozdogan, Y. Knyazikhin, and R. B. Myneni. 2002a. Multiscale analysis and validation of the MODIS LAI product I. Uncertainty assessment. Remote Sensing of Environment 83:414-430.

Tian, Y., C. E. Woodcock, Y. Wang, J. L. Privette, N. V. Shabanov, L. Zhou, Y. Zhang, W. Buermann, J. Dong, B. Veikkanen, T. Hame, K. Andersson, M. Ozdogan, Y. Knyazikhin, and R. B. Myneni. 2002b. Multiscale analysis and validation of the MODIS LAI product. II. Sampling strategy. Remote Sensing of Environment 83:431-441.

van Leeuwen, W. J. D., A. R. Huete, C. L. Walthall, S. D. Prince, A. Begue, and J. L. Roujean. 1997. Deconvolution of remotely sensed spectral mixtures for retrieval of LAI, fAPAR and soil brightness. Journal of Hydrology 188-189:697-724.

Villeneuve, P. V., S. A. Gerstl, and G. P. Asner. 1998. Estimating nonlinear mixing effects for arid vegetation scenes with MISR channels and observation directions. Proceedings of the International Geoscience and Remote Sensing Symposium 3:31-35.

Webster, R., P. J. Curran, and J. W. Munden. 1989. Spatial correlation in reflected radiation from the ground and its implications for sampling and mapping by ground based radiometry. Remote Sensing of Environment 29:67-78.

Wessman, C. A. 1992. Spatial scales and global change: bridging the gap from plots to GCM grid cells. Annual Review of Ecology and Systematics 23:175-200.

Wessman, C. A., and G. P. Asner. 1998. Ecosystems and the problems of measurement at large spatial scales. Pages 346-371 *in* P. M. Groffman and M. L. Pace, editors. Successes, Limitations, and Frontiers in Ecosystem Ecology. Springer-Verlag, New York.

Wessman, C. A., J. D. Aber, D. L. Peterson, and J. M. Melillo. 1988. Remote sensing of canopy chemistry and nitrogen cycling in temperate forest ecosystems. Nature 335:154-156.

Wessman, C. A., C. A. Bateson, and T. L. Benning. 1997. Detecting fire and grazing pattern in tallgrass prairie using spectral mixture analysis. Ecological Applications 7:493-511.

Wu, J. 1999. Hierarchy and scaling: extrapolating information along a scaling ladder. Canadian Journal of Remote Sensing 25:367-380.

Zagolski, F., V. Pinel, J. Romier, D. Alcayde, J. P. Gastellu-Etcheforry, G. Giordano, G. Marty, and E. Mougin. 1996. Forest canopy chemistry with high spectral resolution remote sensing. International Journal of Remote Sensing 17:1107-1128.

PART II

CASE STUDIES

CHAPTER 9

CARBON FLUXES ACROSS REGIONS

Observational Constraints at Multiple Scales

BEVERLY E. LAW, DAVE TURNER, JOHN CAMPBELL,
MICHAEL LEFSKY, MICHAEL GUZY, OSBERT SUN,
STEVE VAN TUYL, AND WARREN COHEN

9.1 INTRODUCTION

Scaling biogeochemical processes to regions, continents, and the globe is critical for understanding feedbacks between the biosphere and atmosphere in the analysis of global change. This includes the effects of changing atmospheric carbon dioxide, climate, disturbances, and increasing nitrogen deposition from air pollution (Ehleringer and Field 1993, Vitousek et al. 1997). Quantification and uncertainty analysis of carbon pools and fluxes by terrestrial biota is needed to guide policy and management decisions. Unanswered questions include: (1) how and where is the terrestrial biosphere currently sequestering carbon? (2) how might forests be managed to maximize carbon sequestration? Managed carbon sequestration would have to be optimized within and among geographic regions with attention to how this might affect biodiversity and how to manage for the effects of "natural" disturbances on carbon storage and fluxes.

Processes in the terrestrial biosphere are dynamic and occur over a wide range of spatial and temporal scales. For example, net ecosystem production is the net effect of several large fluxes: photosynthetic uptake, and release of carbon dioxide (CO_2) by respiration from autotrophs (plants) and heterotrophs (e.g., microbial decomposition). Scales range from micrometers and microseconds (e.g., cellular processes such as photosynthesis) to kilometers and centuries (e.g., decomposition of recalcitrant pools of soil carbon) (Figure 9.1). Disturbance can have a significant effect on CO_2 loss to the atmosphere through decomposition of necromass such as that left from logging of forests, or pulses of CO_2 to the atmosphere from fire, and through manufacturing of forest products. Interannual variation in climate can influence photosynthesis and respiration differently such that net CO_2 uptake can

J. Wu, K.B. Jones, H. Li, and O.L. Loucks (eds.),
Scaling and Uncertainty Analysis in Ecology: Methods and Applications, 167–190.

change significantly. These factors complicate any simple scheme to quantify carbon storage and fluxes from ecosystems, yet it is important to know how disturbance and climate interact to affect biogeochemical processes for more informed management at the regional scale.

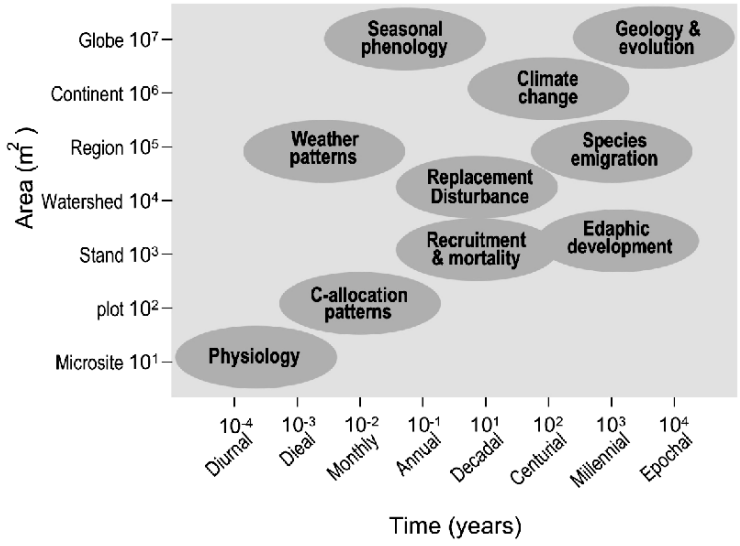


Figure 9.1. *Temporal and spatial scales of major processes affecting forest ecosystems.*

Scaling strategies in terrestrial processes often involve upscaling process data at a scale smaller than the scale of interest (e.g., leaf-level photosynthesis), and combining this information with structural and environmental data to quantify process rates at the scale of interest (Aber et al. 1993, Jarvis 1995, Wu and Li, Chapters 1 and 2). Such strategies take into account the feedbacks among components (e.g., atmosphere and vegetation), and linkages across scales (e.g., leaf-stand-landscape), and require that models are used and tested at each scale.

In this chapter, we demonstrate an approach to using field observations, remote sensing tools, and a biogeochemistry model (Biome-BGC) in a spatially nested hierarchy (Wu 1999) to improve predictions of carbon pools, productivity, and net ecosystem production (NEP) for every square kilometer of forests in a region. We examine uncertainty in a variety of ways for the different levels of data analysis.

9.2 ISSUES IN SCALING ECOSYSTEM PROCESSES

The interaction of processes operating at different spatial and temporal scales is one of the greatest challenges to regional estimates of biogeochemical processes. At any scale there is heterogeneity in types and rates of processes. For example, photosynthetic rates vary within a tree canopy (Reich et al. 1997). At larger scales,

the rates vary among different developmental stages of a given tree species as an indirect result of disturbance, and among different plant functional groups or biomes. Field observations are necessarily limited in scope, and logistical considerations make it infeasible to measure everything everywhere all of the time. Moreover, we simply do not understand some processes (e.g., respiration, carbon allocation within plant) as well as we understand others (e.g., photosynthesis). Simple aggregation and extrapolation schemes to larger spatial and longer temporal scales using field observations miss the critical role of feedbacks, which is an essential feature of scaling (Jarvis 1995, Wu and Li, Chapter 2).

Scientists have studied ecological systems for decades through observational studies, experiments, and development of models that incorporate their understanding of ecosystem function. Process models can be used to extend knowledge across time and space and to test hypotheses about coupling of processes and responses to environmental conditions. Considerable efforts have been expended to develop biogeochemical models to examine terrestrial ecosystem responses to global change (Melillo et al. 1993, Cramer et al. 2001). However, the connections between climate, soil conditions, and vegetation dynamics are poorly understood and are highly simplified in most models.

Biogeochemistry models are quantitative representations of our understanding of the storage and transport of carbon, water and nutrients through soil, vegetation and the atmosphere. They incorporate nonlinearity of processes, and the multiple scales and interactions of processes. They generally operate in two-dimensional space and are typically run with hourly to daily climate data over years to centuries. A limitation of such process models is that they require mass and energy balance, sometimes at temporal and spatial scales that may not make sense relative to how processes actually occur. Mass and energy balance also cannot necessarily be quantified with field measurements to test model assumptions. For example, flux sites that measure ecosystem energy components have found that on average, they can only account for ~80% of net radiation (Wilson et al. 2002). As a result, models are often evaluated by comparison with measurable budget components (e.g., Law et al. 2001a).

Biogeochemistry models use input parameters for the physiology, biochemistry, structure, and allocation patterns of vegetation functional types, or biomes. For single-stand simulations it is possible to measure many of the required model parameters, but as spatial coverage increases, data availability decreases, and generalized biome parameterizations are applied. For example, parameterization may be simplified to constant foliar nitrogen across a biome or life form in a region using data from the literature. Similarly, allocation of carbon to plant tissues may be assigned as fixed fractions across age classes and climatic zones. Undocumented parameter selection and unknown model sensitivity to parameter variation for larger-resolution simulations are currently a major limitation to regional and global modeling (White and Running 1994, White et al. 2000). Although some ecosystem process models are dynamic and converge towards carbon, nitrogen and water balances, they can result in the right answer for the wrong reasons, or a predicted variable such as net primary productivity can be quite inaccurate because of a variety of uncertainties in model structure or parameters.

Decision rules are usually developed for selecting parameters in regional modeling. The procedure is usually to identify the simplest parameterizations (the default variables) and to then test the model to determine which parameters need more specification. Another method for parameterization is data assimilation, which has been used for a long time in atmospheric research, but it is relatively new to scaling ecosystem processes (e.g., Cescatti 1997). Data assimilation is the process of finding the model representation that is most consistent with the observations. Data assimilation usually proceeds sequentially in time. The model organizes and propagates forward the information from previous observations. Information from new observations is used to modify the model state, and to be as consistent with them and the previous observations (e.g., time series in Kalman filter, a Bayesian approach). As more data become available, such an approach may make sense for some parameters in regional carbon cycle scaling applications.

Another scaling issue is the mismatch of scales between observations and predictions. Model output variables include carbon storage in live and dead pools, net primary production (NPP), and net ecosystem production (NEP). Computational logistics and availability of spatial data for running the models may require simplifications that include linear aggregation of input and output variables. "Big leaf" models such as Biome-BGC assume a homogeneous 2-dimensional layer of foliage for resource use, carbon uptake and transpiration over a grid cell that can range from 30 m to 1 km to 0.5 degree (longitude and latitude) on a side. Evaluation of modeled NPP across a region is often conducted by comparing 1 km mean NPP values with 1 hectare means from tree structure measurements scaled by allometry, a mismatch in spatial scales when in reality, NPP may be heterogeneous within 1 km (e.g., clearcut and mature forests within 1 km; Turner et al. 2003). Temporal mismatches in scale also occur, whereby time-integration of available data and model output differ. Limited availability of field estimates of NPP or biomass across regions and continents has resulted in comparisons between model averages over years with field estimates over a variety of single or multiple years. The mismatch in time and space has uncertainties associated with it, yet this is difficult to quantify and overcome.

9.3 AN APPROACH TO SCALING AND UNCERTAINTY ANALYSIS OF ECOSYSTEM PROCESSES IN FORESTS – A CASE STUDY

The goal of the regional TERRA-PNW project is to estimate carbon storage, NPP and NEP for every square km of forest across a region over several climate years with improved accuracy, and explain sensitivity of NEP to cover type, forest age, disturbance, and interannual variability in climate. The approach is to spatially distribute a biogeochemistry model (Biome-BGC) which assimilates a wide range of spatially explicit information about the environment, using a combination of remote sensing and field observations as model input and for model testing. To treat the complexity of the carbon cycle, our modeling approach takes advantage of the near decomposability (*sensu* Wu 1999) of the ecological hierarchy (Table 9.1). At the level of the ecosystem, our emphasis is on NEP (annual time step). The associated

processes of photosynthesis and heterotrophic respiration are treated separately at the daily time step. At the ecosystem level, the key dynamic is the successional trend from negative NEP to positive NEP over multiple years. Because our emphasis here is primarily on annual NEP, we do not treat the longer time frames and processes relevant to interactions within holons at the upper levels of the hierarchy, e.g., interactions among age classes within a landscape mediated by fire, or interactions among ecoregions within a region mediated by climate change. However, differentiation of these upper levels remains useful because it permits specification of unique model parameterizations for ecophysiological constants such as specific leaf area (SLA).

Table 9.1. *Delineation of the ecological hierarchy.*

Level	*Examples*
Region	Pacific Northwest
Ecoregion	Coast Range, West Cascades
Landscape	Intensive Management, Wilderness
Ecosystem	Young, Mature, Old Growth Stands
Functional Group	Producers, Decomposers

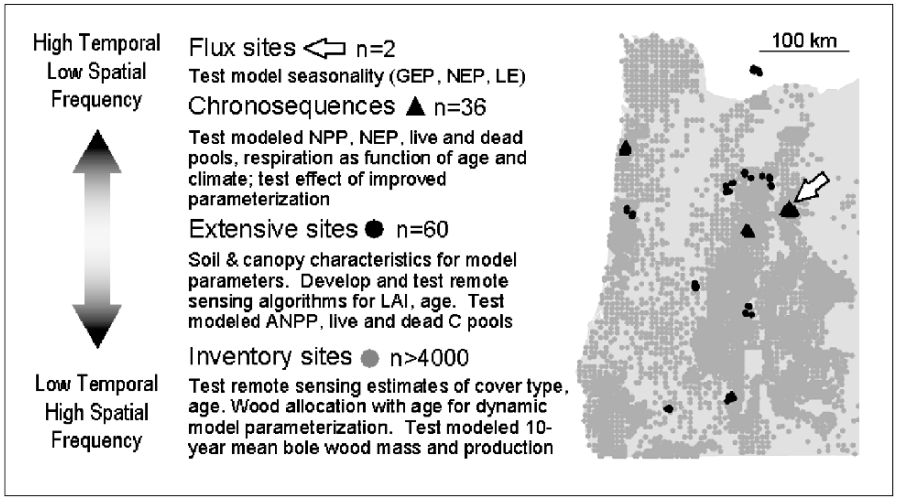


Figure 9.2. *Hierarchical approach to collecting the field data used to develop model parameters, develop remote sensing algorithms, and validate model output. Due to confidentially requirements, mapped locations are only approximate.*

Biome-BGC is fundamentally a daily time step model of coupled carbon, nitrogen, and water cycles (version 4.1.2; Thornton et al. 2002). It requires spatial data on land cover classification, stand age, and a reference leaf area index (LAI), all provided by satellite remote sensing, and it is driven with a distributed daily

climatology (DAYMET model). Model outputs included gross photosynthesis (GPP), net primary production (NPP of foliage, above- and belowground wood, fine roots), heterotrophic respiration (Rh) and NEP.

Model sensitivity tests are used to determine critical input variables to measure in the field for a range of forest types and developmental stages. White et al. (2000) conducted a sensitivity analysis with Biome-BGC and found that simulated NPP for all biomes is significantly affected by variation in foliar and fine root C:N, and NPP of woody biomes is strongly controlled by leaf nitrogen in Rubisco, maximum stomatal conductance, and SLA while non-woody biomes are sensitive to fire mortality and litter quality.

The scaling strategy is to use a spatially nested hierarchy (Wu 1999) to optimize field measurements for model parameters and testing. Field observations range from inventory data (many locations, few variables), to extensive sites, and intensive sites (chronosequences and tower flux sites, greater frequency and types of measurements, fewer locations) (Figure 9.2). Some field measurements are relatively easy to make, and are needed for the wide range of vegetation types and environmental conditions. For example, the model parameters foliar C:N and SLA can be measured at mid-season at many locations (extensive sites). More difficult measurements, such as stomatal conductance, are carried out at fewer intensive sites, or values are obtained from the literature. Remote sensing is a useful tool for obtaining spatially distributed vegetation characteristics (Wessman and Bateson, Chapter 8). A large pool of field data is required to develop and test remote sensing algorithms for vegetation mapping, so field observations are needed at many locations to cover the domain of application (e.g., LAI, forest type, and forest age at extensive sites).

To aid diagnostics, model outputs are evaluated with observations at a variety of spatial and temporal scales, starting with the most intensive observations and followed by more distributed sites that have less information. Then necessary improvements in model structure and parameters are identified and implemented, and model testing is reiterated at multiple scales.

We demonstrate the scaling approach over an east-west swath across central Oregon (300 km × 50 km) that covers a strong climatic gradient from the mild coastal conditions where water is not limiting to growth, to the Cascade Mountains where snowfall and freezing temperatures occur, to the semi-arid east side of the Cascade Mountains where temperatures are more continental (as in Figure 9.3).

9.4 FIELD OBSERVATIONS

9.4.1 Flux Sites: Measurements

Eddy covariance flux sites, such as the AmeriFlux network of sites (currently 80 sites in North, Central and South America) provide net CO_2 and water vapor exchange data. Flux systems comprise three-axis sonic anemometers that measured wind speed and virtual temperature, and infrared gas analyzers that measure concentrations of water vapor and CO_2 above the canopy. Fluxes are averaged half-hourly, and data are

evaluated for quality (Law et al. 2001a, Baldocchi 2003). Additional biological measurements typically made at the sites provide data for evaluating flux components (e.g., transpiration, respiration). Flux data are aggregated daily to examine seasonal trends in simulated and observed GPP, net ecosystem production (NEP), and latent energy flux (LE, evaporation and transpiration), and diagnose potential causes for discrepancies.

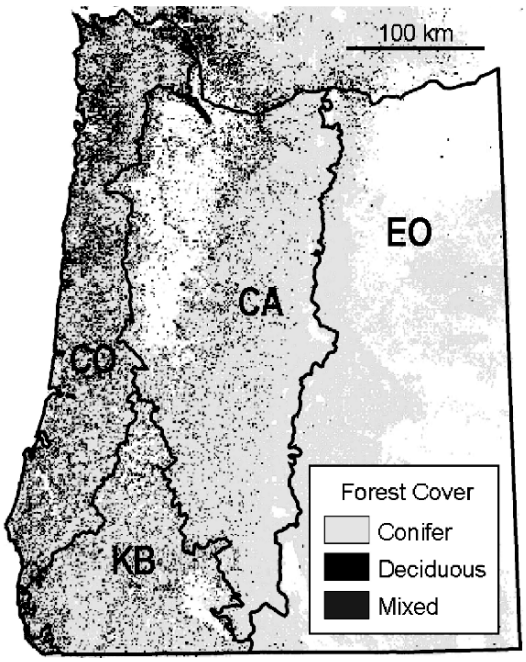


Figure 9.3. *Forest cover type map derived from remote sensing and supplementary GIS data. Grey lines denote ecoregions in each of which are recognized five forest classes (conifer, deciduous, mixed, semi-open, and open). Semi-open, and open together represent only 8% of the forested area and as such do not resolve on this figure.*

9.4.2 Flux Sites: Uncertainty Analysis

In previous studies (Anthoni et al. 1999) we quantified uncertainty in eddy flux estimates of NEP by combining systematic errors geometrically, and estimated that the overall uncertainty of the daytime carbon dioxide flux was ~±12% of the mean half-hourly flux. Nighttime fluxes are more problematic due to low wind conditions in tall canopies, and when data are screened to remove these periods, the cumulative error can result in substantial uncertainty in annual estimates of NEP. Therefore, the flux data are most useful for testing models when aggregated to a daily or monthly timestep, a range consistent with the time-step of Biome-BGC.

9.4.3 Chronosequence Plots: Measurements

To expand the range of forests types for which we had reliable estimates of NEP, 36 additional study plots were established across a strong precipitation gradient in the study region. The chronosequence plots are consist of 3 independent replicates of 4 age classes blocked by 3 forest types. Forest ages range from 10 to 300 years and are classified as either initiation, young, mature or old.

The mass balance approach we used to estimating NEP is:

$$NEP = (NPP_A - R_{WD}) + (\Delta C_{FR} + \Delta C_{CR} + \Delta C_{soil} - fine\ litterfall) \quad (9.1)$$

where NPP_A is aboveground net primary production (wood and foliage by both over- and understory plants), R_{WD} is the respiration from woody debris (decomposition of coarse and fine woody debris, stumps and snags), ΔC_{FR} is the net change in fine root C (not different from zero in this study), ΔC_{CR} is the difference between the net growth live coarse roots and the decomposition of coarse roots attached to stumps, ΔC_{soil} is the net change in mineral soil C (not different from zero in this study), and fine litterfall includes leaves and twigs <1 cm diameter falling to the ground in one year.

Procedures for measuring the components of equation 1 are detailed in Law et al. (2003) and generally rely on radial stem growth and allometric biomass equations for estimating woody production, optical measures of LAI and leaf turnover for estimating foliar production, volume inventory and decay functions for estimating dead wood respiration, and soil coring for estimating change in soil and fine root C.

9.4.4 Chronosequence Plots: Uncertainty Analysis

Both an experimental and measurement uncertainty was assessed for the NEP values (Table 9.2). For the purpose of describing the range of behavior exhibited by a certain condition class (forest type and age in this study) the most useful measure of uncertainty is an expression of the variance among true replicates of the condition, i.e., "experimental" uncertainty. This was calculated for all measured parameters, including NEP, simply as the standard deviation among the replicated plots (3 per age class). Computing the experimental uncertainty is appropriate only when there are true plot replicates (see Hurlbert 1984).

For the purposes of model validation, there is a desire to know the measurement error, which can stem from both the instrument error (e.g., calibration of carbon dioxide gas analyzer) and the error that arises from sample design (e.g., variation among soil cores used to estimate plot-level fine root mass). To assign measurement uncertainty to a composite parameter such as NEP, it is necessary to know the uncertainty associated with its components such as wood production or coarse woody debris decomposition. However, it is not possible or practical to account for all sources of uncertainty. For instance, measurement uncertainty exists in the coring of tree boles to determine their radial growth, however we know from prior analysis

that this error is insignificant compared to the extrapolation of these data to uncored trees on the same plot based on diameter-increment regressions. Consequently, the uncertainty in the radial increment component of NPP_A is based solely on the diameter-increment regressions.

Another example of using sampling error as the primary source of measurement error is in the assessment of plot-level LAI, which was measured with an LAI-2000 at 39 points regularly stratified throughout each plot. The standard deviation of these 39 measurements served as the plot-level uncertainty. Potentially, uncertainties in LAI could also be attributed to error in clumping corrections at both leaf and stand scale, but such errors are not quantifiable in practice. Fractional values of LAI (e.g., 0.5) make little difference biologically and in process modeling as indicated by model sensitivity tests (Thornton et al. 2002).

Table 9.2. *Means, experimental uncertainties, and measurement uncertainties for field estimates of aboveground net primary production (NPP_A), belowground net primary production (NPP_B), and net ecosystem production (NEP), for each age class and cover type represented by the chronosequence plots.*

	NPP_A ($gC\ m^{-2}\ yr^{-1}$)		NPP_B ($gC\ m^{-2}\ yr^{-1}$)		NEP ($gC\ m^{-2}\ yr^{-1}$)	
Cascade Head						
Initiation	793	(83, 22)*	273	(83, 95)	555	(95, 37)
Young	801	(26, 48)	259	(99, 73)	393	(104, 63)
Mature	657	(44, 59)	202	(22, 70)	423	(55, 73)
Old	486	(145, 49)	217	(68, 81)	238	(98, 71)
HJ Andrews						
Initiation	315	(24, 34)	252	(52, 118)	199	(23, 35)
Young	476	(127, 31)	234	(45, 66)	288	(115, 44)
Mature	478	(103, 40)	274	(74, 99)	314	(170, 54)
Old	318	(53, 56)	218	(34, 128)	–24	(148, 83)
Metolius						
Initiation	114	(42, 8)	94	(52, 43)	–129	(110, 17)
Young	231	(27, 19)	169	(47, 43)	117	(59, 31)
Mature	323	(151, 36)	162	(75, 62)	169	(200, 46)
Old	180	(71, 27)	152	(34, 45)	34	(121, 35)

** The first value in parentheses is the experimental uncertainty (1 SD of the mean of 3 replicate plots). The second value in parentheses is the average measurement uncertainty calculated for each site-age combination (measurement uncertainty determined for each plot by Monte Carlo simulation as 1 SD of 1000 standard normal iterations, accounting for covariance among equation components).*

Once an appropriate measurement uncertainty was assessed for each of the components of NEP (Equation 9.1), Monte Carlo simulations were used to determine a final aggregate uncertainty. By randomly sampling within the probable distribution of each variable (set by its own measurement uncertainty), a single

Monte Carlo run generates one probable value for NEP. Repeating the simulation 1000 times generates a distribution of probable NEP values. A final measurement uncertainty for NEP is expressed as the standard deviation of this distribution. In this study we assumed that uncertainty about each component of NEP had a standard normal distribution. When using Monte Carlo simulations it is important to consider the covariance among component variables. Some Monte Carlo models accommodate a correlation matrix that quantifies the covariance among equation components. A simpler alternative, employed here, was to combine equation components known to be computationally linked (such as understory wood and foliage mass which are both derived from stem diameter) into one variable before running the Monte Carlo simulation. In this study, experimental uncertainties in NEP averaged 44% and measurement uncertainties averaged 19% of the mean NEP, with the highest uncertainties in the oldest forests (Table 9.2).

9.4.5 Extensive Plots: Measurements

There is generally a large gap in field observations between the relatively low number of sites where it is feasible to make intensive measurements, and the large number of inventory sites where only a few measurements are made. To bridge this gap, we established 60 additional plots using a hierarchical random sample design that allowed maximum representation of forest types that exist in the region, the age classes present, and the climate space. A single visit to the plots provided data on soil and canopy C and N, maximum LAI, biomass and aboveground productivity for the range of environmental conditions and forest types. LAI measurements from these plots were used to develop the regressions that predicted LAI from remote imagery.

9.4.6 Extensive Plots: Uncertainty Analysis

Uncertainty computations for the extensive plots followed the same procedures described above for the intensive plots. The uncertainty in field estimates of NPP_A and live mass aboveground (LM_A) at the extensive plots averaged 8 and 9% of the means (1 SD), respectively.

9.4.7 FIA/CVS Inventory Plots: Measurements

Federal forest inventories are repeated on a large number of forested plots in Oregon that are visited relatively infrequently. Current Vegetation Survey (CVS) plots are on federal national forest lands (4468 CVS plots in Oregon) and Forest Inventory & Analysis plots (FIA) are on private lands (1120 plots). CVS and FIA sampling intervals are 10 and 8-12 years, respectively. The measurements made on these plots are primarily tree structural dimensions and species. These data are used to estimate biomass during each measurement period, and ~10-year mean stemwood production. Limited measurements of some variables, such as wood increment (33% of trees) and tree height (23% of trees), reduce confidence in the accuracy of the estimates of

biomass and productivity but remain valuable since the large number of plots can be used to evaluate trends in stemwood biomass and growth across climatic zones and forest types, and to determine relative accuracy of model predictions in these different conditions. For the inventory plots, stemwood mass was calculated from

$$Biomass_b = Volume_b \times Wood\ Density \tag{9.2}$$

where $Volume_b$ is stemwood volume, and wood density is the dry density of wood.

The ~10 year mean NPP of aboveground stemwood was estimated from

$$NPP_{Aw} = Biomass_{w2} - Biomass_{w1} \tag{9.3}$$

Where NPP_{Aw} is aboveground NPP of stemwood and $Biomass_{w2}$ and $Biomass_{w1}$ are aboveground woody biomass at current and previous time steps, respectively. Previous and current height of unmeasured trees was modeled using height-diameter equations developed in the region from forest inventory data (Garman et al. 1995). The study area was divided into four geographic regions, each corresponding to commonly acknowledged physiographic zones in Oregon (Oregon Coast Range, Western Cascades, and Eastern Cascades – after Franklin and Dyrness 1973). When possible, physiographic zone and species-specific allometric equations were applied to estimate volume. Wood density data were acquired for most of the major hardwood and softwood species of western Oregon through wood density surveys conducted by the U.S. Forest Service (USDA Forest Service 1965, Maeglin and Wahlgren 1972).

9.4.8 FIA/CVS Inventory Plots: Uncertainty Analysis

Error estimates for NPP_{Aw} are based on uncertainty in radial growth propagated through the allometric models. In Van Tuyl et al. (2005), we made estimates of the potential magnitude of error associated with using generic wood densities and non-site specific allometry. The error associated with using generic versus plot-specific wood densities on 36 plots was estimated to be about 10%. An empirical comparison of volume equations used in this study suggests that errors as high as 40% of the mean could result from using equations not developed in the study area. These results suggest that site-specific volume allometry is much more important to making quality estimates of biomass and NPP than are site-specific wood density.

9.5 REMOTE SENSING OBSERVATIONS

The role of remote sensing in this study, as in many regional studies, is to provide large-domain spatial data layers that the biogeochemistry process model requires: LAI, stand age, and forest type. In particular, the prediction of stand age throughout this ecologically diverse region required different approaches in different stand types. While reasonable continuous estimates of stand age have been made using remote sensing of closed-canopy Douglas-fir/western hemlock forests in western

Oregon (Cohen et al. 2001), similar efforts in the open forests east of the Cascade Mountains were met with only limited success. Therefore, questions of scaling over these varying stand types must inherently address the varying data precisions in each area.

9.5.1 Land Cover

Forest cover for the study area was created by updating the 1988 forest cover layer created by Cohen et al. (2001) using the same land cover classes (Table 9.3). Non-forest areas (primarily urban and agricultural areas totaling 24% of land area) were defined using masks taken from Cohen et al. (2001) and supplemented with information from the National Land Cover Database (NLCD) for the eastern portion of the study area (Vogelmann 1998). For the purposes of this study, all forested areas in the East Cascades ecoregion were considered closed coniferous forest, and their extent was fixed by the NLCD coverage. Of over 8.8 million hectares of forest, 3% was in an open condition, 8% was semi-open, 5.5% was deciduous, and 16% was mixed forest (Figure 9.3). The resulting land cover information was validated using 24 aerial photos distributed throughout the western study area, with an average accuracy of 82%, and a range of 49% to 97%.

Figure 9.4. *Leaf area index (LAI) map derived from Landsat EMT+.*

9.5.2 Forest Age

The date of stand replacing disturbance can serve as a surrogate for stand age, with the caveat that stand re-establishment periods can vary according to resource availability and competition for resources (e.g., Law et al. 2003). To increase the accuracy of the estimates of age in the earliest stages of forest succession (when carbon flux changes rapidly), the continuous estimates of stand age were combined with age estimates based on mapping of disturbance through change detection.

A map of forest age was created by using an existing 1988 dataset of Cohen et al. (2001) extending it to the east side of the Cascade Mountains. It was updated to reflect age in Year 2000, and masked to remove areas that had changed since 1988. Regression between the 2000 tasseled-cap image and the ages in 1988 were used to estimate the remaining ages needed for new conifer areas. This relationship explained 68% of variance of log-transformed age, with an RMS residual of 0.392 log-years, which were similar to the 65.9% of variance and 0.592 RMS residual of 0.57 reported by Cohen et al. (2001).

Table 9.3. *Forest cover definitions for remote sensing land cover classification.*

Cover Class	*Definition*
Non-forest	Forested Cover < 0
Open	Total Forest Cover < 30%
Semi-open	30% < Total Forest Cover < 70%
Deciduous	Total Cover > 70% and Conifer Cover < 30%
Mixed	Total Cover > 70% and 30% < Conifer Cover < 70%
Conifer	Total Cover and Conifer Cover > 70%

9.5.3 Leaf Area Index

The remote sensing estimates of leaf area index (LAI) ranged from 1 to 12 in the West Cascades and Coast Range, and 0.5 to 8 in the East Cascades (Figure 9.4). To construct LAI algorithms, LAI was measured at the 96 extensive and intensive plots following methods in Law et al. (2001b). Polygons were hand digitized around each of the plots in reference to the Landsat ETM+ scene to ensure that a homogenous region was being referenced in the comparison of spectral characteristics and LAI. Both the tasseled-cap index and NDVI indices were calculated from the ETM+ mosaic and stepwise multiple regressions were used to determine the best set of variables for predicting LAI. The resulting equation uses brightness raised to the second power and wetness raised to the power of 11.606, explains 80% of variance, and has an RMSE of 1.668 (Figure 9.5). Subsequent analysis of the residuals for the East Cascades ecoregion indicated that LAI in those plots was underestimated by ~20%. Using a combined data set of 24 plots collected in 1999 and 2001, a new coverage was calculated for the East Cascades. The resulting equation uses only the wetness variable, raised to the power of 14.876, explains 82% of variance and has an RMSE of 0.742 (Figure 9.5). Both equations explain similar percentages of

variance, but the RMSE for the East Cascades is lower, probably due to the lower number of observations in this ecoregion.

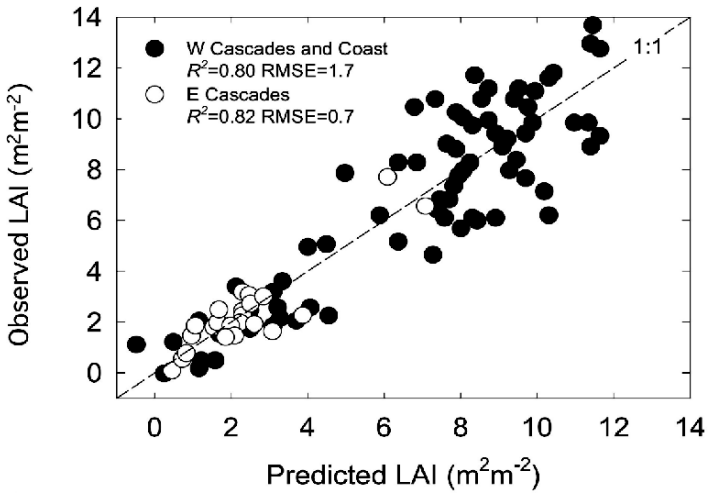


Figure 9.5. *Predicted and observed leaf area index (LAI) for field plots in the Eastern Cascades and combined plots of the Coast Range and Western Cascades.*

9.6 IMPLEMENTATION OF THE DISTRIBUTED MODELING

9.6.1 Overview

The Biome-BGC model (Thornton 1998, Thornton et al. 2002) was selected for this application because it includes the complete carbon cycle, it can assimilate input data from multiple sources (notably plot level measurements of parameters such as foliar nitrogen concentration), and it disaggregates carbon cycle processes sufficiently enough to allow comparisons with a wide variety of observations. The model has been previously tested at individual plots in coniferous forests of the PNW region (Running 1994, Law et al. 2001a).

Biome-BGC was run on a 25 m grid covering most of Oregon west of the Cascade Mountains. Much of the forested portion of the Pacific Northwest is characterized by clearcut patches smaller than 1 km^2 (Cohen et al. 2002), so high spatial resolution is essential to characterize spatial patterns in carbon flux (Cohen et al. 1996, Turner et al. 2000). Conversely, a 1 km resolution is suitable to capture much of the significant variation in climatic variables.

9.6.2 Climate Inputs

For climate inputs, Biome-BGC requires daily estimates of minimum and maximum temperature, precipitation, vapor pressure deficit, and solar radiation. A daily time

step, 1-km grid for these variables over the period of 1980-1997 was used in this application. This data set was generated by interpolation of meteorological station data using the DAYMET program (Thornton et al. 1997, Thornton and Running 1999, Thornton et al. 2000).

9.6.3 Model Spinup

Biome-BGC was specifically designed to simulate changes in carbon fluxes over long periods. The model must therefore be run over a 1000+ year "spinup" to bring the slow turnover soil carbon pools into near steady state. Disturbances such as clearcut harvests are then imposed, accounting for tree carbon affected by the disturbance. For each model run in this analysis, two successive disturbances separated by 60-90 years (depending on location) were simulated at the end of each spin-up such that 1/3 of the live tree carbon was transferred to the coarse woody debris pool at each disturbance. The model was then run forward to the age specified by the remote sensing classification. The 18-year climate time series was run repeatedly in these analyses and manipulated such that the last year of secondary succession was always 1997.

9.6.4 LAI Optimization

LAI is prognostic in BIOME-BGC, and although LAI can be remotely sensed, it cannot simply be prescribed in the model because of the continuous interaction among the various model compartments. Earlier analysis in Pacific Northwest conifer stands has shown a strong linear relationship between stand LAI and a site water balance index based on annual precipitation, annual potential evapo-transpiration, and soil water holding capacity (Grier and Running 1977). To achieve agreement between remotely sensed LAI and simulated LAI for a given cell in this application, a secant method (Cheney and Kincaid 1985) was used with model runs at different soil depths to iteratively solve for the soil depth that minimizes the difference between a reference LAI (e.g., from remote sensing) and simulated LAI at a specified stand age. An initial value of soil depth to seed the iterations was taken from a digital map of soil depths based on the State Soil Geographic (STATSGO) database (Kern et al. 1997). In young stands that may not have achieved equilibrium LAI, the minimum possible soil depth was constrained by the distribution of STATSGO soil depths in the ecoregion. Using this approach, the fit between remotely sensed LAI and Biome-BGC LAI was good ($r^2 = 0.97$) with an RMSE of 0.5 LAI units.

9.6.5 Integrating 1 km and 30 m Data

A critical issue in the distributed model implementation was the scale mismatch between land cover and LAI data at the 25 m resolution and the climate data at 1 km resolution. Because of computational constraints associated with the model spinups, a unique model run could not be made at each 25 m cell in the region of interest.

Thus, a separate model run was made only once in each 1 km cell for each combination of cover type and age class. For use in the soil depth determination, a reference LAI was determined for each cover type × age class combination within each 1 km cell. The remotely sensed reference LAI was the mean for all 25 m cells belonging to each cover type × age class combination within the 1 km cell.

9.6.6 Ecophysiological Inputs

Biome-BGC requires a set of ecophysiological constants for model initialization. We created a generic set of constants for each cover type based on the values in White et al. (2000). The White et al. (2000) analysis determined that the model was particularly sensitive to values of foliar C:N and SLA, so we also created ecoregion-specific sets of constants where foliar C:N and SLA values were means based on field measurements at the intensive and the chronosequence plots.

9.6.7 Comparisons of Simulations with the Survey Data

The comparison of simulated stemwood production with observations from the FIA survey data was constrained by a number of factors. Besides the initial problem of converting information on distributions of diameter and growth increment into wood production, these factors included (1) plot location, (2) convergence of model output and inventory data on a common parameter, and (3) achieving overlap in time between the observations and the simulations.

Perhaps the most significant issue for model evaluation inaccuracy in some plot locations. Federal law currently prohibits release of FIA plot locations for both private and public lands (recent amendment of the Food Security Act). Thus, researchers outside of the FIA program are extremely limited in their ability to conduct analyses of the data in a spatial context. In addition, locations of the CVS plots on public land were not determined with Global Positioning Systems, so the locations are somewhat uncertain.

Because much of the forested land in the study area is publicly owned, the set of CVS plots was used for the purposes of comparing survey-based and simulated wood production. CVS locations were accepted as reported and the coordinates were used to determine an associated 1 km cell within the climate grid. In the model spinups used to determine soil depth, the reference LAI was the mean value for all 25 m cells in the relevant cover class within the 1 km cell. After the soil depth was selected, the model spinup completed, and the disturbances imposed, the model was run to the age specified by the CVS data.

The problem with temporal overlap in the CVS and simulated data is caused by the nonuniformity in the year of visit to the permanent plots, and the observation that there is large interannual variation in bolewood production based on climate variation in the Pacific Northwest (Turner et al. 2003). CVS plots are visited on a roughly 10-year interval, and the growth increment is reported for the previous 10 years. There are also delays in getting the data into the FIA database. Thus data at the time it was received from FIA (in the year 2001) may have been five or more

years old. The model NPP_{Aw} can be aggregated over any interval desired, and for this study the mean NPP_{Aw} over the 10 yr period from 1988-1997 was selected for the comparisons. Because of the many constraints on achieving a implemented in this Biome-BGC application such that allocation to fine roots and leaves one-to-one comparison in space and time between the CVS data and the simulation data, comparisons were made more generally by examining relationships of NPP_{Aw} to stand age within the ecoregions.

An age-specific allocation scheme was increased in older stands, except those of the East Cascades ecoregion where empirical data indicated that allocation to fine roots is greater in young stands (Law et al. 2003). The ecological rationale for increased root and leaf allocation with age follows the nutrient limitation hypothesis, i.e., nutrients become more limited in late succession because they are increasingly sequestered in the biomass. Allocation to fine roots thus increases, and correspondingly stemwood production decreases. This is not an appropriate rationale for water-limited ecosystems, where relatively large allocation to roots throughout stand development is critical for survival.

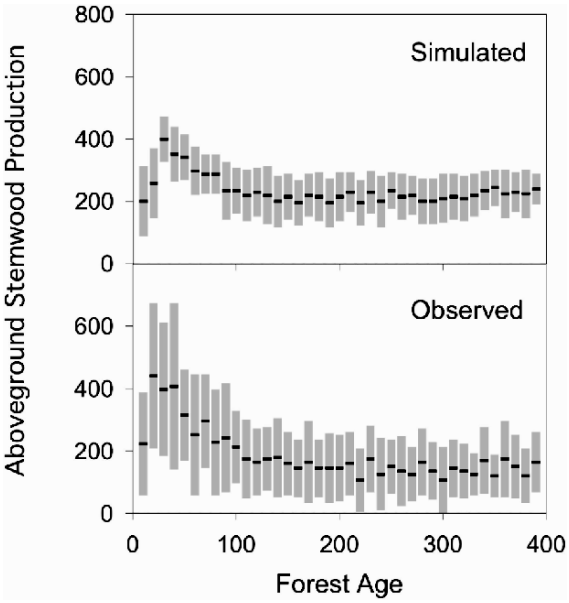


Figure 9.6. *Simulated and observed aboveground stemwood production in relation to stand age for the West Cascades CVS inventory plots. Grey bars are the standard deviation of 6 to 36 plots depending on age class.*

To implement the allocation shift in the model simulations for the Coast Range and West Cascades ecoregions, a nonlinear increase in the allocation to leaves and fine roots was prescribed in late succession. The relevant parameters were a maximum stemwood production rate in young stands, a lower stable stemwood

production rate in older stands, and an age that indicated the midpoint of the transfer. These parameters were based on a cubic polynomial fit to the CVS NPP_{Aw} data by ecoregion (Figure 9.6). With the dynamic allocation implemented in the model, the ratio of NPP_{Aw} in old and young stands came into much closer agreement with the observations (Table 9.4).

Table 9.4. *Model/data comparisons (CVS plot data) for the ratio of aboveground net primary production for wood (NPP_{Aw}) at the low and high extreme points of the cubic polynomial fit for NPP_{Aw} versus forest age. Units are g C m^{-2} y^{-1}. N is the number of observations, CI is the 95% confidence interval for the predictions at the forest ages corresponding to the high (young) and low (old) extremes of the fitted cubic polynomial model.*

	N	*Ratio*	*Young*	*CI*	*Old*	*CI*
Observations						
Coast Range	383	0.50	705	272 -1137	354	–64 - 772
West Cascades	1677	0.44	469	214 - 724	208	–45 - 462
Model-Before Dynamic Allocation						
Coast Range	373	0.92	353	263 - 443	324	243 - 413
West Cascades	1626	0.86	330	182 - 477	285	138 - 432
Model-After Dynamic Allocation						
Coast Range	373	0.69	456	353 - 559	314	211 - 416
West Cascades	1626	0.51	366	235 - 497	185	55 - 315

Across all ecoregions, the model NPP_{Aw} values were similar to the CVS data in that the highest magnitudes were in the Coast Range ecoregion, slightly lower values in the West Cascades ecoregion, and much lower values in the East Cascades. There was generally more scatter in the CVS observations than in the simulations. This may occur in part because of a tendency for the scaling approach to under-represent sites with low and high LAI (due to the necessity of averaging LAI within each cover class by age class combination over the 1 km grid cells). The permanent plot data will be further utilized for model improvement by examining the relationships of NPP_{Aw} to climate indices such as annual potential evapotranspiration, and evaluating the degree to which the model is responding in a similar fashion.

9.6.8 Comparisons of Simulations with the Extensive Plot Data

Comparisons were made between observations and modeled NPP_{Aw}, total aboveground NPP (NPP_A), and stem mass at the 96 extensive plots (including the intensive chronosequence plots). The model was run with the same protocols as for the CVS plots but used field observations of LAI as the reference LAI in determining soil depth. The comparisons (Table 9.5) showed good agreement for

NPP_{Aw} and NPP_A, but a tendency for the model to underestimate stem mass at older stand ages. A potential cause of stem mass underestimation is overestimation of tree mortality and those relationships are being explored with model sensitivity analyses and with evaluation of mortality estimates in the literature.

Table 9.5. *Regression statistics (observed vs. modeled) for extensive plot and chronosequence plot comparisons (b = slope and a = intercept). Flux units are g C m^{-2} y^{-1}, mass units are g C m^{-2}. See text for NPP abbreviations.*

Variable	*a*	*SE*	*b*	*SE*	R^2	*RMSE*
Extensive Plots (N=75)						
NPP_{Aw}	90	26	0.80	0.08	0.59	116
NPP_A	59	32	0.93	0.06	0.75	121
LM_{Aw}	3770	863	0.57	0.04	0.74	8950
Chronosequence Plots (N=36)						
NPP_{Aw}	62	31	0.81	0.09	0.69	100
NPP_A	36	37	0.93	0.08	0.82	103
NPP	20	79	1.10	0.11	0.73	207
LM_{Aw}	4240	1210	0.54	0.05	0.76	9630
NEP	99	38	0.55	0.12	0.37	183

Table 9.6. *Effects of alternative parameterization schemes on regression statistics for comparison of observed and modeled aboveground net primary production (NPP_A; b = slope and a = intercept). In the default case, all sites were run with the same set of ecophysiological constants. In the ecoregion case, there was a unique parameterization of the ecophysiological constants for each ecoregion, and in the site-specific case, there was a unique parameterization for each site.*

Parameterization	*a*	*SE*	*b*	*SE*	R^2	*RMSE*
Default	107	36	0.92	0.07	0.70	148
Ecoregion	49	33	0.94	0.06	0.75	121
Site Specific	63	41	0.89	0.08	0.63	150

To reveal the benefits of using the ecoregion-specific observations of foliar N concentration, SLA, and leaf retention, a comparison of observed and modeled NPP_A was also made for a model run using a generic conifer parameterization. Without the ecoregion parameterization, the model produced significant additive bias, and the RMSE was considerably higher (Table 9.6). The use of even more specialized (site-specific) values of foliar N, SLA, and leaf retention also showed improved fit over the generic conifer parameterization, but the RMSE was just as high due to large model error at a few sites. For each parameterization scheme, regressions of model-error against climate indices such as annual precipitation and summer precipitation were weak at best, except for the site-specific parameterization scheme. Possibly, other parameters linked to foliar C:N and SLA must be changed in parallel to achieve consistent improvements. Further analyses of these model errors in relation to climatic gradients may be helpful for diagnostic purposes.

9.6.9 Comparisons of Simulations with Chronosequence Data

At the chronosequence plots, comparisons were made between field-based and modeled NPP_{Aw}, NPP_A, total NPP, stem mass, and NEP (Table 9.5, Figure 9.7). The agreement between model and observations was best for NPP_A and NPP. For other variables the correlations were positive (with regression slopes >0.5), but further work on model development is needed. Future work will focus on the mortality parameter and on components of heterotrophic respiration, which strongly influence modeled NEP.

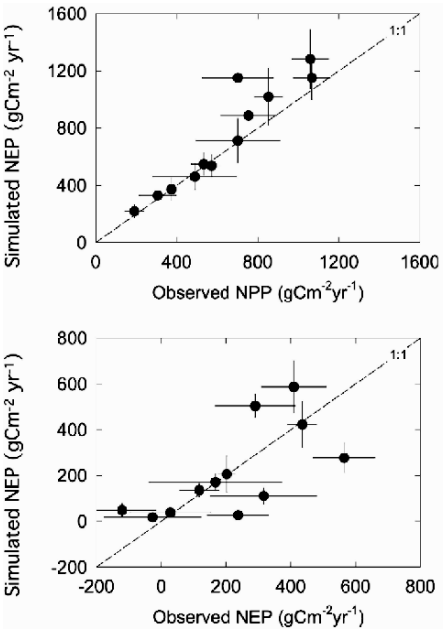


Figure 9.7. *Simulated and observed net primary production (NPP) and net ecosystem production (NEP) for the chronosequence study plots. Values are means and standard deviation of 3 replicate plots per site × age class combination.*

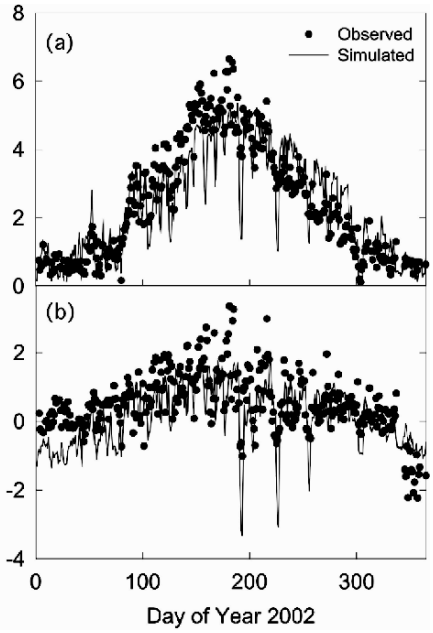


Figure 9.8. *Comparison of flux tower observations and model simulations of gross primary production (GPP) and net ecosystem production (NEP) at the Metolius young ponderosa pine site in 2002.*

9.6.10 Comparisons of Simulations with Flux Tower Data

Flux tower estimates of evapotranspiration were initially used to evaluate the generic parameterization derived from White et al. (2000). In examining observed and modeled evapotranspiration, it became evident that the default maximum stomatal conductance parameter was too high, and it was therefore reduced significantly. For the year 2001 comparisons, the time series plots of daily GPP at the young pine tower site suggested a slight underestimation of GPP in mid-growing

season (Figure 9.8). The model also consistently underestimated NEP throughout most of the growing season. There remain significant uncertainties in the nighttime NEE estimates from the tower data and the modeled NEE is the small net of the large GPP and heterotrophic respiration (errors in both). Thus, the comparisons must be considered tentative. Comparisons will be made with results of continuing efforts to quantify heterotrophic respiration fluxes in the field (Law et al. 2001a, 2003).

9.6.11 NEP Surfaces

After completion of model testing and parameterization based on the complete suite of observational data, the model will be run wall-to-wall over the east-west swath in western Oregon. Preliminary test areas in the three ecoregions (e.g., Figure 9.9) show the effects of management and environmental gradients. NEP is relatively low in areas recently clearcut for harvest, and highest in young stands (age 30-100) that have a closed canopy and have lost most residues from their stand-originating disturbance. NEP on the drier east side of the Cascades tends to be relatively low. Besides analysis of within region heterogeneity in NEP, the remote sensing/ modeling scaling approach will also permit analysis of interannual variation in NEP (Turner et al. 2003).

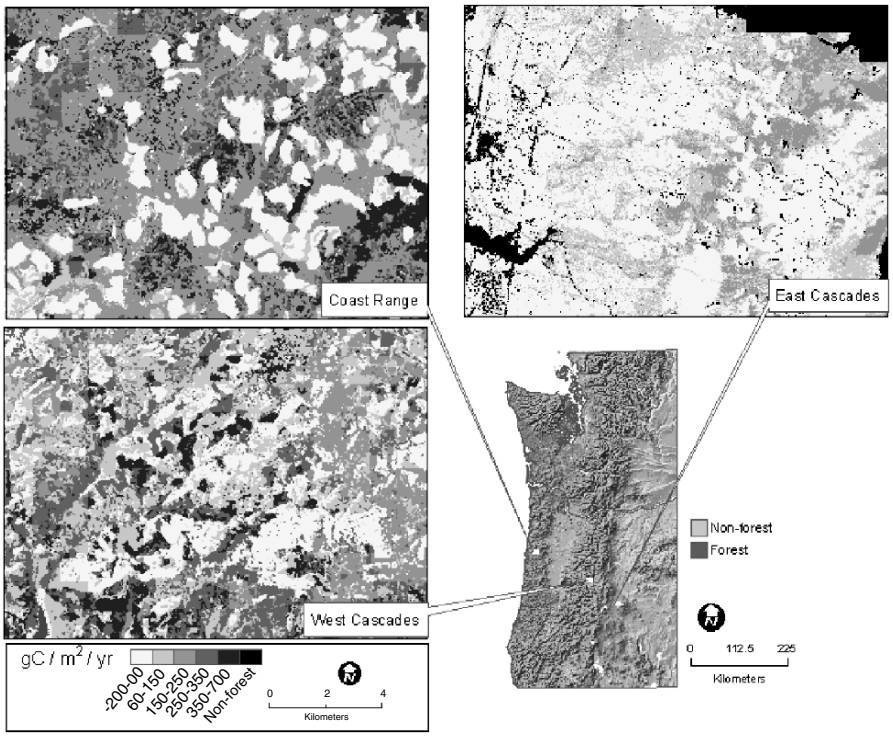


Figure 9.9. *Simulated net ecosystem production (NEP) for selected areas in the study region.*

9.7 CONCLUSIONS

Scaling in space and time is essential if we are to address global change issues relevant to terrestrial carbon cycling. We demonstrated a scaling strategy that uses a spatially nested hierarchy of observations for model parameterization and testing, along with a simulation model that provides a means of integrating environmental information at a range of spatial and temporal scales. Observations that include flux towers, intensive field sites, and inventory data increase understanding of ecological processes and permit the iterative process of model testing and improvement, recognizing there are uncertainties in field observations as well as model estimates. Uncertainty estimates for field observations, that are aggregates of multiple measurements, can be calculated by determining upper and lower boundaries for each component measurement and propagating these estimates through to a single variable (NEP in this case) with Monte Carlo models. Field observations over a range of environmental conditions are necessary for accuracy assessment of remote sensing estimates of vegetation characteristics. Knowledge of model sensitivities to key parameters helps to determine measurements that should be made at inventory sites (e.g., wood increment and density), extensive sites (e.g., foliar and soil carbon and nitrogen), and intensive sites (e.g., A-Ci curves – photosynthetic response to internal CO_2). In coniferous forests of the Pacific Northwest, disturbance history and environmental gradients are the major controls on carbon pools and fluxes. Thus, for regional analysis of carbon, nitrogen and water cycling it is critical to have spatial data layers such as remotely sensed estimates of cover type, changes in cover or disturbance, and spatially distributed climate. Advancements of the approach might include more iterative model development and testing and data assimilation techniques.

ACKNOWLEDGEMENTS

This study was funded by U.S. EPA - Science to Achieve Results (STAR) Program (Grant # R-82830901-0) and the Department of Energy (Grant # FG0300ER63014). Thanks to Peter Anthoni and Meredith Kurpius for tower flux data analysis, and to the following people for field data collection and analysis: Jesse Bablove, Jason Barker, Aaron Domingues, David Dreher, Marie Ducharme, Chris Dunham, Colin Edgar, Isaac Emery, Nathan Gehres, Angie Hofhine, Julie Horowitz, Nicole Lang, Erica Lyman-Holt, Darrin Moore, Adam Pfleeger, Lucia Reithmaier, Jennifer Sadlish, Matthew Shepherd, Nathan Strauss, and Vernon Wolf.

REFERENCES

Aber, J. D., C. Driscoll, C. A. Federer, R. Lathrop, G. Lovett, J. M. Melillo, P. Steudler, and J. Vogelmann. 1993. A strategy for the regional analysis of the effects of physical and chemical climate change on biogeochemical cycles in northeastern (U.S.) forests. Ecological Modeling 67: 37-47.

Anthoni, P. M., B. E. Law, and M. H. Unsworth. 1999. Carbon and water vapor exchange of an open-canopied ponderosa pine ecosystem. Agricultural and Forest Meteorology 95:151-168.

Baldocchi, D. D. 2003. Assessing ecosystem carbon balance: problems and prospects of the eddy covariance technique. Global Change Biology 9:479-492.

Cescatti, A. 1997. Modelling the radiative transfer in discontinuous canopies of asymmetric crowns. II. Model testing and application in a Norway spruce stand. Ecological Modelling 101:275-284.

Cheney, W., and D. Kincaid. 1985. Numerical Mathematics and Computing. Brooks/Cole, Monterey CA.

Cohen, W. B., M. E. Harmon, D. O. Wallin, and M. Fiorella. 1996. Two decades of carbon flux from forests of the Pacific Northwest. BioScience 46:836-844.

Cohen, W. B., T. K. Maiersperger, T. A. Spies, and D. R. Oetter. 2001. Modeling forest cover attributes as continuous variables in a regional context with Thematic Mapper data. International Journal of Remote Sensing 22:2279-2310.

Cohen W. B., T. A. Spies, R. J. Alig, D. R. Oetter, T. K. Maiersperger, and M. Fiorella. 2002. Characterizing 23 years (1972-1995) of stand replacement disturbance in western Oregon forests with Landsat imagery. Ecosystems 5:122-137.

Cramer, W., A. Bondeau, F. I. Woodward, I. C. Prentice, R. A. Betts, V. Brovkin, P. M. Cox, V. Fisher, J. A. Foley, A. D. Friend, C. Kucharik, M. R. Lomas, N. Ramankutty, S. Sitch, B. Smith, A. White, and C. Young-Molling. 2001. Global response of terrestrial ecosystem structure and function to CO_2 and climate change results from six dynamic global vegetation models. Global Change Biology 7:357-373.

Ehleringer, J. R., and C. Field, editors. 1993. Scaling Physiological Processes: Leaf to Globe. Academic Press, Orlando, FL.

Franklin, J. F., and C. T. Dyrness. 1973. Natural Vegetation of Oregon and Washington. Pacific Northwest Forest and Range Experiment Station. USDA Forest Service General Technical Report PNW-8.

Garman, S. L., S. A. Acker, J. L. Ohmann, and T. A. Spies. 1995. Asymptotic Height-diameter Equations for Twenty-four Tree Species in Western Oregon. Research Contribution 10. Forest Research Laboratory, Oregon State University, Corvallis, OR.

Grier, C. C., and S. R. Running. 1977. Leaf area of mature northwestern coniferous forests: relation to site water balance. Ecology 58:893-899.

Hurlbert, S. H. 1984. Pseudoreplication and the design of ecological field experiments. Ecological Monographs 54:187-211.

Jarvis, P. G. 1995. Scaling processes and problems. Plant, Cell and Environment 18:1079-1089.

Kern, J. S., D. P. Turner, and R. F. Dodson. 1997. Spatial patterns in soil organic carbon pool size in the Northwestern United States. Pages 29-43 *in* R. Lal, J. M. Kimbal, R. Follett, and B. A. Stewart, editors. Soil Processes and the Carbon Cycle. CRC Press, Boca Raton, FL.

Law, B. E., O. J. Sun, J. Campbell, S. Van Tuyl, and P. Thornton. 2003. Changes in carbon storage and fluxes in a chronosequence of ponderosa pine. Global Change Biology 9:510-524.

Law, B. E., P. Thornton, J. Irvine, S. Van Tuyl, and P. M. Anthoni. 2001a. Carbon storage and fluxes in ponderosa pine forests at different developmental stages. Global Change Biology 7:755-777.

Law, B. E., S. Van Tuyl, A. Cescatti, and D. D. Baldocchi. 2001b. Estimation of leaf area index in open-canopy ponderosa pine forests at different successional stages and management regimes in Oregon. Agricultural and Forest Meteorology 108:1-14.

Maeglin, R. R., and H. E. Wahlgren. 1972. Western Wood Density Survey: Report Number 2. USDA Forest Service Research Paper FPL-183.

Melillo, J. M., A. D. McGuire, D. W. Kicklighter, B. Moore, C. J. Vorosmarty, and A. L. Schloss. 1993. Global climate change and terrestrial net primary production. Nature 363:234-240.

Reich, P. B., M. B. Walters, and D. S. Ellsworth. 1997. From tropics to tundra: global convergence in plant functioning. Proceedings of the National Academy of Science 94:13730-13734.

Running, S. W. 1994. Testing FOREST-BGC ecosystem process simulations across a climatic gradient in Oregon. Ecological Applications 4:238-247.

Thornton, P. E. 1998. Regional ecosystem simulation: combining surface- and satellite-based observations to study linkages between terrestrial energy and mass budgets. Ph.D. Dissertation. School of Forestry, University of Montana, Missoula, MT.

Thornton, P. E., and S. W. Running. 1999. An improved algorithm for estimating incident daily solar radiation from measurements of temperature, humidity, and precipitation. Agricultural and Forest Meteorology 93:211-228.

Thornton, P. E., H. Hasenauer, and M. A. White. 2000. Simultaneous estimation of daily solar radiation and humidity from observed temperature and precipitation: an application over complex terrain in Austria. Agricultural and Forest Meteorology 104:255-271.

Thornton, P., B. E. Law, D. A. Ellsworth, H. Gholz, A. Goldstein, D. Hollinger, and K. T. Paw U. 2002. Modeling the effects of disturbance history and climate on carbon and water budgets in evergreen needleleaf forests. Agricultural and Forest Meteorology 113:185-222.

Thornton, P. E., S. W. Running, and M. A. White. 1997. Generating surfaces of daily meteorological variables over large regions of complex terrain. Journal of Hydrology 190:214-251.

Turner, D. P., W. B. Cohen, and R. E. Kennedy. 2000. Alternative spatial resolutions and estimation of carbon flux over a managed forest landscape in western Oregon. Landscape. Ecology 15:441-452.

Turner, D. P., M. Guzy, M. Lefsky, S. Van Tuyl, O. Sun, C. Daly, and B. E. Law. 2003. Effects of land use and fine scale environmental heterogeneity on net ecosystem production over a temperate coniferous forest landscape. Tellus 55B:657-668.

U.S. Department of Agriculture, Forest Service. 2001. Forest Inventory And Analysis National Core Field Guide, Volume 1: Field Data Collection Procedures for Phase 2 Plots, Version 1.5. Internal Report U.S. Department of Agriculture, Forest Service, Forest Inventory and Analysis, Washington, D.C.

Van Tuyl, S., B. E. Law, D. Turner, and A. Gitelman. 2005. Variability in net primary production and carbon storage in biomass across Oregon forests – an assessment integrating data from forest inventories, intensive sites, and remote sensing. Forest Ecology and Management 209:273-291.

Vitousek, P. M., J. D. Aber, R. W. Howarth, G. E. Likens, P. A. Matson, D. W. Schindler, W. H. Schlesinger, and D. Tilman. 1997. Human alteration of the blobal nitrogen cycle: sources and consequences. Ecological Applications 7:737-750.

Vogelmann, J., T. Sohl, and S. Howard. 1998. Regional characterization of land cover using multiple sources of data. Photogrammetric Engineering and Remote Sensing 64:45-57.

White, J. D., P. Thornton, S. W. Running, and R. Nemani. 2000. Parameterization and Sensitivity Analysis of the Biome-BGC Terrestrial Ecosystem Model: Net Primary Production Controls. Earth Interaction 4, paper No. 3.

White, M. A., and S. W. Running. 1994. Testing scale dependent assumptions in regional ecosystem simulations. Journal of Vegetation Science 5:687-702.

Wilson, K., D. D. Baldocchi, E. Falge, M. Aubinet, P. Berbigier, C. Bernhofer, H. Dolman, C. Field, A. Goldstein, A. Granier, D. Hollinger, G. Katul, B. E. Law, T. Meyers, J. Moncrieff, R. Monson, J. Tenhunen, R. Valentini, S. Verma, and S. Wofsy. 2002. The diurnal centroid of ecosystem energy and carbon fluxes at FLUXNET sites. Agricultural and Forest Meteorology 113:223-243.

Wu, J. 1999. Hierarch and scaling: extrapolation information along a scaling ladder. Canadian Journal of Remote Sensing 25:367-380.

CHAPTER 10

LANDSCAPE AND REGIONAL SCALE STUDIES OF NITROGEN GAS FLUXES

PETER M. GROFFMAN, RODNEY T. VENTEREA, LOUIS V. VERCHOT, AND CHRISTOPHER S. POTTER

10.1 INTRODUCTION

Nitrogen (N) gas fluxes have great relevance to soil fertility, water quality and air quality. Analysis of these fluxes presents several conceptual and practical scaling challenges because they are mediated by microorganisms at the scale of microns and seconds but have relevance at relatively large spatial (meters to kilometers and larger) and temporal (years, decades) scales. In this chapter, we evaluate three scaling issues that arose as part of an analysis of the effects of N deposition on gaseous N loss from temperate forest ecosystems in the northeastern U.S.

How does this chapter fit into the context of this book and the topic of "scaling and uncertainty analysis in ecology?" It occurs to us that there are three main groups of scientists grappling with scaling issues: (1) those with an inherent conceptual interest in scaling, (2) those interested in micro-scale processes (e.g., N gas fluxes) that are relevant at large scales and (3) those interested in solving large scale problems (e.g., nitrate delivery to coastal waters) that are regulated by micro-scale processes. We fall solidly in the second group, researchers who have been struggling to measure N gas fluxes at micro-scales being asked to evaluate the importance of our results to large-scale problems such as the fate of atmospheric N deposition or nitrate delivery to coastal waters (these knotty problems are defined below). Given that N gas fluxes are miserable to measure at micro-scales (lousy methods, absurd variability over small spatial and temporal scales), we, and most other micro-scale researchers, are uncomfortable scaling our miserable data to larger scales. That is, if you take a bad number measured at a small scale and extrapolate it to a very large scale, do you end up with a "very bad number" or a "big bad number" or what? So, be warned gentle reader, that landscape and regional scale studies of N gas fluxes are "not for the squeamish." But, given the difficulty of our challenge, we are pleased to contribute to a book that includes representatives from all three scaling

J. Wu, K.B. Jones, H. Li, and O.L. Loucks (eds.),
Scaling and Uncertainty Analysis in Ecology: Methods and Applications, 191–203.

motivation groups. In our view, exchange of ideas and challenges among these groups is the key to making progress in this critically important area of environmental science.

Our case study addresses three distinct scaling challenges: (1) how to account for landscape scale variability in regional studies – an experimental design issue, (2) how to account for the episodic nature of gas flux – a temporal scaling issue, and (3) how to validate landscape and regional scale flux estimates – an uncertainty and validation issue. The experimental design issues that we address are also discussed in the chapters by Wu and Li (Chapters 1 and 2), Bradford and Reynolds (Chapter 6), and Peters et al. (Chapter 7). Li and Wu (Chapter 3) provide a relevant discussion of uncertainty and error analysis that is relevant to our third challenge.

10.2 N GAS FLUXES AND ATMOSPHERIC DEPOSITION – SOME BACKGROUND

Soil-atmosphere N gas fluxes are the most poorly characterized component of the terrestrial N cycle (Mooney et al. 1987, Schlesinger 1997). There are three gases that are produced; nitric oxide (NO), nitrous oxide (N_2O) and dinitrogen (N_2), as a by-product of multiple N transformations that occur in soil (Firestone and Davidson 1989). The most important transformations that lead to gas flux are nitrification, an aerobic process, and denitrification, an anaerobic process. Given that these processes have complex regulating factors and high variability in time and space, N gas fluxes often exhibit extreme variation (Foloronuso and Rolston 1984, Parkin 1987, Robertson et al. 1988). Moreover, it is difficult to measure gas fluxes without disturbing the physical soil environment and/or the biological transformations that produce the fluxes, leading to frequent concerns that observed results are artifacts of a particular method (Groffman et al. 1999).

Fluxes of N gases influence several ecosystem (10 m), landscape (100m), regional (>100 km) and global scale processes. At the ecosystem scale, N gas fluxes can deplete soil stocks of inorganic N, an essential, and frequently limiting (to plant growth) nutrient (Vitousek and Howarth 1991). At the landscape scale, these fluxes can prevent or mitigate the movement of excess inorganic N from terrestrial environments (e.g., highly fertilized agricultural fields) into water bodies where they can cause overgrowth of aquatic plants and eutrophication (Lowrance 1998). At regional and global scales, N_2O is a "greenhouse" gas that can influence the earth's radiative budget and plays a role in stratospheric ozone destruction (Prather et al. 1995). Nitric oxide is a highly reactive gas that is a precursor to tropospheric ozone formation and is readily converted to reactive N and deposited back to the earth's surface in precipitation (NRC 1992).

In addition to the greenhouse effect, an additional regional scale phenomenon affected by N gas fluxes is atmospheric deposition. Human activities have greatly increased the global production of reactive nitrogen through fertilizer use and fossil fuel combustion, leading to enrichment of the atmosphere and increased rates of reactive N deposition to the earth's surface (Vitousek et al. 1997). There is concern that enriched deposition can create a series of adverse consequences in the environment, resulting in N "saturation" or an N "cascade" affecting forests,

groundwater and freshwater and coastal aquatic ecosystems (Aber et al. 1989, Galloway et al. 2003). One of these effects may be to enhance N gas fluxes, contributing to the greenhouse effect (N_2O) or increasing ozone levels and N deposition (NO).

The fate of N deposition in terrestrial ecosystems is one of the greatest current mysteries in environmental science. Many studies have found that a very high percentage (>90%) of the N deposited on terrestrial ecosystems is retained, i.e., not exported from the ecosystem via hydrologic pathways (Boyer et al. 2002, van Breemen et al. 2002). The specifics of this vast retention, which consists of N storage in soils and vegetation and gaseous losses, are not well characterized. Quantifying the contribution of N gas fluxes to this retention is of great interest because this N is removed from the ecosystem while N stored in soils and plants remains available for cycling within, and export from, the ecosystem. There is also interest in determining if N gas fluxes could be a sensitive indicator, or early warning symptom of the onset of N saturation.

Previous studies have suggested that N gas fluxes in northeastern forest soils are low (Bowden 1986, Bowden et al. 1991), but only a small number of sites have been studied. Moreover, these studies have, for the most part, only measured N_2O. Fluxes of NO have recently been shown to be much higher than those of N_2O in forest plots receiving long-term experimental N additions (Venterea et al. 2003a). Fluxes of N_2 are basically unknown, but could be an important component of ecosystem retention of atmospheric deposition, with no negative environmental impact.

10.3 A REGIONAL SCALE STUDY OF THE IMPORTANCE OF N GAS FLUXES TO THE FATE OF ATMOSPHERIC N DEPOSITION

We received funding from the US Environmental Protection Agency's Science To Achieve Results (STAR) program on Regional Scale Analysis and Assessment to investigate the "effects of N deposition on gaseous N loss from temperate forest ecosystems." Our project has four objectives: (1) to determine the importance of gaseous loss of N from temperate forest ecosystems, (2) to determine the impacts of N deposition on gaseous loss of N from these ecosystems, (3) to test a mechanistic model that relates N gas emissions to N availability and soil moisture content, and (4) to develop a new and more mechanistic version of the daily NASA-CASA ecosystem model for N gas emissions that can be applied at the regional level using satellite remote sensing and other spatial data sets in a geographic information system (GIS) format. This new simulation model will be used to assess trends in N cycling over gradients of N deposition in the northeast US and to project changes in N gas fluxes with changing air pollution.

The project takes advantage of an N deposition gradient in the northeastern US that runs from West Virginia (high deposition - ~12 kg N ha^{-1} y^{-1}) north and east to Maine (low deposition - ~5 kg N ha^{-1} y^{-1}). Our approach was to make monthly *in situ* measurements of gas fluxes using chamber (~0.10 m^2) methods (Venterea et al. 2003a) along with measurements of ancillary N cycle processes at five sites along the gradient: Fernow Experimental Forest, WV, Catskill Preserve, NY, Harvard Forest, MA, Hubbard Brook, NH, and Bear Brook, ME. We then used the

data to modify existing process (hole in the pipe) and ecosystem (NASA-CASA) scale models and then use regional data sets to run the models at the regional scale.

The project presents three distinct scaling challenges: (1) how to account for landscape scale variability at each point along the regional gradient, i.e., how to determine the "representative flux" at each site, (2) how to account for the episodic nature of gas flux, e.g., bursts of flux in response to rainfall events, given a monthly sampling program, and (3) how to validate the landscape and regional scale estimates of flux that we produce given that we have no large-area, independent way to measure flux at landscape and regional scales. In the sections below we discuss how we have addressed each of these scaling challenges.

10.3.1 Challenge #1 – How To Account for Landscape Variability along a Regional Gradient

This challenge arises from the fact that although our five sites are aligned along a marked N deposition gradient, there is considerable natural variability in N dynamics and gas flux at each site due to variation in soils, vegetation, geology, elevation, aspect, and land use history. To establish a "representative flux" at each of our five sites, it was necessary to account for these factors in the selection of our monthly sampling locations.

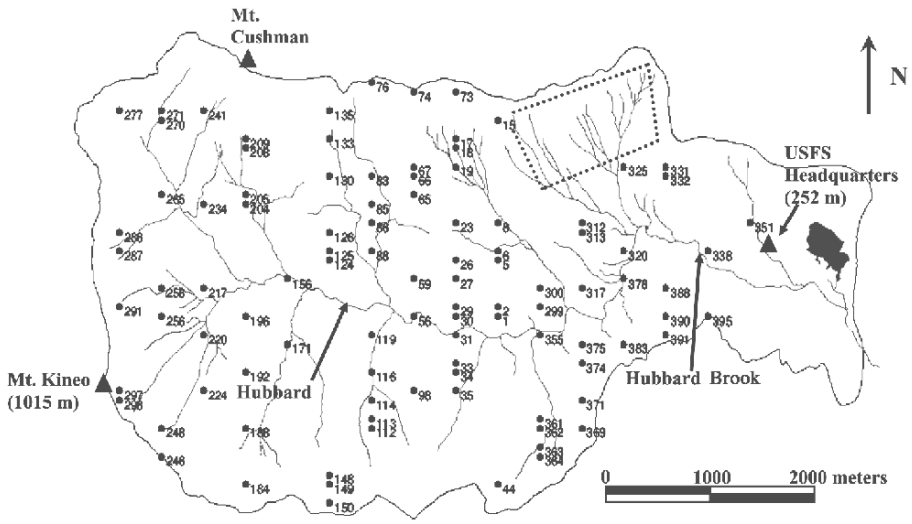


Figure 10.1. *The Hubbard Brook Experimental Forest, showing 100 plot locations utilized in the "valley-wide" study. Numbers are plot designations previously established by Schwarz et al. (2003). Dashed lines are approximate boundaries of areas used in previous watershed-scale studies (Likens and Bormann 1995). From Venterea et al. (2003b).*

At the Hubbard Brook Experimental Forest in New Hampshire, we designated the Hubbard Brook valley, a 3,160-ha catchment within the White Mountain

National Forest of central New Hampshire, USA (43° 56' N, 71° 45' W), as the representative landscape unit for this location along our regional N deposition gradient. While the HBEF has been the site of numerous watershed/ecosystem scale studies (Likens and Bormann 1995), the vast majority of these studies have taken place in a series of small watersheds in the northeast corner of the Hubbard Brook valley. For this study, we participated in a multi-investigator effort to characterize variation in ecosystem properties across the entire valley. To accomplish this, we sampled 100 randomly selected plots, a subset of 400 plots that had been established for an earlier valley-wide vegetation study (Figure 10.1, Schwarz et al. 2003), incubated samples in the laboratory, and measured potential net N mineralization and nitrification and N_2O production (Venterea et al. 2003b). We then examined relationships between these N cycle variables, which have been shown to be strongly related to N gas fluxes (Davidson et al. 2002, Venterea et al. 2003a) and landscape parameters (elevation, aspect, dominant tree species).

Our hypothesis was that there were not going to be strong landscape scale patterns in N cycling at Hubbard Brook. The forest is relatively uniform northern hardwood forest composed of yellow birch (*Betula alleghaniensis* Britton), sugar maple (*Acer saccharum* Marsh.), red spruce (*Picea rubens* Sarg.) and American beech (*Fagus grandifolia* Ehrh.) with small amounts of paper birch (*Betula papyrifera* Marsh.), balsam fir (*Abies balsamea*), red maple (*Acer rubrum*), Eastern hemlock (*Tsuga canadensis*), white ash (*Fraxinus americana*), and striped maple (*Acer pensylvanicum* L.) (Schwarz et al. 2003). Soils are acidic (pH 3.5 - 5.5) and consist of well-drained, Typic Haplorthods of sandy loam texture derived from glacial till (USDA 1996). The implications of this hypothesis for our regional project were profound – if sustained, it meant that we could sample anywhere (or at least randomly), e.g., in a nice flat site close to the road, rather than have to establish sites "all over the damn place . . ."

Somewhat surprisingly (given the relative uniformity of soils and vegetation described above), we observed strong, coherent landscape patterns of N cycling across the landscape of the Hubbard Brook valley. All process rates were higher on south facing than on north facing slopes, and at high elevation than at low elevation (Figure 10.2). These patterns were driven by the effects of aspect and elevation on soil moisture and the distribution of vegetation. The results are consistent with many other studies that have found strong patterns in N cycling with elevation and aspect (Schimel et al. 1985, Burke 1989, Groffman and Tiedje 1989, Bohlen et al. 2001).

The nitrification and N_2O fluxes were higher than we expected, with nitrification representing over 50% of net mineralization and N_2O flux representing more than 0.70% of net nitrification. These results suggest that N cycling at Hubbard Brook is relatively dynamic, and that the potential for N gas fluxes is relatively high. These results were surprising given that this site is towards the low end of our regional N deposition gradient and that streamwater nitrate losses at this site are low (Aber et al. 2002).

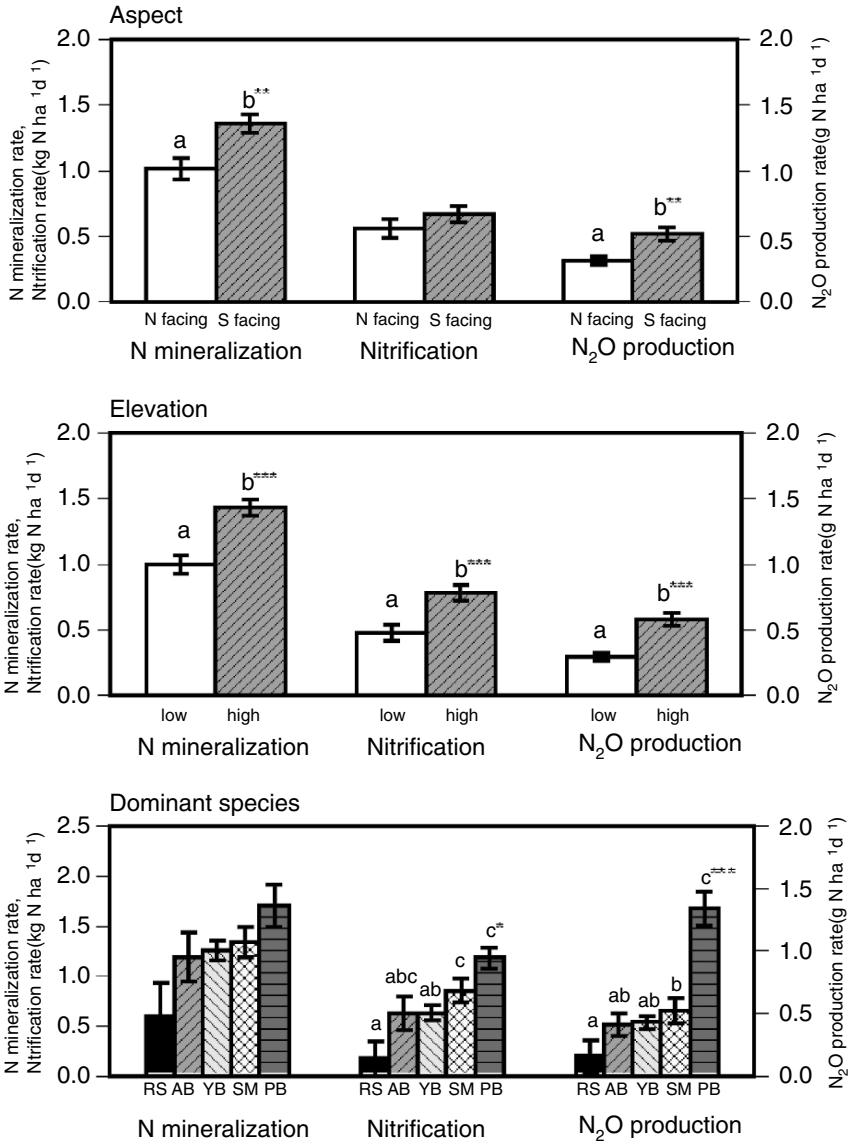


Figure 10.2. *Landscape patterns in N cycle processes at the Hubbard Brook Experimental Forest: Top – Aspect, Middle – Elevation (< or > 600 m), and Bottom – Vegetation type (> 50% basal area). Mineralization and nitrification data from Venterea et al. (2003b). RS = red spruce, AB = American beech, YB = yellow birch, SM = sugar maple, and PB = paper birch.*

The results from our "valley-wide" study were used as a basis for picking sites for our monthly measurements of *in situ* gas fluxes. They increased our confidence

that we were measuring representative fluxes at this point along our regional N deposition gradient because we had evidence that our design encompassed the major landscape-scale factors influencing flux at this location. This confidence depends on the idea that the laboratory-based potential fluxes that we measured are indicative of field fluxes, which is well supported conceptually and practically (Davidson et al. 2002, Venterea et al. 2003a), and that we have not missed any key landscape-scale factors influencing flux at this point along the gradient e.g., dead moose carcasses may be hotspots of N gas flux. The intensive sampling that we did (100 plots) suggests that this is not the case, but this is impossible to verify (see discussion of our third scaling challenge below; also see Wu and Li, Chapters 1 and 2, Bradford and Reynolds, Chapter 6, Peters et al., Chapter 7, and Wagenet 1998 for discussions of landscape scale experimental designs). We were also left to grapple with questions about the number of sites and measurements necessary to produce well-constrained estimates of flux. Unfortunately, few of our sites for monthly flux measurements were able to be located in flat locations near the road!

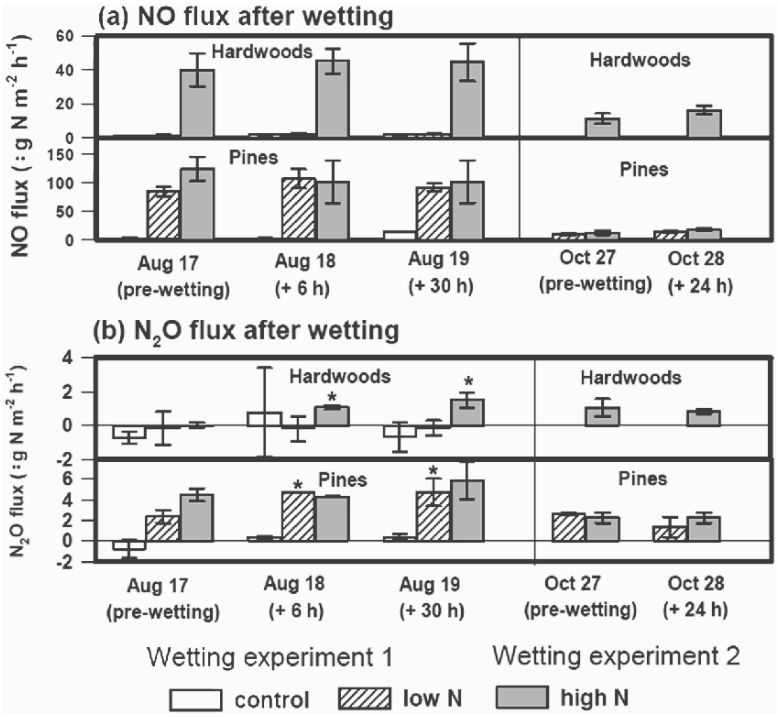


Figure 10.3. *Response of (a) NO flux and (b) N_2O flux to 25-mm of water added on August 17 and October 27 of 2001 to control, low N (50 kg N ha^{-1} y^{-1}) and high N (150 kg N ha^{-1} y^{-1}) plots at the Harvard Forest, MA. Asterisks indicate if post-wetting fluxes are significantly different from pre-wetting fluxes at $p < 0.05$. From Venterea et al. (2003a).*

It is important to note that Hubbard Brook was only one of the five points along our regional N deposition gradient. At the other points on the gradient, analyses of landscape-scale controls on flux and site selection were based on previous work at the sites by other investigators. Our regional study was greatly aided by the fact that numerous N cycling studies have been carried out at these sites.

10.3.2 Challenge #2 – How to Account for the Episodic Nature of N Gas Fluxes?

N gas fluxes are notoriously episodic, with short bursts of production occurring following rainfall or thawing events accounting for a high percentage of annual flux (Groffman et al. 2000). The best way to produce accurate evaluations of flux is to make continuous (greater than daily) measurements (Papen and Butterbach-Bahl 1999, Groffman et al. 2000). However, continuous monitoring of flux is difficult and expensive even at one site, and it is certainly not possible at landscape and regional scales. We approached this challenge with a mixture of field campaigns to assess the episodic nature of flux and simulation modeling to accomplish temporal scaling.

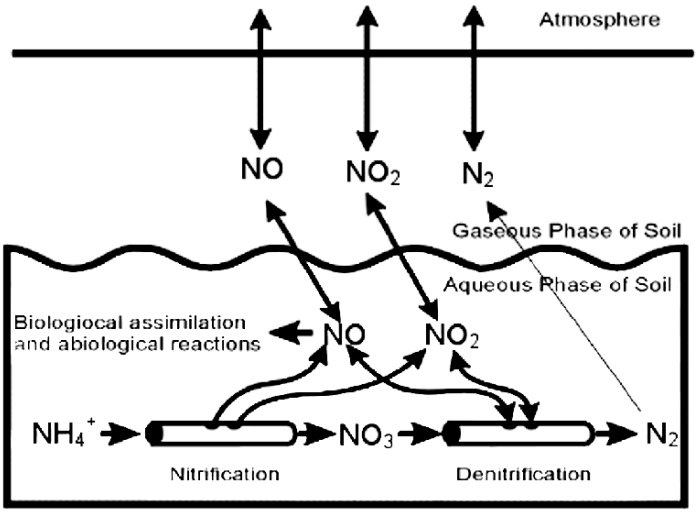


Figure 10.4. *The "hole-in-the-pipe" conceptual model of N gas fluxes. From Firestone and Davidson (1989).*

At the Harvard Forest, MA site, we assessed episodic fluxes of NO and N_2O associated with wetting events (Figure 10.3), diurnal temperature changes and N additions. Data from these assessments, combined with data from our monthly sampling allow us to parameterize flux models for these gases based on the "hole-in-the-pipe" formulation devised by Firestone and Davidson (1989). In this model, the overall rate of N transformation is depicted as the flow of water through a pipe, and N gases leak out through holes in the pipe (Figure 10.4). The overall rate of N transformation is controlled by soil and organic matter quality variables and is

represented by measurements of gross and/or net N mineralization and nitrification. The size of the holes is controlled by more transient soil conditions such as pH, temperature and water content. The hole-in-the-pipe model can be run at a daily time step and is thus capable of depicting the episodic nature of N gas fluxes. We suggest that process models, carefully calibrated with site-specific field data, are useful tools for depicting the episodic nature of N gas fluxes. Until we have technology that allows for continuous measurement of fluxes, at multiple sites, these models will continue to be important tools in landscape and regional scale studies of these fluxes.

It is interesting to note that the fluxes of NO at the Harvard Forest site were high, up to 8% of inputs, while N_2O fluxes were much lower. Previous studies at this site also found low N_2O fluxes (Bowden et al. 1991, Magill et al. 1997). Our new NO data suggest that N gas fluxes are larger, and more responsive to N deposition, than previously thought. Fluxes of NO were not responsive to wetting events, but N_2O fluxes were.

10.3.3 Challenge #3 – How to Validate Landscape and Regional Scale Estimates of Flux

In addition to serving as tools for depicting the episodic nature of N gas fluxes, i.e., for temporal extrapolation, we use models as spatial extrapolation tools to produce landscape and regional scale estimates of flux. We are in the process of linking our process models with the NASA-CASA model (Figure 10.5), which is an aggregated representation of major ecosystem C and N transformations (including gas fluxes) that can be run at regional scales when driven by a set of gridded coverages at 1-km spatial resolution (Potter et al. 1996, 1997).

The N gas emission components of the NASA-CASA model have been re-evaluated in the context of our field measurements. Revisions are underway in the CASA framework, based in part on recent validation/comparison studies (Davidson et al. 2000, Parton et al. 2001). Regional driver data sets that will be used for extrapolation include nitrogen deposition isopleth maps, daily climate drivers, soils and satellite-based estimates of leaf area index, and land cover and vegetation type. The NASA-CASA model will generate predicted nitrification rates in forest soils, which will in turn be used to predict NO and N_2O emission fluxes from soil surfaces as a function of simulated soil water content, temperature, pH, bulk density, and texture. Field measurements of these parameters at our five experimental sites will be used to make model calibration checks of the trace gas algorithms.

While the NASA-CASA model will produce landscape and regional scale estimates of flux, we have no way to directly validate these estimates because there is no way to independently assess flux at these scales. While we feel that our models are conceptually sound and empirically robust, it is possible that we have overlooked critical controlling factors at the landscape and regional scale, e.g., dead moose carcasses that may be hotspots of N gas flux in our northern sites. Given our inability to truly validate our estimates, some alternative approaches are possible. First, it is possible to predict fluxes for new sites within our region and then validate these predictions with field measurements. Second, we can evaluate our flux

estimates in the context of what is known about other fluxes at these well-studied sites. For example, if our estimates of gas flux are much higher than precipitation inputs and/or streamwater outputs of N, we will suspect that our estimates are too high. Finally, we can develop and apply other, independent modeling/extrapolation approaches to the region and see if estimates are similar. Clearly, none of these alternative validation approaches is very satisfying, but it is currently the best we can do. We can also apply standard methods of uncertainty analysis to our model results (reviewed by Li and Wu, Chapter 3), but again, these do not allow for true validation of our landscape and regional scale flux estimates.

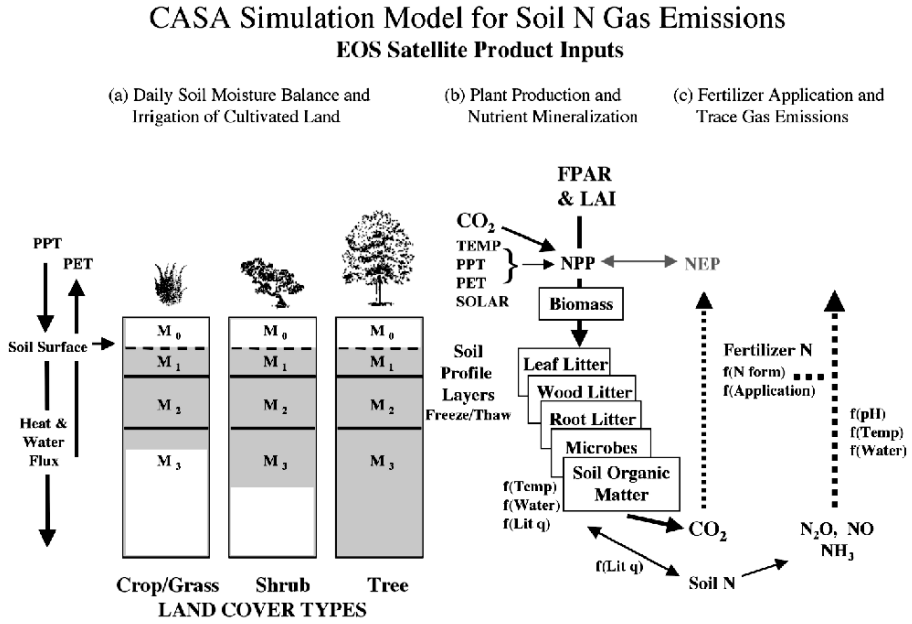


Figure 10.5. *The NASA-CASA ecosystem carbon and nitrogen process model showing how landscape and regional scale data can be used to drive ecosystem process (including N gas fluxes) models.*

10.4 CONCLUSIONS

Landscape and regional scale studies of N gas fluxes are difficult. Are the scaling challenges inherent in these studies insurmountable? One way to evaluate this is to ask if we have successfully addressed the objectives of our project. We have met our first objective (i.e., to determine the importance of gaseous loss of N from temperate forest ecosystems) with data suggesting that N gas fluxes, especially NO, are more important in northeastern forest soils than previously thought. For our second objective, we have evaluated the response of N gas fluxes to N deposition, with fertilization studies and by comparison of sites along our regional deposition

gradient (still underway). This comparison is greatly facilitated by our landscape experimental designs, which allow us to establish representative fluxes for each site along the regional gradient. These designs will also allow us to evaluate the importance of deposition as a driver of flux compared to "local factors" such as elevation, aspect and vegetation type.

Our third objective, to develop models of N gas fluxes, has also been achieved, by collection of data to parameterize flux models for our sites. It is important to note that hole-in-the-pipe type modeling is ongoing at many sites around the world, providing many opportunities for comparison and synthesis (Verchot et al. 1999, Davidson et al. 2000).

We will also achieve our fourth objective, producing regional scale estimates of N gas fluxes. However our ability to validate these estimates is indirect and incomplete and therefore our confidence in their accuracy is low. We will compare our estimates of flux with regional deposition estimates and budgets (Boyer et al. 2002, Driscoll et al. 2003) and will carry out spatial validation, i.e., prediction of flux at new sites. However, true validation will await the development of new methods (e.g., micrometeorological towers, aircraft-based measurements, new isotope approaches) that allow for independent measurement of fluxes at ecosystem, landscape and regional scales.

ACKNOWLEDGEMENTS

This work was funded by grants from the EPA (NCERQA Grant R827674), the National Science Foundation (Grant NSF DEB-9810221, Hubbard Brook LTER) and the Andrew W. Mellon Foundation. Special thanks to Jianguo Wu for organizing the workshop and book of which this chapter is a part.

REFERENCES

Aber, J. D., K. J. Nadelhoffer, P. Steudler, and J. M. Melillo. 1989. Nitrogen saturation in northern forest ecosystems. Bioscience 39:378-386.

Aber J. D., S. V. Ollinger, C. T. Driscoll, G. E. Likens, R. T. Holmes, R. J. Freuder, and C. L. Goodale. 2002. Inorganic N losses from a forested ecosystem in response to physical, chemical, biotic and climatic perturbations. Ecosystems 5:648-658.

Bohlen, P., P. Groffman, C. Driscoll, T. Fahey, and T. Siccama. 2001. Plant-soil-microbial interactions in a northern hardwood forest. Ecology 4:965-978.

Bowden, R., J. Melillo, P. Steudler, and J. Aber. 1991. Effects of nitrogen additions on annual nitrous oxide fluxes from temperate forest soils in the northeastern United States. Journal of Geophysical Research-Amospheres 96:9321-9328.

Bowden, W. B. 1986. Gaseous nitrogen emissions from undisturbed terrestrial ecosystems: an assessment of their impact on local and global nitrogen budgets. Biogeochemistry 2:249-279.

Boyer E. W., C. L. Goodale, N. A. Jaworski, and R. W. Howarth. 2002. Anthropogenic nitrogen sources and relationships to riverine nitrogen export in the northeastern USA. Biogeochemistry 57:137-169.

Burke, I. C. 1989. Control of nitrogen mineralization in a sagebrush steppe landscape. Ecology 70:1115-1126.

Davidson, E. A., M. Keller, H. E. Erickson, L. V. Verchot, and E. Veldkamp. 2000. Testing a conceptual model of soil emissions of nitrous and nitric oxides. Bioscience 50:667-680.

Driscoll, C., D. Whitall, J. Aber, E. Boyer, M. Castro, C. Cronan, C. Goodale, P. Groffman, C. Hopkinson, K. Lambert, G. Lawrence, and S. Ollinger. 2003. Nitrogen pollution in the northeastern United States: sources, effects and management options. BioScience 53:357-374.

Firestone, M. K., and E. A. Davidson. 1989. Microbial basis for N_2O and NO production and consumption. Pages 7-21 *in* M. O. Andreae and D. S. Schimel, editors. Exchange of Trace Gases Between Ecosystems and the Atmosphere. Wiley, New York.

Foloronuso, O. A., and D. E. Rolston. 1984. Spatial variability of field-measured denitrification gas fluxes and soil properties. Soil Science Society of America Journal 49:1087-1093.

Galloway, J. N., J. D. Aber, J. W. Erisman, S. P. Seitzinger, R. W. Howarth, E. B. Cowling, and B. J. Cosby. 2003. The nitrogen cascade. BioScience 53:341-356.

Groffman, P. M., and J. M. Tiedje. 1989. Denitrification in north temperate forest soils: relationships between denitrification and environmental factors at the landscape scale. Soil Biology and Biochememistry. 21:621-626.

Groffman, P. M., R. Brumme, K. Butterbach-Bahl, K. E. Dobbie, A. R. Mosier, D. Ojima, H. Papen, W. J. Parton, K. A. Smith, and C. Wagner-Riddle. 2000. Evaluating annual nitrous oxide fluxes at the ecosystem scale. Global Biogeochemical Cycles 14:1061-1070.

Groffman, P. M., E. Holland, D. D. Myrold, G. P. Robertson, and X. Zou. 1999. Denitrification. Pages 272-288 *in* G. P. Robertson, C. S. Bledsoe, D. C. Coleman, and P. Sollins, editors. Standard Soil Methods for Long Term Ecological Research. Oxford University Press, New York.

Likens, G. E., and F. H. Bormann. 1995. Biogeochemistry of a Forested Ecosystem, 2nd edition. Springer-Verlag, New York.

Lowrance, R. 1998. Riparian forest ecosystems as filters for nonpoint-source pollution. Pages 113-141 *in* M. L. Pace and P. M. Groffman, editors. Successes, Limitations and Frontiers in Ecosystem Science. Springer-Verlag, New York.

Magill, A. H., J. D. Aber, J. J. Hendricks, R. D. Bowden, J. M. Melillo, and P. A. Steudler. 1997. Biogeochemical response of forest ecosystems to simulated chronic nitrogen deposition. Ecological Applications 7:402-415.

Mooney, H. A., P. M. Vitousek, and P. A. Matson. 1987. Exchange of materials between terrestrial ecosystems and the atmosphere. Science 238:926-932.

National Research Council (NRC). 1992. Rethinking the Ozone Problem in Urban and Regional Air Pollution. National Academy Press, Washington, D.C.

Papen, H., and K. Butterbach-Bahl. 1999. A 3-year continuous record of nitrogen trace gas fluxes from untreated and limed soil of a N-saturated spruce and beech forest ecosystem in Germany: 1. N_2O emissions. Journal of Geophysical Research 104:18487-18503.

Parkin, T. B. 1987. Soil microsites as a source of denitrification variability. Soil Science Society of America Journal 51:1194-1199.

Parton, W. J., E. A. Holland, S. J. Del Grosso, M. D. Hartman, R. E. Martin, A. R. Mosier, D. S. Ojima, and D. S. Schimel. 2001. Generalized model for NO_x and N_2O emissions from soils. Journal of Geophysical Research – Atmospheres 106:17403-17419.

Potter, C. S., P. A. Matson, P. M. Vitousek, and E. A. Davidson. 1996. Process modeling of controls on nitrogen trace gas emissions from soils worldwide. Journal of Geophysical Research 101:1361-1377.

Potter, C. S., R. H. Riley, and S. A. Klooster. 1997. Simulation modeling of nitrogen trace gas emissions along an age gradient of tropical forest soils. Ecological Modelling 97:179-196.

Prather, M., R. Derwent, D. Ehhalt, P. Fraser, E. Sanhueza, and X. Zhou. 1995. Other trace gases and atmospheric chemistry. Pages 73-126 *in* J. T. Houghton, L. G. MeiroFilho, B. A. Callander, N. Harris, A. Kattenburg, and K. Maskell, editors. Climate Change. Cambridge University Press, New York.

Robertson, G. P., M. A. Huston, F. C. Evans, and J. M. Tiedje. 1988. Spatial variability in a successional plant community: patterns of nitrogen availability. Ecology 69:1517-1524.

Schimel, D. S., M. A. Stillwell, and R. G. Woodmansee. 1985. Biogeochemistry of C, N and P in a soil catena of the shortgrass steppe. Ecology 66:276-282.

Schlesinger, W. H. 1997. Biogeochemistry: An Analysis of Global Change. Academic Press, New York.

Schwarz, P. A., T. J. Fahey, and C. E. McCulloch. 2003. Factors controlling spatial variation of tree species abundance in a forested landscape. Ecology 84:1862-1878.

U.S. Department of Agriculture Forest Service. 1996. Hubbard Brook Ecosystem Study: Site Description and Research Activities. USDA Forest Service NE-INF-96-96R.

Van Breemen, N., E. W. Boyer, C. L. Goodale, N. A. Jaworski, K. Paustian, S. P. Seitzinger, K. Lajtha, B. Mayer, D. Van Dam, R. W. Howarth, K. J. Nadelhoffer, M. Eve, and G. Billen. 2002. Where did all the nitrogen go? Fate of nitrogen inputs to large watersheds in the northeastern U.S.A. Biogeochemistry 57/58:267-293.

Venterea, R. T., P. M. Groffman, L. V. Verchot, A. H. Magill, J. D. Aber, and P. A. Steudler. 2003a. Nitrogen oxide gas emissions from temperate forest soils receiving long-term nitrogen inputs. Global Change Biology 9:346-357.

Venterea, R. T., G. M. Lovett, P. M. Groffman, and P. A. Schwarz. 2003b. Landscape patterns of soil nitrate production in a northern hardwood-conifer forest. Soil Science Society of America Journal. 67:527-539.

Verchot, L., E. Davidson, J. Cattanio, I. Ackerman, H. Erickson and M. Keller. 1999. Land use change and biogeochemical controls of nitrogen oxide emissions from soils in eastern Amazonia. Global Biogeochemical Cycles 13:31-46.

Vitousek, P. M., and R. W. Howarth. 1991. Nitrogen limitation on land and in the sea: how can it occur. Biogeochemistry 13:87-115.

Vitousek, P. M., J. D. Aber, R. W. Howarth, G. E. Likens, P. A. Matson, D. W. Schindler, W. H. Schlesinger, and D. G. Tilman. 1997. Human alteration of the global nitrogen cycle: sources and consequences. Ecological Applications 7:737-750.

Wagenet, R. J. 1998. Scale issues in agroecological research chains. Nutrient Cycling in Agroecosystems 50:23-34.

CHAPTER 11

MULTISCALE RELATIONSHIPS BETWEEN LANDSCAPE CHARACTERISTICS AND NITROGEN CONCENTRATIONS IN STREAMS

K. BRUCE JONES, ANNE C. NEALE, TIMOTHY G. WADE, CHAD L. CROSS, JAMES D. WICKHAM, MALIHA S. NASH, CURTIS M. EDMONDS, KURT H. RIITTERS, ROBERT V. O'NEILL, ELIZABETH R. SMITH, AND RICK D. VAN REMORTEL

11.1 INTRODUCTION

There have been numerous papers reporting relationships between watershed and landscape characteristics and chemical, physical, and biological attributes of streams (see summary in Lee et al. 2001). Some of these studies have shown strong linkages between stream and near-site landscape conditions (Lammert et al. 1999, Stauffer et al. 2000, Talmage et al. 2002), while others show stronger linkages between stream and watershed-scale, landscape conditions (Roth et al. 1996, Allan et al. 1997, Herlihy et al. 1998, Davies et al. 2000, Jones et al. 2001a). Moreover, there are differences in the importance of environmental variables in explaining variation in stream condition among the various studies (for a comparison, see Roth et al. 1996 and Lee et al. 2001).

Differences in results are not surprising given the wide range of scales and biophysical settings, and the potential for variation in environmental conditions (e.g., precipitation and flow) among the years in which these studies were conducted. Scaling is one of the top challenges facing environmental managers and ecologists alike (Wu 1999, Wu and Hobbs 2002).

Stream chemistry, including nitrogen concentration, seems to be controlled by overall watershed conditions (Carpenter et al. 1998, Herlihy et al. 1998, Jones et al. 2001a), although point sources and atmospheric nitrogen deposition also can be significant sources of nitrogen in streams (Behrendt 1996 and Smith et al. 1997, respectively). Studies of the importance of land cover and land use in riparian zones is well documented (Lowrance et al. 1984, Peterjohn and Correll 1984), but most

J. Wu, K.B. Jones, H. Li, and O.L. Loucks (eds.),
Scaling and Uncertainty Analysis in Ecology: Methods and Applications, 205–224.

have been limited in geographic scope and biophysical variability. Moreover, little is known on how effective forested riparian buffers are in filtering nitrogen in different biophysical settings across broad regions.

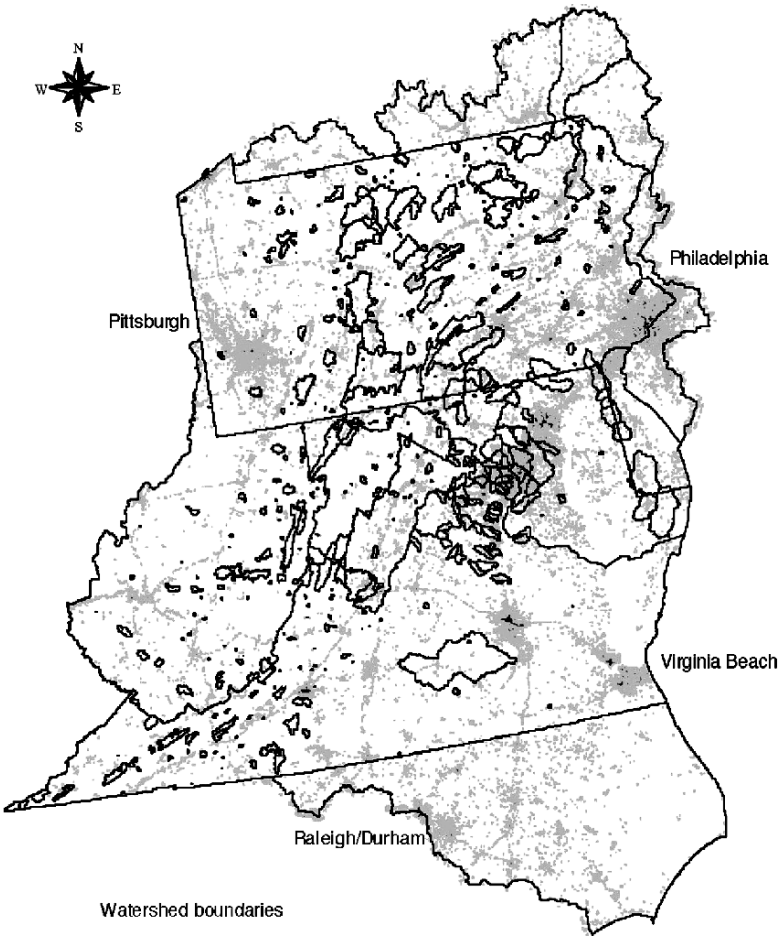


Figure 11.1. *The Mid-Atlantic study area and watersheds used in the analysis.*

Excess nutrient loading into streams, from atmospheric inputs and land surfaces, is a key issue with regards to maintaining the chemical and biological condition of streams (Carpenter et al. 1998, Jones et al. 2001a). High levels of nutrients in streams can create both risks to human health and biological condition (Ator and Ferrari 1997). Smith et al. (2001) found a strong relationship between landscape characteristics, especially the amount of urban and agriculture on greater than 3% slopes, and relatively high concentrations of fecal coliforms in South Carolina streams. They hypothesized that excess loadings of nutrients from animal and

human waste was the likely cause of this pattern. Atmospheric nitrate deposition is considered to be a significant source of nitrogen in streams, especially in the northeastern U.S (EPA 1996, Smith et al. 1997, Clark et al. 2000).

The scaling relationships of factors affecting stream nitrogen and changing importance of explanatory variables across a broad geographic region are important in determining what management actions might be effective in reducing the risk of excess nutrient loadings to streams (Hunsaker and Levine 1995). The Environmental Protection Agency (EPA), through its responsibilities relative to the Clean Air Act, has been trying to reduce atmospheric inputs of nitrogen to streams through implementation of emission standards (EPA 1996, Holland et al. 1999).

In this chapter, we analyze relationships between total stream nitrogen and explanatory variables representing three different scales – regional atmospheric nitrogen deposition, watershed level land surface characteristics, and land cover composition within the riparian zone – using a statistical approach that was used to design the NASA space shuttle. We also discuss potential sources of error in the application of the statistical approach. Finally, we discuss how changes in scaling functions might be used as an indicator of overall watershed condition.

11.2 METHODS

11.2.1 Stream Data and Watershed Delineation

Four hundred seventy-seven (477) observations of total nitrogen were obtained from the 1993 and 1994 Environmental Monitoring and Assessment Program (EMAP) stream surveys in the Mid-Atlantic region (370 total samples, EPA 2000) and from the STORET database (107 total samples) from approximately the same time period using similar collection methods. Dates of collection were similar to those for land cover acquisition (see below). Watersheds for each of the 477 samples were delineated from 30 m Digital Elevation Model (DEM) data using ArcInfo GIS watershed delineation software (ESRI 1996). Each stream sample represented the pour point for the upstream watershed or catchment area. The watersheds were widespread across the region, varied considerably in size, and represented a wide range of biophysical settings (Figure 11.1).

11.2.2 Explanatory Variables

We used a set of variables similar to those used by Jones et al. (1997, 2001a) (Table 11.1). The variables represented three spatial scales: (1) wet nitrate deposition, which had a broad regional pattern (see Jones et al. 1997), (2) watershed-level characteristics of land cover, elevation, slope, and soil characteristics, and (3) riparian zone land cover characteristics. We used an atmospheric wet nitrate deposition model (SAMAB 1996) to estimate watershed level atmospheric inputs of nitrogen. This model used the EPA wet deposition network data from 1993 (see Holland et al. 1999) to create a broad deposition grid of the region. Grid values were then modified and calculated on 180 × 180 m grid cells derived from a resample of

the DEM based on the strong correlation between wet nitrate deposition and elevation (SAMAB 1996). Atmospheric wet nitrate deposition values were then calculated for each watershed in a GIS.

Table 11.1. *Independent variables used in the regression tree analysis. See Jones et al. (1997, 2001a) for calculation of variables and implementation of soil loss and nitrate deposition models, respectively.*

Metric	*Explanation*
Agriculture on steep slopes (AGS3)	Percent of watershed with agricultural land cover occurring on greater than 3% slopes
Agricultural land cover (ALC)	Percent of watershed with agricultural land cover
Forest land cover (FLC)	Percent of watershed with forest land cover
Soil loss (POSO)	Proportion of watershed with the potential for soil losses greater than 2240 kg ha^{-1} yr^{-1}
Road Density (RD)	Average number of kilometers of roads per square km of watershed
Roads near streams (RXS)	Proportion of total stream length having roads within 30 m
Riparian agriculture (RIPA)	Percent of total watershed with agricultural land cover within 30 m of the stream
Riparian forest (RIPF)	Percent of watershed area with forest land cover within 30 m of the stream
Slope gradient (SG)	Average percent slope gradient for the watershed
Forest Fragmentation (FFRG)	Of all the pairs of adjacent pixels in the watershed that contain at least one forest pixel, the percentage for which the other pixel is not forest
Wetland land cover (WLC)	Percent of watershed with wetland land cover
Nitrate Deposition (ND)	Estimated average annual wet deposition of nitrate (kg ha^{-1} yr^{-1})

Watershed and riparian-level land cover metrics were generated in a GIS using the National Land Cover Database (NLCD, Vogelmann et al. 2001). The NLCD was generated from Landsat Thematic Mapper data (30 meter resolution) that consisted of approximately 21 land-cover classes for the Mid-Atlantic region (Vogelman et al. 2001). However, similar to Jones et al. (2001a and b), we reclassified the NLCD data into six classes: agriculture, urban, wetland, water, forest, and barren. Descriptions of the variables (landscape metrics) are given in Table 11.1 and can also be found in Jones et al. (1997), Jones et al. (1999), and Jones et al. (2001a). A metric for soil loss was generated in a GIS at the watershed scale by applying the RUSLE soil loss equation (Jones et al. 1997).

11.2.3 Regression Tree Analysis

We performed Regression Tree Analysis (RTA) using CART® software (Breiman et al. 1984, Steinberg and Colla 1995) to investigate existing patterns between the dependent variable, total nitrogen, and the set of watershed explanatory variables

listed above. Total nitrogen concentration values were log-transformed using a natural log transformation prior to any analyses.

RTA is a binary, recursive partitioning process whereby the data are consecutively split into two child nodes using a set of splitting rules. The goal of each split is to reduce the variance within each parent node whereas the overall goal of the RTA is to produce terminal nodes containing data with maximized homogeneity (i.e., low variance). Improvement is measured by how much the variance is reduced and at each split; CART® evaluates all possible splits for all variables. With our analysis, for example, CART® examined a potential 5247 splits (477 possible nitrogen values × 11 variables) to determine the most effective root node split. In the tree structure produced by RTA, "yes" answers always go to the left and "no" answers always go to the right. These child nodes then either become terminal nodes with no further splits or become parent nodes themselves. The mean value for each terminal node is then the predicted value for watersheds with the same attributes as those in the terminal node. We used the default modes in running the analyses, including the 1-SE rule to define the tree. Additionally, we evaluated the consistency of the tree against a set of 25 randomly selected trees generated from the Bagging program in CART®. We compared the sum of residuals squared and absolute deviations from the initial tree against those generated from the committee of trees to evaluate the consistency of variables selected for the tree. CART® also provides a list of surrogate and competing variables for each split in the tree, along with their potential contribution to the resolving power of the overall tree. Surrogate variables are usually correlated to the selected variables and can be almost as good in resolving a particular split in the tree as the selected variables. When surrogate variables explain nearly as much variability as the selected variables for particular breaks in the initial tree, randomly generated trees (e.g., through the Bagging program) often yield tree splits with alternating variables.

The analysis was performed with no pre-conceived idea of stratifying the sample (e.g., by ecoregion) but one of the goals of this analysis was to identify emerging patterns among groups of watersheds via the RTA and then to investigate the spatial patterns demonstrated by the terminal nodes to determine whether they shared common biophysical characteristics.

We classified total nitrogen concentrations into three condition classes based on results of stream surveys conducted by EMAP in the Mid-Atlantic Highlands (EPA 2000), and summarized conditions for each terminal node. These classes were good (<5.9 Ln N mg/L), fair (5.9 Ln N mg/L to 6.5 Ln N mg/L), and poor (>6.5 Ln N mg/L).

11.2.4 Spatial Analysis of Terminal Node Members

In order to evaluate the spatial relationship among the nodal sample points identified by RTA, a nearest-neighbor analysis of complete spatial randomness was conducted. We used the nearest-neighbor algorithm of Clark and Evans (1954) with the edge-correction methodology of Donnelly (1978) as implemented in the "Animal Movement Analysis" ArcView extension (Hooge and Eichenlaub 1997).

This ArcView extension calculates a nearest-neighbor index for a given polygon theme. In practice, this polygon theme would be a pre-defined boundary (e.g., watershed boundary, county boundary, or a defined sample-site boundary). However, inasmuch as the nodal point theme generated by the RTA analysis spans across large geographic regions, polygon boundaries were defined by connecting the outermost coordinates of each set of nodal points (the "minimum convex polygon") (Hooge and Eichenlaub 1997). The strength of this method is that each analysis of a point theme focuses on the distribution of points as they appear spatially, without imposing a political or other boundary condition.

The index of Clark and Evans (1954) is generated by calculating the ratio of the mean nearest-neighbor distance among all sample points (numerator) and the expected mean nearest-neighbor distance under the assumption of a Poisson process (denominator). Under complete spatial randomness, this ratio will tend toward unity, whereas a ratio less than one is indicative of a clustered pattern and ratio greater than one is indicative of a uniform pattern (Hooge and Eichenlaub 1997).

11.2.5 Ecoregion Characterization of Watersheds

Since existing evidence suggests a relationship between stream condition and ecoregions (Hughes et al. 1987, Herlihy et al. 1998, Griffith et al. 1999), we characterized ecoregion composition for each RTA terminal node. Unlike Herlihy et al. (1998) who assigned an ecoregion designation based on the stream sampling location, we calculated the percentage of the watershed surface in different ecoregions. We used a digital coverage of Omernik Ecoregions (Omernik 1987) to determine the proportion of each ecoregion on each of the 477 watersheds.

11.3 RESULTS

11.3.1 Tree Structure

The RTA analysis yielded a tree structure with eight terminal nodes and an overall strength (1-relative resubstitution error) of 0.615 (Figure 11.2). The sum of residuals squared and absolute deviations from the initial tree (30.99 and 29.03, respectively) were similar to those generated from the committee of 25 randomly generated trees (29.73 and 26.80, respectively).

The most important variable (improvement of 0.45) and first split on the tree was the percentage of forest at the watershed scale. Three hundred watersheds were split into a group that had greater than 68.3% forest (right side of the tree) and 177 were in a group with less than or equal to 68.3% forest (left side of the tree, Figure 11.2). Three of the four terminal nodes on the right side of the tree had streams with average total nitrogen values in the good class (range 5.1 to 6.4 Ln N mg/L), whereas three of the four terminal nodes on the left side of the tree had poor average values (range 5.9 to 8.1 Ln N mg/L, Figure 11.2). The percentage of watershed area in agricultural land cover and forest fragmentation were surrogate values (association >0.75) for forest land cover, but were not selected in the model because

the amount of forest provided the greatest improvement in the model at this level in the tree. Moreover, the percentage of forest at the watershed scale was the first splitting variable in all 25 randomly generated trees in the Bagging analysis.

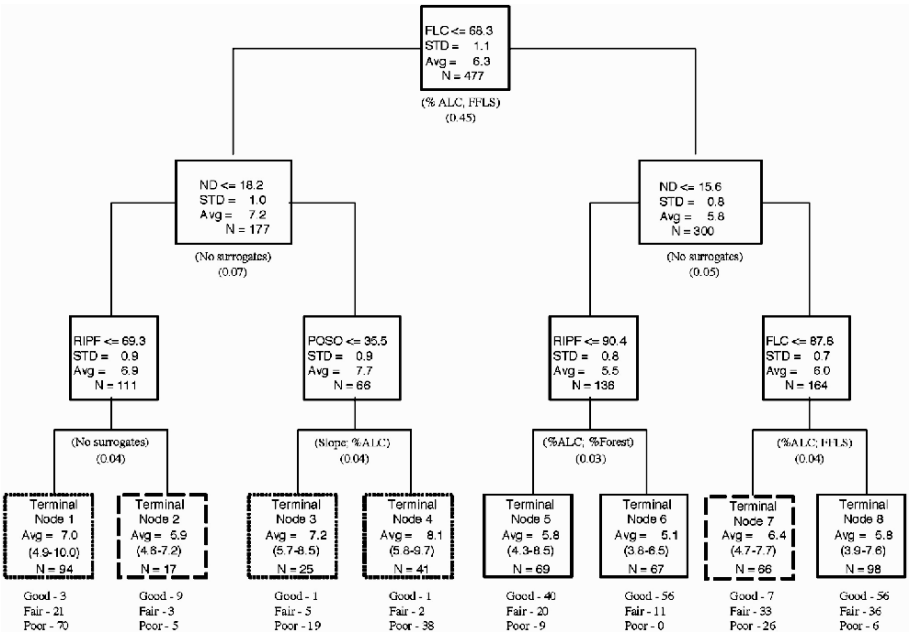


Figure 11.2. *Regression tree results for 477 Mid-Atlantic streams and associated watersheds. Variable breaks (threshold values), total number of paired stream/watershed samples, and the average total nitrogen concentration (Ln mg/L) are given at each interior node. Variable abbreviations are given in Table 11.1. Variables inside parentheses indicate surrogate variables (greater than >75% association). Numbers inside parentheses below surrogate variable names are the relative importance of that split in the model. Terminal nodes with solid line boxes indicate average nitrogen values in the good range; dashed line boxes indicate fair range; dotted line boxes indicate poor range (see discussion in Methods section). Numbers (and range) of stream samples in good, fair, and poor condition are given below each terminal node.*

At the next level on both sides of the tree, total wet nitrate deposition was the most important variable determining total nitrogen concentration in Mid-Atlantic streams (0.07 and 0.05 improvement in the model on the left and right sides of the tree, respectively) and there were no surrogate variables (Figure 11.2). Lack of a surrogate was expected because atmospheric nitrogen deposition was not correlated with any of the other variables used in the analysis. On the right side of the tree, watersheds with relatively low N deposition were further divided based on the amount of riparian forest; 69 watersheds had less than or equal to 90.4% riparian forest and 67 had greater than 90.4% riparian forest (Figure 11.2). Percent agriculture and percent forest at the watershed scale were surrogate values for this

split (Figure 11.2). Both terminal nodes (nodes 5 and 6) had average stream total N values in the good class, but those in the terminal node with greater amounts of riparian forest (Node 6) had the lowest total N concentrations (5.1 Ln N mg/L versus 5.8 Ln N mg/L, Figure 11.2). Additionally, the majority of watersheds in terminal node 6 had total stream N concentrations in the good condition class, with no watershed in poor condition, whereas terminal node 5 had a greater percentage of watersheds in fair and poor condition (Figure 11.2).

The group of watersheds on the right side of the tree with relatively higher atmospheric N deposition were further divided into a group of 98 watersheds with greater than 87.8% forest and a group with less than or equal to 87.8% forest; the former had average total N concentration in streams in the good range and the later in the fair range (Figure 11.2). Terminal node 8 also had a higher percentage of watersheds in good condition than did terminal node 7 (Figure 11.2). Surrogate variables included the amount of agriculture at the watershed scale, forest fragmentation, and the amount of agriculture on steep slopes.

On the left side of the tree, watersheds with relatively higher atmospheric N deposition were split into a group of 25 watersheds with relatively lower potential soil loss and a group of 41 watersheds with relatively higher potential soil loss (Figure 11.2); percent agriculture was a surrogate for potential soil loss in this part of the tree. Although both terminal nodes had streams with average total N concentration in the poor range, those watersheds in terminal node 3 (lower potential soil loss) had on average lower total N concentrations in streams than terminal node 4 (7.2 versus 8.1 Ln N mg/L), and the former had a greater percentage of watersheds in good or fair condition (Figure 11.2). Finally, watersheds with relatively lower atmospheric N deposition were divided into a group of 94 watersheds with less than or equal to 69.3% riparian forest and a group of 17 watersheds with greater than 69.3% riparian forest (Figure 11.2). Terminal node 2 had an average total N concentration in streams in the fair range whereas terminal node 1 had an average in the poor range, and the former had a higher percentage of watersheds in the good and fair range than did the later (Figure 11.2). There were no surrogate variables for the split of these watersheds and stream samples.

11.3.2 Geography and Characteristics of Terminal Nodes

Of the eight terminal node point themes analyzed, five were found not to differ significantly from a random distribution of points (Table 11.2). Node 1 points were found to be more clustered than would be expected by chance, and nodes 2 and 5 were found to be more uniformly spaced than would be expected by chance. However, few data points were available for analysis for node 2, and hence one should exercise caution when interpreting the results for this particular set of points (Table 11.2). Terminal node 5 sites tended to be widely dispersed in forested areas in the southern half of the Region (Figure 11.3). Terminal node 1 sites, which had on average relatively high concentrations of stream N, consisted of geographic clusters in the Washington, D.C. area and in valleys in the Appalachian Mountain region (Figure 11.3).

Table 11.2. *Summary of the nearest-neighbor analysis for the eight nodal point themes generated by Regression Tree Analysis. The R-statistic and |z|-statistic for the Clark and Evans (1954) test are shown along with a generalized p-value obtained from a normal-curve statistical table (Zar 1999); significant p-values (<0.05) are shown with an asterisk. The result column indicates whether the point theme is indicative of a random, clustered, or uniform spatial distribution.*

Node No.	*Sample Size*	*Polygon Boundary Pts*	*Polygon Area (m^2)*	*R-Statistic*	*|z|-Statistic*	*P-Value*	*Result*
1	94	9	1.828×10^{11}	0.751	4.624	<0.001*	Clustered
2	17	7	9.845×10^{10}	1.596	4.700	<0.001*	Uniform
3	26	5	8.197×10^{10}	0.986	0.134	>0.5	Random
4	41	9	7.445×10^{10}	0.969	0.386	>0.5	Random
5	69	12	1.173×10^{11}	1.130	2.059	<0.02*	Uniform
6	67	11	1.091×10^{11}	1.078	1.215	>0.2	Random
7	66	12	1.007×10^{11}	1.095	1.475	>0.2	Random
8	97	9	9.136×10^{10}	1.011	0.202	>0.2	Random

Table 11.3*. Summary of independent variable values for each terminal node. Values are means ± SD. Mean differences between the amount of riparian forest and total forest in watersheds are also given, as is the average log N concentration for each terminal node. See Table 11.1 for variable names.*

Metric	*Node1*	*Node2*	*Node3*	*Node4*	*Node5*	*Node6*	*Node7*	*Node8*
AGS3	20.9	10.2	16.0	33.7	9.3	1.8	10.9	1.7
	±14.4	±9.1	±6.9	±12.6	±6.1	±3.3	±5.7	±2.0
ALC	47.6	32.6	34.0	62.0	15.5	3.2	17.6	2.9
	±18.2	±14.1	±14.9	±14.1	±8.5	±5.5	±7.8	±2.8
FLC	40.1	55.4	60.0	35.1	81.3	95.9	78.7	96.0
	±12.9	±10.2	±11.1	±13.7	±8.6	±6.1	±5.7	±3.3
POSO	37.1	22.6	26.8	52.7	34.6	34.8	20.6	17.7
	±18.1	±12.5	±7.4	±12.9	±17.3	±24.8	±10.3	±14.9
RD	225.2	215.5	208.2	221.6	124.3	94.0	163.2	96.5
	±154.8	±88.2	±67.5	±56.6	±54.9	±89.3	±53.4	±51.5
RIPA	26.4	7.9	17.8	36.5	13.7	1.2	10.2	1.9
	±15.6	±6.3	±9.9	±15.1	±12.1	±1.9	±9.3	±3.4
RIPF	45.9	81.2	63.0	43.2	74.3	97.8	81.7	94.2
	±17.0	±9.5	±13.9	±18.9	±14.0	±2.9	±12.4	±7.5
SG	4.9	2.9	6.7	6.2	12.7	15.9	8.4	12.3
	+4.1	+2.2	+2.9	+2.6	+7.3	+7.6	+4.1	+5.2
FFLS	26.1	19.1	16.7	31.4	8.3	2.6	10.4	2.5
	+10.8	+7.1	+10.3	+15.0	+3.9	+3.0	+5.2	+2.3
WLC	0.5	0.6	20.3	19.7	13.4	13.3	19.7	19.8
	±0.9	±0.7	±0.9	±1.0	±1.3	±1.6	±1.6	±2.1
ND	14.2	14.4	16.7	31.4	8.3	2.6	10.4	2.5
	+2.2	+1.6	+10.3	+15.0	+3.9	+3.0	+5.2	+2.3
RIFL-FLC	5.8	25.7	3.1	8.1	−7.0	1.9	3.0	−1.7
LogNConc.	7.0	5.9	7.2	8.1	5.8	5.1	6.4	5.8

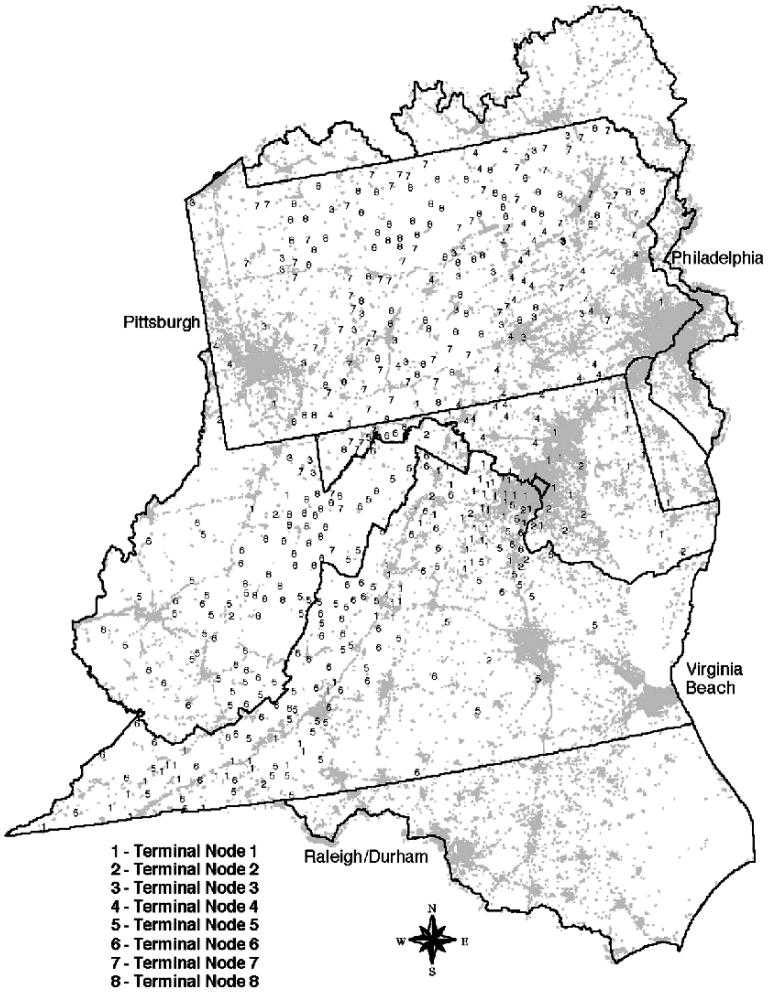


Figure 11.3. *Spatial distribution of terminal node members across the Mid-Atlantic region. Terminal node numbers correspond to numbers from the regression tree in Figure 11.2.*

Although only terminal node 1 watersheds were spatially clustered, there appeared to be a geographic pattern related to a combination of the amount of forest and wet nitrate deposition (Figure 11.3, Table 11.3). Those terminal nodes with watersheds with the greatest amount of forest and least amount of agriculture had the lowest average concentrations of total nitrogen in streams (terminal node 6, Table 11.3), but relatively higher atmospheric nitrate deposition increased total stream nitrogen even when the average amount of forest exceeded 78% (Table 11.3). For example, watersheds in terminal node 8 had on average 96% or more forest (Table 11.3) yet average total nitrogen values were only slightly in the good range (5.8 Ln

N mg/L). Moreover, watersheds in terminal node 7 had on average stream N concentrations in the fair range (6.4 Ln N mg/L, Figure 11.2). Watersheds in terminal nodes 7 and 8 tended to be in the northern part of the region where atmospheric wet nitrate deposition was the greatest (Table 11.3, Figure 11.3). The four terminal nodes with 68.3% or less forest had average total stream nitrogen values in the poor range, except for terminal node 2, which was in the fair range (but on average, only 0.1 Ln N mg/L from the good range, Table 11.3). The highest total stream nitrogen values were found in terminal node 4 which were in northern areas of region (Figure 11.3); these areas had high atmospheric nitrate deposition and relatively low amounts of forest (Table 11.3). Terminal node 2 had on average 25.7 percent more riparian forest than total forest in the watershed, by far the greatest difference in these two metrics (Table 11.3), and tended to be in central and southwest portion of the Region, but not entirely (Figure 11.3).

Table 11.4. *Average percent of watershed area composed of Omernik Level IV ecoregions summarized by regression tree terminal nodes. n = the number of watersheds in the terminal node. The total number of Ecoregions = 43.*

Terminal Node	*% Area*	*Terminal Node*	*% Area*
Node 1 (30 Ecoregions, n = 94)		*Node 2 (15 Ecoregions, n = 17)*	
Piedmont Uplands	23.7	Chesapeake Rolling Coastal Plain	29.4
No. Limestone Dolomite Valleys	14.2	Forested Hills and Mountains	11.8
So. Limestone Dolomite Valleys	13.9	No. Shale Valleys	10.6
Other Valleys	17.1	Other Valleys	7.1
Other Uplands and Plateaus	14.5	Uplands and Plateaus	11.0
Node 3 (19 Ecoregions, n = 26)		*Node 4 (16 Ecoregions, n = 41)*	
No. Limestone Dolomite Valleys	10.5	No. Shale Valleys	32.8
No. Sandstone Ridges	11.8	Piedmont Uplands	16.7
No. Shale Valleys	12.6	No.Limestone Dolomite Valleys	11.1
Pittsburgh Low Plateau	11.5	Other Valleys	8.2
Uplands and Valleys/Mixed La	10.6	Ridges and Mountains	14.3
Node 5 (26 Ecoregions, n = 69)		*Node 6 (19 Ecoregions, n = 67)*	
Northern Inner Piedmont	12.0	Cumberland Mountains	16.4
Northern Sandstone Ridges	10.6	Northern Sandstone Ridges	13.8
Other Mountains and Ridges	43.7	Southern Sandstone Ridges	11.2
Low Plateaus and Uplands	4.6	Other Ridges and Mountains	28.9
Valleys and Coastal	11.5	High and Interior Plateaus	20.0
Other Plateaus	6.8		
Node 7 (20 Ecoregions, n = 66)		*Node 8 (14 Ecoregions, n = 97)*	
Uplands and Valleys	18.5	Unglaciated Allegheney High Plains	26.7
Unglaciated_Allegheney_High_Plains	12.8	Forested Hills and Mountains	25.5
Northern Shale Valleys	10.1	Northern Sandstone Ridges	13.3
Mountains and Ridges	21.2	Northern Dissected Ridges	12.0
Other High Plateaus	8.2	Glaciated Allegheney Plateau	10.2
Low Plateaus	24.3	Coastal	4.7

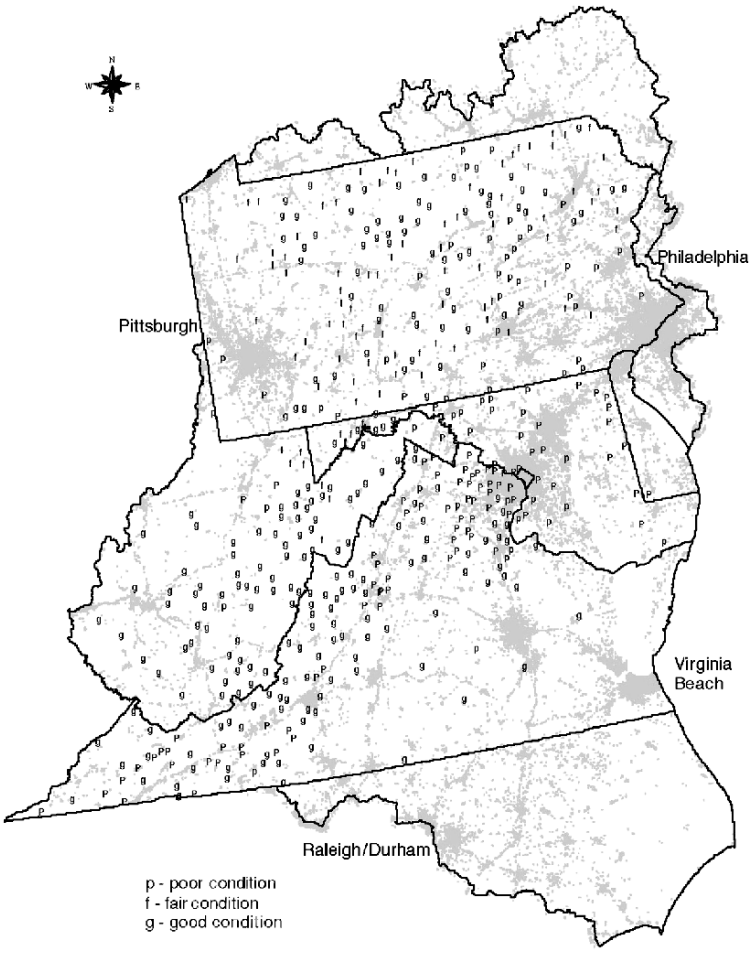


Figure 11.4. *Spatial distribution of condition classes for stream nitrogen based on average condition classes for terminal nodes given in Figure 11.2. p = poor, f = fair, g = good.*

A reclassification of terminal nodes into stream N condition classes (poor, fair, good) resulted in a geographic pattern that supports conclusions described above (Figure 11.4). Many of the good condition sites were in the mountainous areas of the Region where the total percentage of forested land cover at the watershed scale was relatively high. However, a relatively large percentage of the fair condition sites were in the northern part of the Region (e.g., Pennsylvania, Figure 11.4) where forests make up the majority of the land cover, but also where wet nitrate deposition is relatively high. Poor condition sites are widely distributed across the Region but tended to be in agricultural and urban areas (Figure 11.4).

There appeared to be no strong pattern of terminal node membership by ecoregion (Table 11.4), although terminal nodes with average N concentration

values in the good range tended to be in ecoregions with ridges, high plateaus, or mountains (Table 11.4). This is not surprising in that these areas are almost entirely forest. Terminal nodes that on average were in poor condition tended to be comprised of ecoregions characterized by valleys, mixed uplands and valleys, low plateaus, or Piedmont uplands (Table 11.4).

11.4 DISCUSSION

In our study, the regression tree analysis (RTA) produced a set of scaling structures that were derived from the entire set of independent variables representing regional, watershed, and riparian level scales. Although RTA has uncovered scaling relationships between environmental characteristics and breeding bird and fish species richness (O'Connor et al. 1996, Rathert et al. 1999), most scaling studies generally apply an a priori set of scales and model assumptions prior to the analysis. For example, many studies have used a hierarchical classification structure, such as ecoregions, to pre-stratify samples prior to analysis and model development (Herlihy et al. 1998, Detenbeck et al. 2000, Jensen et al. 2000). Moreover, many models are implemented from conceptual constructs with little empirical testing of variables and associated scales of importance (Aber 1997). Finally, RTA determines how scaling relationships among variables changes as biophysical conditions changes across the Region. Although results of multiple linear regression analysis have shown forest and atmospheric nitrogen deposition to be important determinants of stream nitrogen in the Mid-Atlantic region (Jones et al. 2001a), these results don't determine how the importance of these variables change in different biophysical settings across the region. Understanding how the importance of environmental variables changes in different biophysical settings is critical in determining area-specific environmental protection and management needs (Kennen et al. 2002). Moreover, an understanding of which environmental factors constrain local-scale, stream conditions help environmental managers determine the scale at which environmental improvements and policies must be pursued. Our results suggest that, in the central Appalachian Mountains and northeast and north central portions of the region, management of atmospheric nitrate deposition is critical in reducing stream nitrogen concentration even in relatively forested watersheds. Our results also suggest that in areas with relatively low atmospheric nitrate deposition it is possible to decrease nitrogen concentrations in streams by increasing the ratio of riparian forest to overall forest in the watershed, even when the overall amount of forest in the watershed would otherwise result in excess nitrogen (e.g., streams in poor condition based on nitrogen concentrations). In at least one watershed (Pohick Creek, Fairfax County, Virginia), this pattern appears to have resulted from zoning policies that protected forest and wetland in the riparian zone (Figure 11.5). Additionally, many of these watersheds were forested at the outflow areas near the stream sample locations. Forested areas near the outflow of the watershed may decrease stream nitrogen more than areas upstream due to greater overall biological assimilation of inorganic stream nitrogen occurring at the former (Wickham et al. 2003).

In our study, total nitrogen concentration in streams was determined first by the amount of forest at the watershed scale and second by the amount of wet nitrate

deposition at the regional scale. The amount of forest is primarily controlled by local-scale conversion of forest to farmland or residential and urban land use, but large, relatively continuous patches of forest exist in mountainous portions of the region because these areas have been difficult to develop (steep slopes and shallow soils that prohibit agriculture and roads, Jones et al. 2001b). Atmospheric wet nitrate deposition exhibits a regional pattern of increasing deposition from southwest to northeast (Clark et al. 2000). Higher deposition rates extend further south in the Appalachian Mountains (due to the strong positive association between elevation and deposition (Nodvin et al. 1995, SAMAB 1996). Although there was some spatial clustering of watershed members in individual terminal nodes (e.g., terminal node 1), differences in the scales and extent of these two major environmental constraints resulted in a relatively loose spatial structure – most terminal nodes were randomly distributed across the region. This is because the amount of forest on watersheds is not determined entirely by regional constraints, like topography and geology, and because large patches of forests do not exhibit the same regional pattern as atmospheric nitrate deposition. Similarly, nitrogen concentration in streams does not exhibit a strong ecoregion affiliation since patterns of atmospheric nitrate deposition were not included in the delineation of ecoregion boundaries, although mountainous ecoregions in the Mid-Atlantic region tend to have high amounts of forest and hence lower stream nitrogen concentrations.

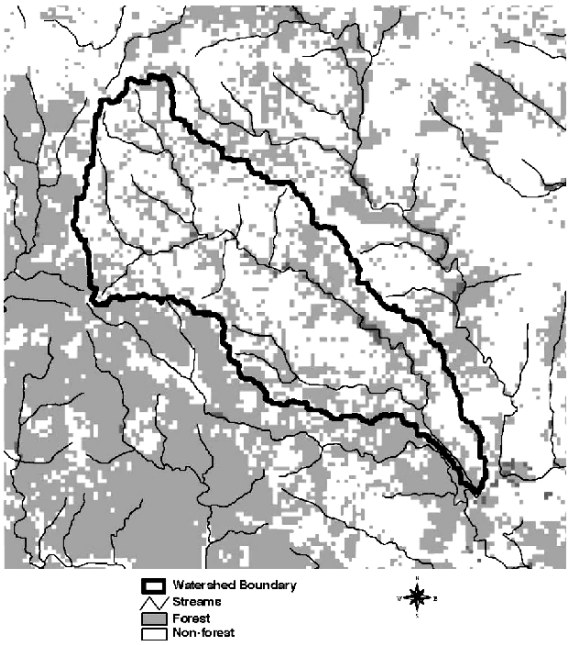


Figure 11.5. *Forest spatial distribution within the Pohick watershed. The Pohick watershed belongs to terminal node 2 and had a N concentration in the good range despite being in a highly developed watershed.*

11.4.1 Model Uncertainty and Sources of Error

The strength of our RTA model (the RTA equivalent of a multiple linear regression r^2 is 1 – the relative resubstitution error) compares favorably to other models predicting stream chemistry from landscape and watershed-scale metrics, where the stream data were based on single samples taken during low or base flow (Herlihy et al. 1998), and to RTA analysis relating multiple scale relationships between fish and bird species richness and environmental factors (Rathert et al. 1999, O'Connor et al. 1996.)

The Bagging analysis in CART® provides a way to evaluate the consistency of the tree generated by the primary analysis, including the importance of certain variables in determining splits in the tree. It compares consistency in the selection of variables and splits between the initial tree and randomly generated trees, as well as the sum of squares and sum of absolute deviations, which reflect the goodness of fit of the initial tree versus the consensus tree. In our study, the initial tree performed as well as the consensus tree. Another statistical routine, Random Forests, also is used to evaluate tree consistency and selection of variables in RTA. However, we did not have access to this program and were thus unable to use it to evaluate our data set.

There are several sources of error that might account for unexplained variance in our model, including errors in the land cover and soil data (Yang et al. 2001, Van Rompaey and Govers 2002, respectively), exclusion of key environmental factors, such as geology and ground water (Winters 2001), exclusion of finer-scale variables affecting local processes (Lawler and Edwards 2002, Rompaey and Govers 2002), and a mismatch in scales between single-season, low flow stream samples and watershed and regional patterns of land surface features and atmospheric nitrate deposition (Jones et al. 2001a). When data on stream nutrients include concentrations during peak events (from fixed, multiple-year stream samples as opposed to one-time samples), correlation coefficients between landscape characteristics and stream nutrient concentrations can increase to nearly 0.90 (Jones et al. 2001a). Moreover, local-scale processes, such as differences in fertilizer applications, become less important than watershed-scale landscape conditions when considering peak flow events (Vuorenmaa et al. 2001). Quantitative relationships between stream nitrogen and watershed- and regional-scale environmental characteristics strengthen because multiple-year stream samples match the temporal scale at which major disturbances occur across the watershed and region. It is during these peak events when the condition of the watershed has the greatest impact on stream condition (Jones et al. 2001a).

Sensitivity analysis may be one way to evaluate the relative importance of the different sources of error listed above in the RTA. This involves changing the values for one of the potential error sources (for example, land cover percentage) while holding the other variables constant. A similar analysis is then performed on the other variables to evaluate other potential sources of error and their impact on the model generated from the RTA. For a more thorough review of sensitivity analysis see Li and Wu (Chapters 3) and Law et al. (Chapters 9).

11.4.2 Changes in Scaling Relationships as an Indicator of Watershed Condition

It may be possible to use scaling functions between streams and watershed areas as an indicator of the health or condition of the aquatic system. Although hierarchy theory provides an important framework for the design of scaling studies (Allen and Starr 1982, O'Neill et al. 1989), the tendency has been to describe ecological processes and patterns as having discrete scales (Turner et al. 1989, O'Neill et al. 1991). Although such an approach helps reduce ecologically complex, middle-numbered systems so that their patterns and processes can be better defined and understood (Wu 1999), it may be important to define how scaling relationships change given certain levels of disturbance. One potential approach would be to evaluate how the strength of the relationships between stream chemistry and biology and overall watershed condition (e.g., the amount of atmospheric nitrate deposition and forest land cover at the watershed scale) changes as a function of increasing precipitation on the watershed. Figure 11.6 is a simple conceptual model of how this approach might work.

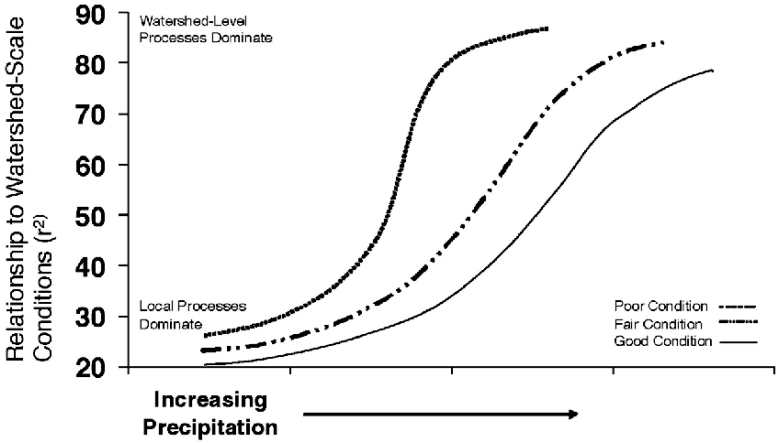


Figure 11.6. *Simple conceptualization of changes in scaling relationships between stream nitrogen and watershed-scale, landscape characteristics with increasing levels of precipitation at a regional scale.*

The strength of the relationship might be determined through a linear or non-linear multivariate approach, where some type of association statistic (e.g., correlation coefficient) would be used to assess changing relationships across scales. Watersheds or catchment areas with natural vegetation have a tendency to decrease energy (resulting from vegetation canopy interception) and increase infiltration (resulting from greater surface roughness) and hence respond (run-off, erosion, and sediment movement) more slowly to increasing precipitation than watersheds with higher amounts of anthropogenic and impervious surface cover (Jennings and Jarnagin 2002). In terms of scaling functions and hierarchy, stream segments with catchment areas consisting of large amounts of natural vegetation would remain

independent of watershed-scale conditions for longer periods of time given broader-scale disturbance events (e.g., broad-scale precipitation events) than stream segments associated with catchments with higher amounts of anthropogenic land cover (Figure 11.6). Local-scale processes affecting nutrient cycling and biological condition would tend to dominate longer in the former than in the latter. Changes in the strengths of associations between stream and watershed level conditions might help identify those areas at greatest risk to broad-scale, disturbance events. They also may be an indicator of the resiliency of the watershed to a broad-scale event or disturbance.

The scaling relationships between streams and broader-scale environmental conditions would likely vary among different biophysical settings. For example, higher gradient watersheds and associated streams would likely respond more rapidly to increasing precipitation than lower gradient watersheds due to higher energy associated with steep hillsides and slopes. Developing an indicator of aquatic ecosystem health based on changing scaling relationships would require significant enhancement of today's water monitoring networks. Fixed, multi-year samples of streams are critical in determining how stream chemistry, physical habitat, and stream biota respond to local versus watershed versus regional environmental conditions given different intensities of disturbance, yet long-term monitoring sites are mostly lacking. The National Water Quality Assessment (NAQWA) program is the only extensive program that systematically collects long-term data on water quality across the United States, but many of these studies have data collections covering only a few years. Moreover, long-term data on stream biology are nearly non-existent. Long-term biological data are critical in understanding the history of disturbance and its potential influence on current biotic community structure (Harding et al. 1998). The American Institute of Biological Sciences (AIBS) has proposed the National Ecological Observatory Network (NEON) to provide an infrastructure for collection of long-term biological and ecological data to fill these gaps. Such a network would dramatically enhance our ability to link stream biological conditions with broader-scale changes in landscape conditions and processes.

ACKNOWLEDGEMENTS

The authors would like to thank the EPA's Office of Research and Development (ORD) for support of this research. We thank Carolyn Hunsaker of the U.S. Forest Service for contributing STORET water quality data used in this study. The U.S. Environmental Protection Agency (EPA), through its Office of Research and Development (ORD), partially funded and collaborated in the research described here.

REFERENCES

Aber, J. D. 1997. Why don't we believe the models? Bulletin Ecological Society America 73:232-233.

Allan, J. D., D. L. Erickson, and J. Fay. 1997. The influence of catacment land use on stream integrity across multiple spatial scale. Freshwater Biology 37:149-161.

Allen, T. F. H., and T. B. Starr. 1982. Hierarchy: Perspectives for Ecological Complexity. University of Chicago Press, Chicago.

Ator, S. W., and M. J. Ferrari. 1997. Nitrate and Selected Pesticides in Ground Water of the Mid-Atlantic Region. U.S. Geological Survey Water Resource Investigative Report 97-4139, Baltimore, MD.

Behrendt, H. 1996. Inventories of point and diffuse sources and estimated nutrient loads – a comparison for different river basins in central Europe. Water Science and Technology 33:99-107.

Breiman, L., J. H. Friedman, O. R. A., and C. J. Stone. 1984. Classification and Regression Trees. Wadsworth, Pacific Grove, CA.

Carpenter, S. R., N. F. Caraco, D. L. Correll, R. W. Howarth, A. N. Sharpley, and V. H. Smith. 1998. Nonpoint pollution of surface waters with phosphorus and nitrogen. Ecological Applications 8:559-568.

Clark, G. M., D. K. Mueller, and M. A. Mast. 2000. Nutrient concentrations and yields in undeveloped stream basins of the United States. Journal of the American Water Resources Association 36:849-860.

Clark, P. J. and F. C. Evans. 1954. Distance to nearest neighbor as a measure of spatial relationship in populations. Ecology 34:445-453.

Davies, N. M., R. H. Norris, and M. C. Thoms. 2000. Prediction and assessment of local stream habitat features using large-scale catchment characteristics. Freshwater Biology 45:343-369.

Detenbeck, N. E., S. L. Batterman, V. J. Brady, J. C. Brazner, V. M. Snarski, D. L. Taylor, and J. A. Thompson. 2000. A test of watershed classification systems for ecological risk assessment. Environmental Toxicology and Chemistry 19:1174-1181.

Donnelly, K. P. 1978. Simulations to determine the variance and edge effect of total nearest-neighbour distance. Pages 91-95 *in* Hodder, I., editor. Simulation Studies in Archaeology. Cambridge University Press, Cambridge.

Environmental Protection Agency (EPA). 1996. Atmospheric Nitrogen Deposition Loading to the Chesapeake Bay: An Initial Analysis of the Cost Effectiveness of Control Options. EPA230-R-96-012, Annapolis, MD.

Environmental Protection Agency (EPA). 2000. Mid-Atlantic Highlands Stream Assessment. U.S. Environmental Protection Agency Region 3, Philadelphia, PA.

ESRI. 1996. Introduction to ArcView GIS. Environmental Systems Research Institute, Redlands, CA.

Griffith, G. E., J. M. Omernik, and A. J. Woods. 1999. Ecoregions, watersheds, basins, and HUCs: how state and federal agencies frame water quality. Journal of Soil and Water Conservation 1999:666-677.

Harding, J. S., E. F. Benfield, P. V. Bolstad, G. S. Helfman, and E. B. D. Jones. 1998. Stream biodiversity: the ghost of land use past. Proceedings of the National Academy of Sciences of the United States 95:14843-14847.

Herlihy, A. T., J. L. Stoddard, and C. B. Johnson. 1998. The relationship between stream chemistry and watershed land cover in the Mid-Atlantic region of the U.S. Water, Air, and Soil Pollution 105:377-386.

Holland, D. M., P. P. Principe, and J. E. Sickles. 1999. Trends in atmospheric sulfur and nitrogen species in the eastern United States for 1989-1995. Atmospheric Environment 33:37-49.

Hooge, P. N. and B. Eichenlaub. 1997. Animal movement extension to ArcView. Version 1.1. Alaska Biological Science Center, U. S. Geological Survey, Anchorage, AK.

Hughes, R. M., E. Rexstad, and C. E. Bond. 1987. The relationship of aquatic ecoregions, river basins, and physiographic provinces to the ichthyogeographic regions of Oregon. Copeia 1987:423-432.

Hunsaker, Carolyn T., and Daniel A. Levine. 1995. Hierarchical approaches to the study of water quality in rivers. BioScience 45:193-203.

Jennings, D. B., and S. T. Jarnagin. 2002. Changes in anthropogenic impervious surfaces, precipitation and daily streamflow discharge: a historical perspective in a Mid-Atlantic subwatershed. Landscape Ecology 17:471-489.

Jensen, M. E., R. Redmond, J. D. Benedetto, P. S. Bourgeron, and I. A. Goodman. 2000. Application of ecological classification and predictive modeling to broad-level assessment of ecosystem health. Journal of Environmental Monitoring and Assessment 64:197-212.

Jones, K. B., A. C. Neale, M. S. Nash, R. D. V. Remortal, J. D. Wickham, K. H. Riitters, and R. V. O'Neill. 2001a. Predicting nutrient and sediment loadings to streams from landscape metrics: a multiple watershed study from the United States Mid-Atlantic Region. Landscape Ecology 16:301-312.

Jones, K. B., A. C. Neale, T. G. Wade, J. D. Wickham, C. L. Cross, C. M. Edmonds, T. R. Loveland, M. S. Nash, K. H. Riitters, and E. R. Smith. 2001b. The consequences of landscape change on ecological resources: an assessment of the United States Mid-Atlantic Region, 1973-1993. Ecosystem Health 7:229-242.

Jones, K. B., K. H. Riitters, J. D. Wickham, R. D. Tankersley, R. V. O'Neill, D. J. Chaloud, E. R. Smith, and A. C. Neale. 1997. An Ecological Assessment of the United states Mid-Atlantic Region: A Landscape Atlas. US EPA/600/R-97/130. Office of Research and Development, U.S. Environmental Protection Agency, Washington, D.C.

Jones, K. B., J. D. Wickham, K. H. Riitters, and R. V. O. Neill. 1999. Taking a broader landscape view: an ecological assessment of the Mid-Atlantic Region of the United States from space. Pages 572-588 *in* Y. Pykh, D. E. Hyatt, and R. J. M. Lenz, editors. Environmental Indices Systems Analysis Approach. EOLSS Publishers Co., Ltd., Oxford, UK.

Kennen, J. G., and M. A. Ayers. 2002. Relation of Environmental Characteristics to the Composition of Aquatic Assemblages Along a Gradient of Urban Land Use in New Jersey, 1996-98. U. S. Geological Survey Water Resources Investigations Report 02-4069, Denver, CO.

Lammert, M., and J. D. Allan. 1999. Assessing biotic integrity of streams: effects of scale in measuring the influence of land use/cover and habitat structure on fish and macroinvertebrates. Environmental Management 23:229-234.

Lawler, J. J., and T. C. Edwards. 2002. Landscape patterns as habitat predictors: building and testing models for cavity-nesting birds in the Uinta Mountains of Utah, USA. Landscape Ecology 17:233-245.

Lee, K. E., R. M. Goldstein, and P. E. Hanson. 2001. Relation between fish communities and riparian zone conditions at two spatial scales. Journal of the American Water Resources Association 37:1465-1473.

Lowrance, R., R. Todd, J. J. Fail, O. Henderickson, R. Lenoard, and L. Asmussen. 1984. Riparian forests as nutrient filters in agricultural watersheds. Bioscience 34:374-377.

Nodvin, S. C., H. V. Miegroet, S. E. Lindberg, N. S. Nicholas, and D. W. Johnson. 1995. Acidic deposition, ecosystem processes, and nitrogen saturation in a high elevation Southern Appalachian watershed. Water, Air, and Soil Pollution 85:1647-1652.

O'Connor, R. J., M. T. Jones, D. White, C. Hunsaker, T. Loveland, B. Jones, and E. Preston. 1996. Spatial partitioning of environmental correlates of avian biodiversity in the conterminous United States. Biodiversity Letters 3:97-110.

O'Neill, R. V., A. R. Johnson, and A. W. King. 1989. A hierarchical framework for the analysis of scale. Landscape Ecology 3:193-205.

O'Neill, R. V., S. J. Turner, V. I. Cullinan, D. P. Coffin, R. Cook, W. Conley, J. Brunt, J. M. Thomas, M. R. Conley, and J. Gosz. 1991. Multiple landscape scales: an intersite comparison. Landscape Ecology 5:137-144.

Omernik, J. M. 1987. Aquatic ecoregions of the conterminous United States. Annals of the Association of American Geographers 77:118-125.

Peterjohn, W. T., and D. L. Correll. 1984. Nutrient dynamics in agricultural watersheds: Observations on the role of a riparian forest. Ecology 65:1466-1475.

Rathert D., D. White, J. Sifneos, and R. M. Hughes. 1999. Environmental associations of species richness in Oregon freshwater fishes. Journal of Biogeography 26:257-273.

Roth, N. E., J. D. Allan, and D. L. Erickson. 1996b. Landscape influences on stream biotic integrity assessed at multiple spatial scales. Landscape Ecology 11:141-156.

Smith, J. H., J. D. Wickham, D. Norton, T. G. Wade, and K. B. Jones. 2001. Utilization of landscape indicators to model potential pathogen impaired waters. Journal of the American Water Resources Association 37:805-814.

Smith, R. A., G. E. Schwarz, and R. B. Alexander. 1997. Regional interpretation of water-quality monitoring data. Water Resources Research 33:2781-2798.

Sourthern Appalachian Man and the Biosphere Program (SAMAB). 1996. The Southern Appalachian Assessment: Atmospheric Technical Report, Chapter 5, Acid deposition. http://www.sunsite.utk.edu/samab/saa/reports/atmospheric/atmospheric.html.

Stauffer, J. C., R. M. Goldstein, and R. M. Newman. 2000. Relationship of wooded riparian zone and runoff potential to fish community composition in agriculture streams. Canadian Journal of Fish and Aquatic Sciences 57:307-316.

Steinberg, D., and P. Colla. 1995. CART: Tree-Structured Non-Parametric Data Analysis. Salford Systems, San Diego, CA.

Talmage, P. J., J. A. Perry, and R. M. Goldstein. 2002. Relation of instream habitat and physical conditions to fish and community of agricultural streams in the Northern Midwest. North American Journal of Fisheries Management 22:825-833.

Turner, M. G., V. H. Dale, and R. H. Gardner. 1989. Predicting across scales: Theory development and testing. Landscape Ecology 3:245-252.

Van Rompaey, A. J. J., and G. Govers. 2002. Data quality and model complexity for regional scale soil erosion prediction. International Journal of Geographical Information Science 16:663-680.

Vogelmann, J. E., S. M. Howard, L. Yang, C. R. Larson, B. K. Wylie, and N. V. Direl. 2001. Completion of the 1990s national land cover data set for the conterminous United States from Landsat Thematic Mapper data and ancillary data sources. Photogrammetric Engineering and Remote Sensing 67:650-662.

Vuorenmaa, J., S. Rekolainen, A. Lepisto, K. Kennttamies, and P. Kauppila. 2001. Losses of nitrogen and phosphorus from agricultural and forest areas in Finland during the 1980s and 1990s. Journal of Environmental Monitoring and Assessment 76:213-248.

Wickham, J. D., T. G. Wade, K. H. Riitters, R. V. O'Neill, J. H. Smith, E. R. Smith, K. B. Jones, and A. C. Neale. 2003. Upstream-to-downstream changes in nutrient export risk. Landscape Ecology 18:193-206.

Winters, T. C. 2001. The concept of hydrologic landscapes. Journal of the American Water Resources Association 37:335-350.

Wu, J. 1999. Hierarchy and scaling: extrapolating information along a scaling ladder. Canadian Journal of Remote Sensing 25:367-380.

Wu, J., and R. Hobbs. 2002. Key issues and research priorities in landscape ecology: an idiosyncratic synthesis. Landscape Ecology 17:355-365.

Yang, L., S. V. Stehman, J. H. Smith, and J. D. Wickham. 2001. Thematic accuracy of MRLC land cover for the eastern United States. Remote sensing of Environment 76:418-422.

Zar, J. H. 1999. Biostatistical Analysis, 4th edition. Prentice Hall, Upper Saddle River, NJ.

CHAPTER 12

UNCERTAINTY IN SCALING NUTRIENT EXPORT COEFFICIENTS

JAMES D. WICKHAM, K. BRUCE JONES, TIMOTHY G. WADE, AND KURT H. RIITTERS

12.1 INTRODUCTION

Nutrient export coefficients are estimates of the total load or mass of nitrogen (N) or phosphorus (P) exported from a watershed standardized to unit area and unit time (e.g., kg ha^{-1} yr^{-1}). They have been widely used as tools for environmental management of lakes, rivers and coastal waters because excessive nutrient enrichment (eutrophication) leads to several negative environmental effects (Carpenter et al. 1998). Yuan and Norton (2003) have shown, for example, that moderate increases in phosphorus concentrations in mid-Atlantic headwater streams shift benthic community structure toward a greater abundance of algae feeders, and others have shown that excessive nutrient input leads to "blooms" of nuisance algae (Paerl 1988, see Carpenter et al. 1998). Because of the potential negative environmental effects of excessive nutrient input, some programs are developing targeted goals of nutrient export by watershed (e.g., Linker et al. 1996).

Nutrient export has been linked strongly to watershed land-cover composition (Beualac and Reckhow 1982, Frink 1991, Panuska and Lillie 1995). As forest is replaced by urban and agriculture, both nutrient export averages and variances tend to increase. Thus, the spatial heterogeneity of land cover in a watershed is an important influence on nutrient export.

Watershed nutrient export coefficients are widely reported in the literature for watersheds ranging in size from 10^2 to 10^5 hectares (Dickerhoff Delwiche and Haith 1983, Lowrance et al. 1985, Clesceri et al. 1986, Jordan et al. 1997, Fisher et al. 1998), but have not been investigated for scale effects despite the wide range in watershed sizes for which they have been reported. Watersheds are analogous to grain (i.e., pixel size) in raster maps. Just as the size of a pixel can be changed, watersheds can also be resolved into few or many units for a given area. The

J. Wu, K.B. Jones, H. Li, and O.L. Loucks (eds.),
Scaling and Uncertainty Analysis in Ecology: Methods and Applications, 225–237.

watersheds used in the studies cited above could have been divided into two or more subunits with nutrient export coefficients estimated for each.

Dividing watersheds into one or more subunits represents one approach to scaling. Scaling nutrient export introduces at least two issues related to error analysis and uncertainty (see Li and Wu, 2003, Chapter 3). First, moving nutrient loads across subwatersheds requires estimation of loss rates as they move from upstream to downstream. In-stream decay of nutrients varies as function of discharge and geographical location (Smith et al. 1997, Peterson et al. 2001). Second, the total load exported from the entire watershed is dependent on whether nutrient export among subwatersheds is assumed to be independent. Dependence exists if nutrient export patterns among subwatersheds are similar. Subwatersheds can be assumed to behave independently if their nutrient export values are not strongly similar. The purpose of this chapter is to examine the effects of these two sources of uncertainty on estimated nutrient export as a function of the number of subwatersheds resolved.

12.2 MODELING NUTRIENT EXPORT

Modeling nutrient export as a function of watershed resolution encompassed three methodological steps. These steps were: (1) development of a model simulating nutrient export, (2) systematic changes to subwatershed resolution, and routing of nutrient export across subwatersheds, and (3) examination of changes in nutrient export as a function of subwatershed resolution.

We selected the Deer Creek watershed in northeastern Maryland (Figure 12.1) to examine scale effects because discharge data were available for the main tributary, and these data were needed to move N and P across subwatershed boundaries. The Deer Creek watershed is located on the western shore of the Chesapeake Bay, and drains into the Susquehanna River just before it empties into the Bay. Deer Creek is about 44,000 hectares, and forest, agriculture, and urban land-cover percentages are about 39%, 60%, and 1%, respectively (Vogelmann et al. 2001). The land-cover classes are well distributed spatially throughout the watershed; neither forest, nor agriculture, nor urban are concentrated in one or more sections of the watershed. A digital copy of the Deer Creek watershed was provided by the state of Maryland, Department of Natural Resources (MD DNR).

We used a nutrient export simulation model based on land-cover composition to estimate total loads of N and P (Wickham et al. 2000). The nutrient export model was constructed by fitting existing, empirical data for watersheds with homogenous (or nearly so) land cover (Reckhow et al. 1980) to exponential, Weibull, normal, and log-normal theoretical distributions. The empirical data were found to best fit log-normal distributions (Wickham et al. 2000). The fitted log-normal distributions were used to model N and P loads for mixed-use watersheds using Equation 5 from Reckhow et al. (1980):

$$L = \sum_{j=1}^{n} \sum_{i=1}^{3} c_i A_i \tag{12.1}$$

where L is the load (mass) of N or P, c is the returned coefficient from statistical resampling of the log-normal distributions for i land-cover classes (urban, agriculture, forest) and A is the area of each land-cover class. The model is iterated (i.e., $j = 1, n$) to estimate mean and variance. Nutrient export coefficients were calculated by dividing the output from Equation 12.1 by the total watershed area. Percentages of forest, agriculture, and urban by subwatershed were acquired from the National Land Cover Data (NLCD) (Vogelmann et al. 2001). A previous evaluation of Equation 12.1 indicated that land-cover composition could be used to simulate nutrient export for individual watersheds where monitored data were lacking (Wickham et al. 2003).

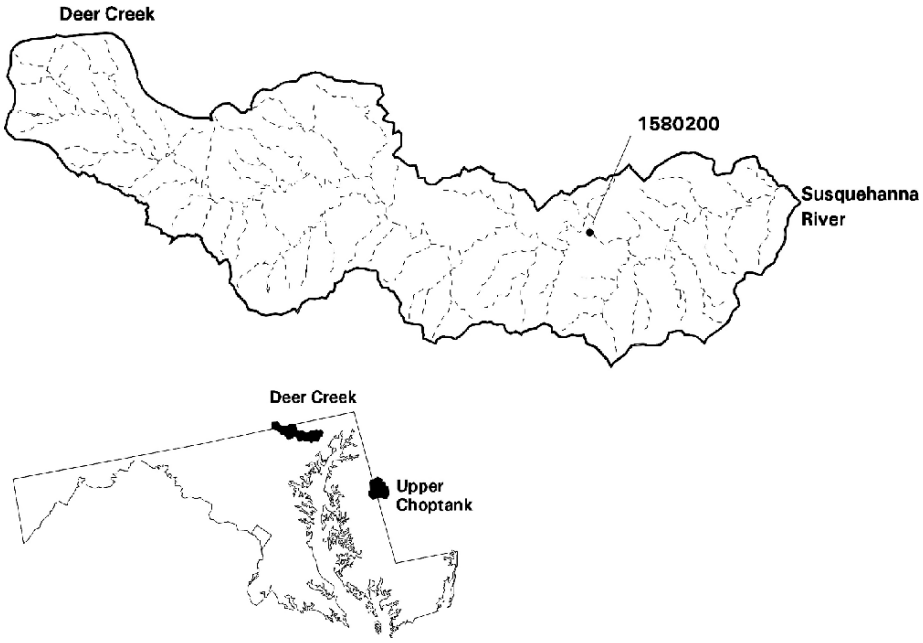


Figure 12.1. *Location Map. The solid lines identify the watershed boundary and the dashed lines identify the stream network. The solid circle identifies the location of the USGS gauging station used for discharge estimates. The location of the Deer Creek watershed within the state of Maryland is shown on the inset map. The location of the upper Choptank watershed is also shown on the inset map because it is referenced in the text.*

The nutrient export model (Equation 12.1) estimates the export from an individual watershed or subwatershed, but does not account for contributions from upstream neighbors (Wickham and Wade 2002). For subwatersheds that have upstream neighbors, some fraction of the exported load from the upstream subwatershed must be included in the load estimated for the downstream subwatershed. Accounting for upstream contributions was accomplished by modifying Equation 12.1:

$$L = \sum_{j=1}^{n} \left(\sum_{i=1}^{3} c_i A_i + \sum_{i=1}^{3} c_i A_i e^{(-d \cdot T)} \right) \quad (12.2)$$

where c, A, i, and j are the same as in Equation 12.1, e is the base of the natural logarithm, d is a decay or in-stream loss coefficient, and T is travel time (in days). The segment of Equation 12.2 after the plus sign represents the load from the upstream subwatershed that is exported to the downstream subwatershed and decayed in the main channel of the downstream subwatershed. The decay coefficient (d) accounts for loss of N and P as they move along the stream channel as a result of biotic and abiotic processes (Hill 1979, Burns 1998). We used the upper and lower bounds of the 90th percentile confidence interval from Smith et al. (1997) for d (0.2981 and 0.4768 for N; 0.1885 and 0.3497 for P) to capture the uncertainty associated with estimating in-stream processing of nutrients. Travel time (T) is a function of stream velocity, which in turn is related to stream discharge (Leopold and Maddock 1953). We chose a stream velocity of 0.6 meters per second based on reported empirical relationships between discharge and velocity (Dewald et al. 1985) and long-term discharge for a gauging station on the main stream of the Deer Creek watershed (station #1580200, *http://water.usgs.gov*).

We resolved the Deer Creek watershed into two to 20 subwatersheds to examine scale effects. Changes to subwatershed resolution were accomplished using an equal stream length criterion: at each resolution, each subwatershed contained approximately equal lengths of the main stream. Identification of equal stream lengths was accomplished by isolating the main tributary as a separate data set, and using GIS software to locate the points along the main tributary that split that stream into two to 20 equal lengths. These points were then used to identify the boundaries of each subwatershed where each point served as the subwatershed outlet.

Equation 12.2 was applied at all subwatershed resolutions from two to 20 under two different assumptions. Under one assumption, the value of c for land-cover class i was selected independently for each subwatershed. Thus, there were 6 (2 watersheds) to 60 (20 watersheds) random draws of c at each iteration of the model, one for each of the three land-cover classes (urban, agriculture, forest) in each subwatershed. Independence was based on the assumption that each subwatershed potentially represented a different ecosystem in regard to nutrient export. Differences in land ownership, topography, soils and other factors from one subwatershed to the next would translate into different rates of nutrient export for each subwatershed. Under the second assumption, only a single value of c for each land-cover class was used for all subwatersheds. That is, only three values of c were used regardless of subwatershed resolution. Comparison of the two outputs measures the effect of the assumption of independence and how that effect changed with subwatershed resolution.

For each subwatershed scale, the nutrient export model was iterated 10,000 times, and the effect of scale was measured as changes in variance of the annual nutrient export coefficients. Nutrient export coefficients show considerable intra-site, temporal variance (Reckhow et al. 1980, Lowrance et al. 1985, Fisher 1998, Panuska and Lillie 1995), and iteration of our model over 10,000 runs was designed

to simulate that variance. We focused on variance, rather than the mean, because variance is a more relevant measure for environmental management. For example, comparison of temporal N and P export data for the upper Choptank River (Fisher et al. 1998) against the Chesapeake Bay nutrient export goals (Linker et al. 1996) indicate that annual N export for the upper Choptank has exceeded the minimum standard for N, ~7.0 kg ha^{-1} yr^{-1}, 4 times in ten years even though the ten-year average for the Choptank is only 6.18 kg ha^{-1} yr^{-1}.

Variance was expressed as the difference between the 5th (Q_5) and 95th (Q_{95}) percentiles. We chose the difference between Q_5 and Q_{95} rather than the classical measure of variance ($\sum (x-\bar{x})^2/n-1$). The classical measure of variance estimates the average departure from the mean. The percentile range more closely matches changes in measured export values that might be realized from one year to the next.

12.3 WATERSHED NUTRIENT EXPORT SCALING RELATIONSHIPS

Scaled relationships between subwatershed resolution and nutrient export variance were quantified using a power law (Schneider 1998, 2001) (Table 12.1). A Euclidean exponent of negative one (–1) translates to a 50% reduction in variance for every doubling of subwatershed resolution (e.g., 5, 10, 20, etc.), whereas an exponent approaching zero would indicate that subwatershed resolution had little effect on nutrient export variance. Exponents close to zero would suggest that scale was not an important factor to consider when quantifying watershed nutrient export.

Table 12.1. *Scaled relationships between nutrient export and subwatershed resolution.*

Assumption	*Nutrient*	*Decay*	*Model*[*]	R^2
Independence	N	–0.2981	$\delta Q = 16.2\ s^{-0.427}$	0.98
	N	–0.4768	$\delta Q = 15.7\ s^{-0.463}$	0.98
	P	–0.1885	$\delta Q = 1.61\ s^{-0.419}$	0.98
	P	–0.3497	$\delta Q = 1.51\ s^{-0.432}$	0.98
Dependence	N	–0.2981	$\delta Q = 21.1\ s^{-0.036}$	0.71
	N	–0.4768	$\delta Q = 19.9\ s^{-0.060}$	0.91
	P	–0.1885	$\delta Q = 1.91\ s^{-0.020}$	0.26
	P	–0.3497	$\delta Q = 1.84\ s^{-0.046}$	0.74

[*] $\delta Q = |Q_5 - Q_{95}|$ and s equals the number of subwatersheds resolved (2 to 20).

Under independence, there was a strong relationship between the scale of subwatershed resolution and nutrient export variance, with exponents less than –0.40 for all combinations of nutrients and in-stream decay rates. Independence assumes spatial heterogeneity is an important factor in controlling nutrient export by treating subwatersheds as distinct ecosystems that have differences in land use

practices, soils, topography, and other factors. The net effect of independence was dampened nutrient export variance (Figures 12.2 and 12.3). Dampening occurred because there was an increased likelihood that at least one subwatershed had high nutrient export, thereby increasing the minimum simulated export values (e.g., Q_5) as subwatershed resolution increased. At the same time, there was a decreased likelihood that many subwatersheds had high nutrient export, thereby decreasing the maximum simulated export values (e.g., Q_{95}). Assuming independence had the effect of reducing nutrient export variance as subwatershed resolution increased.

Another pattern in Figures 12.2 and 12.3 is variance stability with increasing subwatershed resolution. One interpretation of the scaling relationship under independence is that it marks the maximum number of watersheds needed to study nutrient export (Palmer 1988, Jelinski and Wu 1996, Marceau 1999). The relationships between subwatershed resolution and nutrient export variance in Figures 12.2 and 12.3 indicate variance stability at about 10 subwatersheds.

When the assumption of independence among subwatersheds was removed, scaled relationships between subwatershed resolution and nutrient export variance disappeared (Figure 12.4), with Q_5 and Q_{95} remaining relatively constant at about 3 and 21 kg ha^{-1} yr^{-1}, respectively. The seven-fold difference between Q_5 and Q_{95} translated into a potential range of 792,000 kg yr^{-1} in annual load when multiplied by the watershed area. In contrast, the same range was only 176,024 kg yr^{-1} under independence (high in-stream decay, 20 subwatersheds).

The relative importance of independence versus dependence and in-stream decay rates in uncovering scaled relationships between nutrient export variance and subwatershed resolution are evident in the exponents. Comparison of model exponents across assumptions for the same in-stream decay rate reveals the importance of independence versus dependence, whereas comparison within assumptions for different decay rates reveals the importance of in-stream processes. The difference in model exponents across assumptions for the same decay rate was about −0.4 in all cases, whereas the difference in model exponents for different decay rates while holding the assumption constant was 0.024 to 0.036 for N and 0.013 to 0.026 for P (Table 12.1). The assumption of independence (or lack thereof) was the primary factor producing the scaled relationship between nutrient export variance and subwatershed resolution.

12.4 IMPORTANCE AND IMPACT OF UNCERTAINTY SOURCES

Scaling nutrient export coefficients as a function of subwatershed resolution (i.e., moving from Equation 12.1 to Equation 12.2) introduced two sources of uncertainty: 1) estimation of in-stream decay rates and 2) and similarity in nutrient export behavior among subwatersheds (independence versus dependence). Both sources of

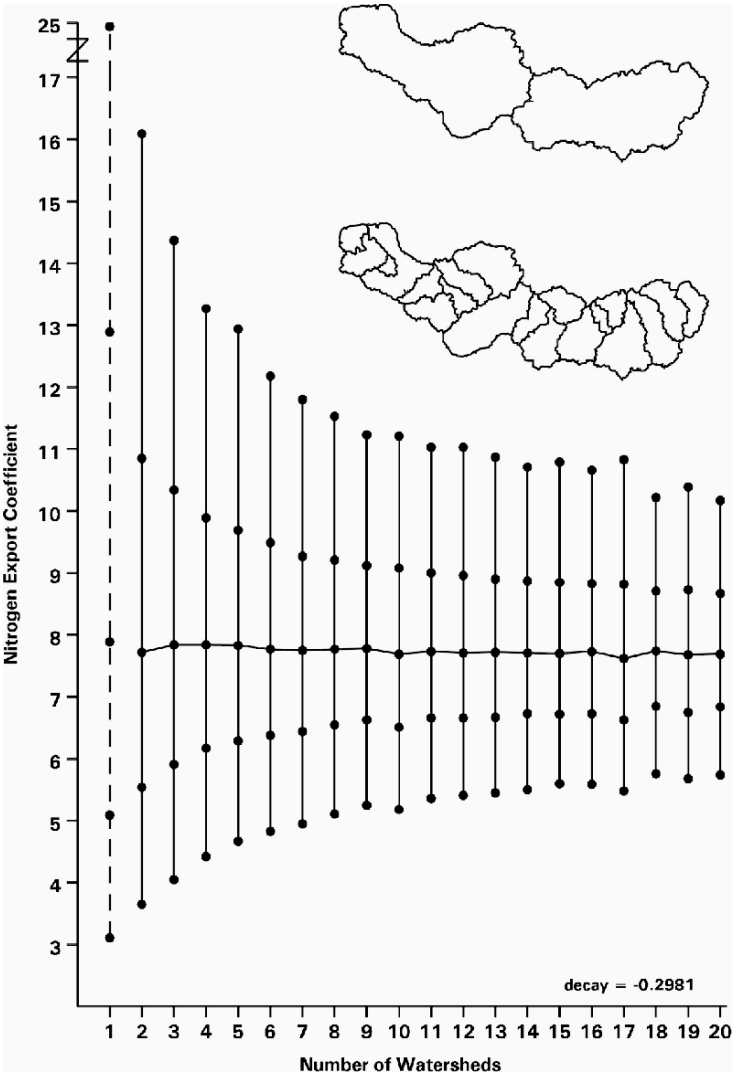


Figure 12.2. *Variance in N export for the terminal subwatershed as a function of the number of subwatersheds resolved. Y-axis units are kg ha^{-1} yr^{-1}. Amplitude of the vertical line is the variance in nutrient export as expressed by the range for the 5th to 95th percentiles. Reading from the bottom to the top, the dots represent the 5th, 25th, 50th, 75th, and 95th percentiles. A line connects the 50th percentiles to aid visual orientation. A dashed line is used to depict N export variance when Deer Creek is treated as a single watershed, because the model does not incorporate decay unless there are two or more subwatersheds. Model results were generated under the assumption of independence of nutrient export among subwatersheds. The inset shows Deer Creek resolved into 2 and 20 subwatersheds.*

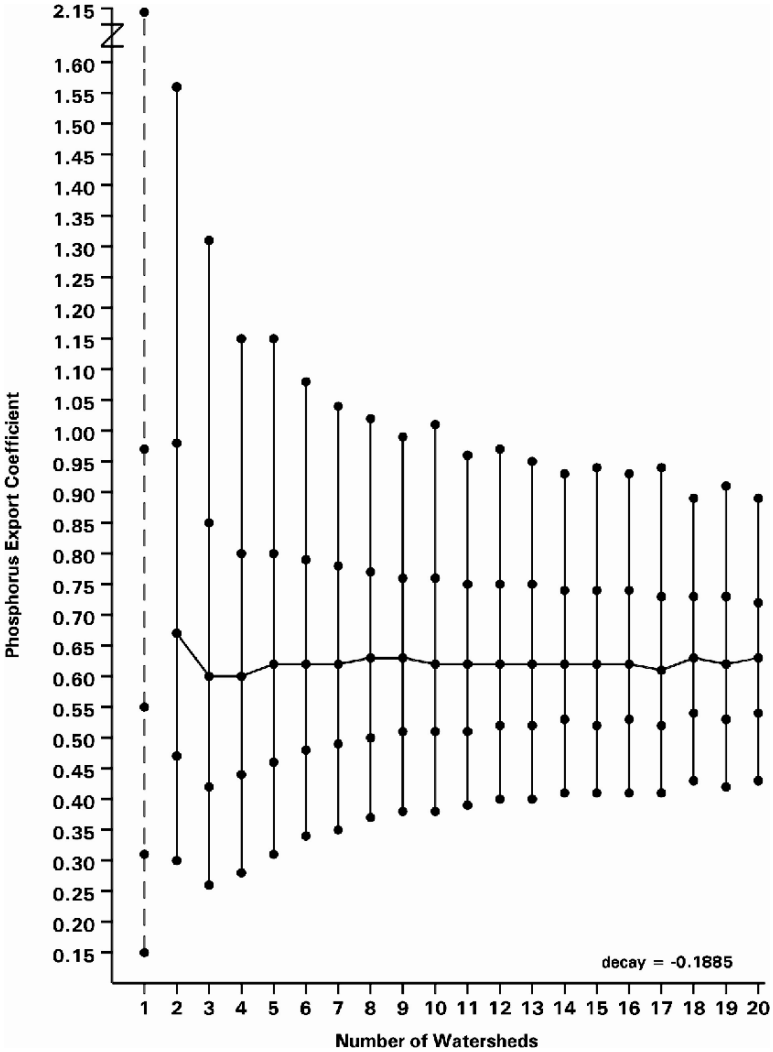


Figure 12.3. *Variance in P export for the terminal subwatershed as a function of the number of subwatersheds resolved. The lines, dots and Y-axis units are the same as in Figure 12.2. Model results were generated under the assumption of independence of nutrient export among subwatersheds.*

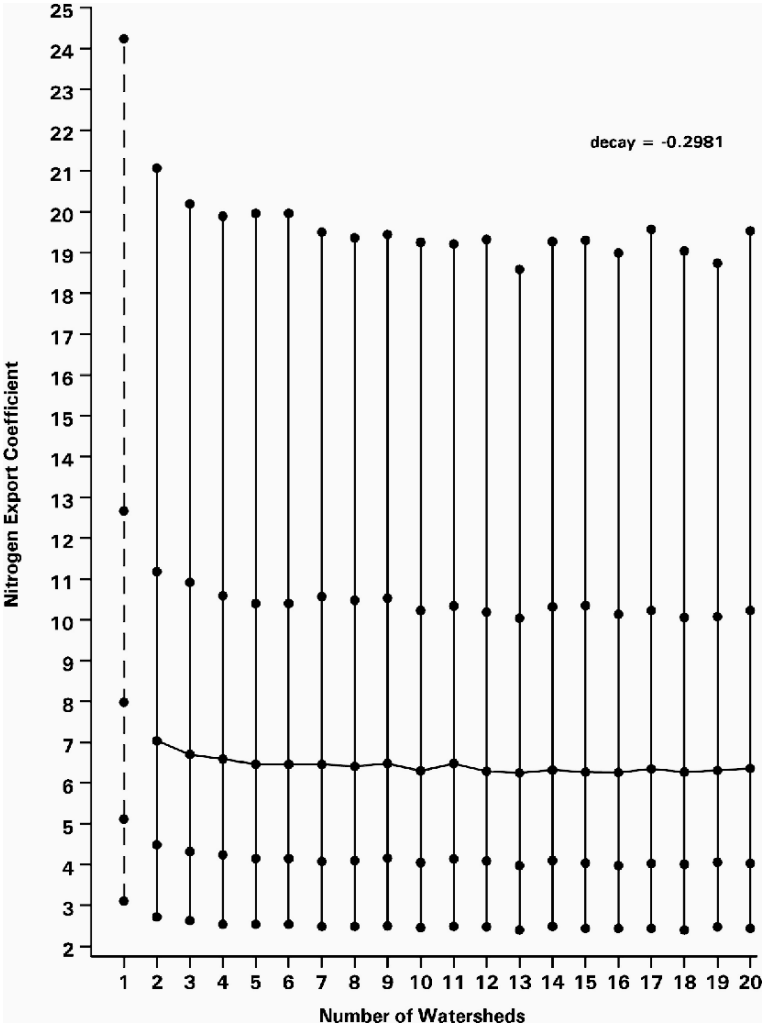


Figure 12.4. *Variance in N export as a function of the number of subwatersheds resolved with the assumption of independence removed. The lines, dots, and Y-axis units are interpreted the same as in Figures 12.2 and 12.3.*

uncertainty are relevant to environmental management issues. Higher in-stream decay rates reduce the total load of N and P exported from Deer Creek to the Susquehanna River and Chesapeake Bay, and independence suggests that ecosystems operate to reduce the magnitude and frequency of extreme events. High rates of in-stream decay and independence provide ecosystem services (Westman 1977, Costanza et al. 1997) that can foster more effective environmental management. Our simulation results suggest that the Chesapeake Bay N management goal (Linker et al. 1996) of ~8.0

kg ha^{-1} yr^{-1} for Deer Creek (Wickham et al. 2003) would be met 80% of the time under the combination of high in-stream decay and independence (Table 12.2).

Table 12.2. *Upper percentiles of N export for 20 subwatersheds.*

	Independence		*Dependence*	
	In-stream Decay		*In-stream Decay*	
Percentiles	*High*	*Low*	*High*	*Low*
Q_{50}	6.87	7.69	5.52	6.36
Q_{60}	7.15	8.10	6.47	7.45
Q_{70}	7.54	8.50	7.94	8.97
Q_{80}	7.99	9.00	9.90	11.16
Q_{90}	8.58	9.70	13.28	15.06

Published in-stream decay rates are highly variable because of their relationship to factors such and discharge and climate (Smith et al. 1997, Preston and Brakebill 1999, Alexander et al. 2000, Peterson et al. 2001). In-stream decay rates are surrogates for biotic and abiotic processes such as assimilation and sedimentation. Rates of in-stream decay have been measured only in a few studies, and the measurements have focused on N (e.g., Hill 1979, 1981, Sjodin 1997, Burns 1998, Peterson et al. 2001). We chose to frame uncertainty surrounding estimation of in-stream decay by selecting the upper and lower 90^{th} percentile confidence intervals for N and P in-stream decay from Smith et al. (1997), and it is unclear if more rigorous analyses (see Li and Wu, Chapter 3) would have improved our results.

Ideally, direct evaluation of independence versus dependence would be based on long-term records of observed data of input (atmospheric, anthropogenic) and output (discharge) of N and P for several subwatersheds. Comparison would be made for subwatersheds with similar land-cover compositions since our model, and the empirical literature on which it is based, show that land-cover composition is a strong driver of nutrient export. Regression relationships of input versus output could be used to test for significantly different slopes between subwatersheds with similar land-cover compositions. Significantly different slopes would suggest that subwatershed-to-subwatershed differences in topography, soils and other factors have resulted in dissimilar patterns of nutrient export (independence) between neighboring subwatersheds. These regression relationships would also be developed for sets of subwatersheds across different locations. Consistency in the regression relationships for several subwatershed sets across different watersheds would be needed to confidently rely on one assumption or the other.

N and P input and output data at the subwatershed level for the Little River watershed near Tifton, Georgia (Lowrance et al. 1985) provide some evidence supporting independence. The size and land-cover compositions of subwatersheds N and O within Little River are similar, but their slopes for N and P input versus export are different for over the years 1979 through 1981. Temporal data on N and P

do not appear to be available for years subsequent to 1981 (*http://www.tifton.uga.edu/sewrl/archived_data.htm*).

Evidence for independence versus dependence could also be inferred from long-term data for a single watershed. For example, long-term data for Deer Creek, if it existed, could be compared to the simulated variances in Figures 12.2 and 12.4 to draw an inference on which assumption was more valid. Data from the upper Choptank (Fisher et al. 1998; see Figure 12.1) have not been collected at a subwatershed level, but the N and P export ranges from its ten-year temporal record for the entire watershed can be compared to the range of simulated values for Deer Creek presented in Figures 12.2 and 12.4. Temporal variance (1981-1990) in N and P export for the upper Choptank were 8.61 (2.89 to 11.5 kg ha^{-1} yr^{-1}) and 0.51 (0.14 to 0.65 kg ha^{-1} yr^{-1}), respectively. The δQ values for N across all decay values and subwatershed resolutions were 4.00 to 12.44 under independence, but were 14.75 to 18.35 under dependence. Likewise, the δQ values for P across all decay values and watershed resolutions were 0.42 to 1.26 under independence, but were 1.49 to 1.79 under dependence. The simulated ranges for N and P export for Deer Creek under independence agree more closely with the temporal record for the upper Choptank.

The temporal data from the upper Choptank and Little River watersheds are limited by a short time span. Variance in its N and P export coefficients for the upper Choptank may increase as more observations are added, favoring dependence. Likewise, a longer temporal record for the Little River might ultimately favor dependence.

12.5 SUMMARY AND CONCLUSION

Our modeling of nutrient export treats each subwatershed as an ecosystem unit that imports, produces, removes, and exports nutrients (Leibowitz et al. 2000). Evidence for scaling relationships between nutrient export coefficients and subwatershed resolution appeared to be most strongly tied to production of nutrients for each subwatershed. Scaling was evident when production of nutrients among subwatersheds was assumed to be dissimilar (i.e., independence), but largely disappeared when nutrient production among subwatersheds was assumed to be similar (i.e., dependence). There is little data to test the assumption of independence, and hence scaling of nutrient export coefficients is hindered more by lack of empirical evidence supporting the plausible theoretical foundation that subwatersheds behave independently in terms of their nutrient export patterns than methodological constraints to test that theoretical foundation (Wu and Li, Chapters 1 and 2).

REFERENCES

Alexander, R. B., R. A. Smith, and G. E. Schwarz. 2000. Effect of stream channel size on the delivery of nitrogen to the Gulf of Mexico. Nature 403:758-761.

Beaulac, M. N. and K. H. Reckhow. 1982. An examination of land use - nutrient export relationships. Water Resources Bulletin 18:1013-1024.

Burns, D. A. 1998. Retention of NO^{-3} in an upland stream environment: a mass balance approach. Biogeochemistry 40:73-96.

Carpenter, S. R., N. F. Caraco, D. L. Correll, R. W. Howarth, A. N. Sharpley, and V. H. Smith. 1998. Nonpoint pollution of surface waters with phosphorus and nitrogen. Ecological Applications 8: 559-568.

Clesceri, N. L., S. J. Curran, and R. I. Sedlak. 1986. Nutrient loads in Wisconsin Lakes: part I. Nitrogen and phosphorous export coefficients. Water Resources Bulletin 22:983-989.

Costanza, R., R. d'Arge, R. de Groot, S. Farber, M. Grasso, B. Hannon, K. Limburg, S. Naeem, R. V. O'Neill, J. Paruelo, R.G. Raskin, P. Sutton, and M. van den Belt. 1997. The value of the world's ecosystem services and natural capital. Science 387:253-260.

DeWald, T., R. Horn, R. Greenspun, L. Manning, P. Taylor, and A. Montalbano. 1985. STORET Reach Retrieval Documentation. U.S. Environmental Protection Agency, Washington, D.C.

Dickerhoff Delwiche, L. L. and D. A. Haith. 1983. Loading functions for predicting nutrient losses from complex watersheds. Water Resources Bulletin 19:951-959.

Fisher, T. R., K. Y. Lee, H. Berndt, and M. M. Norton. 1998. Hydrology and chemistry of the Choptank River basin. Water, Air, and Soil Pollution 105:387-397.

Frink, C. R. 1991. Estimating nutrient exports to estuaries. Journal of Environmental Quality 20:717-724.

Hill, A. R. 1979. Denitrification in the nitrogen budget of a river ecosystem. Nature 281:291-292.

Jelinski, D. E. and J. Wu. 1996. The modifiable area unit problem and implications for landscape ecology. Landscape Ecology 11:129-140.

Jordan, T. E., D. L. Correll, and D. E. Weller. 1997. Effects of agriculture on discharges of nutrients from coastal plain watersheds of the Chesapeake Bay. Journal of Environmental Quality 26:836-848.

Leibowitz, S. G., C. Loehle, L. Bai-Lain, and E. M. Preston. 2000. Modeling landscape functions and effects: a network approach. Ecological Modelling, 132:77-94.

Leopold, L. B. and T. Maddock Jr. 1953. The Hydraulic Geometry of Stream Channels and Some Physiographic Implications. US Geological Survey Professional Paper 252, US Government Printing Office, Washington, DC.

Linker, L. C., G. C. Stigall, C. H. Chang, and A. S. Donigian. 1996. Aquatic accounting: Chesapeake Bay watershed model quantifies nutrient loads. Water Environment and Technology 8:48-52.

Lowrance, R. R., R. A. Leonard, L. E. Asmussen, and R. L. Todd. 1985. Nutrient budgets for agricultural watersheds in the southeastern Coastal Plain. Ecology 66:287-296.

Marceau, D. J. 1999. The scale issue in social and natural sciences. Canadian Journal of Remote Sensing 25:347-356.

Paerl, H. W. 1988. Nuisance phytoplankton blooms in coastal, estuarine, and inland waters. Limnology and Oceanography 33:823-847.

Palmer, M. W. 1988. Fractal geometry: a tool for describing spatial patterns in plant communities. Vegetatio 75:91-102.

Panuska, J. C., and R. A. Lillie. 1995. Phosphorus Loadings from Wisconsin Watersheds: Recommended Phosphorus Export Coefficients for Agricultural and Forested Watersheds. Research Management Findings, Number 38. Bureau of Research, Wisconsin Department of Natural Resources, Madison, WI.

Peterson, B. J., W. M. Wollheim, P. J. Mulholland, J. R. Webster, J. L. Meyer, J. L., Tank, E. M. Martí, W. B. Bowden, H. M. Valett, A. E. Hershey, W. H. McDowell, W. K. Dodds, S. K., Hamilton, S. Gregory, and D. D. Morrall. 2001. Control of nitrogen export from watersheds by headwater streams. Science 292:80-90.

Preston, S. D. and J. W. Brakebill. 1999. Application of spatially referenced regression modeling for evaluation of total nitrogen loading in the Chesapeake Bay Watershed. USGS Water Resources Investigation Report 99-4054, U.S. Geological Survey, Reston, VA.

Reckhow, K. H., M. N. Beaulac, and J. T. Simpson. 1980. Modeling Phosphorus Loading and Lake Response under Uncertainty: A Manual and Compilation of Export Coefficients. EPA/440/5-80/011. U.S. Environmental Protection Agency, Washington, D.C.

Schneider, D. C. 1998. Applied scaling theory. Pages 253-270 *in* D. L. Peterson and V. T. Parker, editors. Ecological Scale: Theory and Applications. Columbia University Press, New York.

Schneider, D. C. 2001. The rise of the concept of scale in ecology. Bioscience 51:545-553.

Sjodin, A. L., W. M. Lewis, and J. F. Saunders. 1997. Denitrification as a component of the nitrogen budget for a large plains river. Biogeochemistry 39:327-342.

Smith, R. A., G. E. Schwarz, and R. B. Alexander. 1997. Regional interpretation of water-quality monitoring data. Water Resources Research 33:2781-2798.

Vogelmann, J. E., S. M. Howard, L. Yang, C. R. Larson, B. K. Wylie, and N. V. Direl. 2001. Completion of the 1990s national land cover data set for the conterminous United States from Landsat Thematic Mapper data and ancillary data sources. Photogrammetric Engineering and Remote Sensing 67: 650-662.

Westman, W. E. 1977. How much are nature's services worth? Science 197:960-963.

Wickham, J. D. and T. G. Wade. 2002. Watershed level risk assessment of nitrogen and phosphorous export. Computers and Electronics in Agriculture 37:15-24.

Wickham, J. D., K. H. Riitters, R. V. O'Neill, K. H. Reckhow, T. G. Wade, and K. B. Jones. 2000. Land cover as a framework for assessing risk of water pollution. Journal of the American Water Resources Association 36:1-6.

Wickham, J. D., T. G. Wade, K. H. Riitters, R. V. O'Neill, J. H. Smith, E. R. Smith, K. B. Jones, and A. C. Neale. 2003. Upstream-to-downstream changes in nutrient export risk. Landscape Ecology 18:193-206.

Yuan, L. L., and S. B. Norton. 2003. Comparing response of macroinvertebrate metrics to increasing stress. Journal of the North American Benthological Society 22:308-322.

CHAPTER 13

CAUSES AND CONSEQUENCES OF LAND USE CHANGE IN THE NORTH CAROLINA PIEDMONT:

The Scope of Uncertainty

DEAN L. URBAN, ROBERT I. MCDONALD, EMILY S. MINOR, AND ERIC A. TREML

13.1 INTRODUCTION

The Triangle Landscape Change Project is an on-going effort at regional assessment centered on the Triangle region of North Carolina, a region framed by the cities of Raleigh, Durham, and Chapel Hill. Like many regions of the eastern United States and elsewhere, the Triangle has an agricultural and industrial past, while its current status is defined by high-tech industries of Research Triangle Park, three major universities, and a growing retirement community. The Triangle is one of the fastest-growing regions in the United States, with some portions experiencing 30-50% population growth in the 1990-2000 decade (Triangle J Council of Governments, *public comm.*).

As a case study for the patterns and consequences of land use change, the Triangle is compelling because its period of explosive growth is rather recent and thus coincides with the period of record of satellite imagery. The availability of imagery is augmented by the inclusion of Duke Forest as a NASA SuperSite; specialized imagery flown for the Forest also encompasses much of the larger region. In addition, a wealth of ancillary ground-based data are available (including the Duke Forest data archives, with monitoring data originating in the 1930's), and so there is a rich geospatial data infrastructure to support large-scale studies of landscape pattern and landscape change.

The Triangle Landscape Change Project embraces a set of related research themes under the umbrella of land use/land cover change (Figure 13.1). Land use pattern provides a framework and template in which we are studying various consequences of changing landscape pattern. These themes include forest dynamics, forest bird communities, and watershed impacts. These themes are coupled in that

J. Wu, K.B. Jones, H. Li, and O.L. Loucks (eds.),
Scaling and Uncertainty Analysis in Ecology: Methods and Applications, 239–257.

forests affect watershed hydrology via transpiring and intercepting leaf area as well as via protective ground cover typically associated with intact forests. Forests also provide the template of bird habitat, in terms of forest composition and structure, while land cover provides a larger context via edge effects on nesting success and potential dispersal limitations for habitat patches isolated by human land uses. Coincidentally, forest bird communities are coupled to watershed impacts indirectly because many forests are preserved as riparian buffers and these buffers represent a significant amount of forest habitat for wildlife.

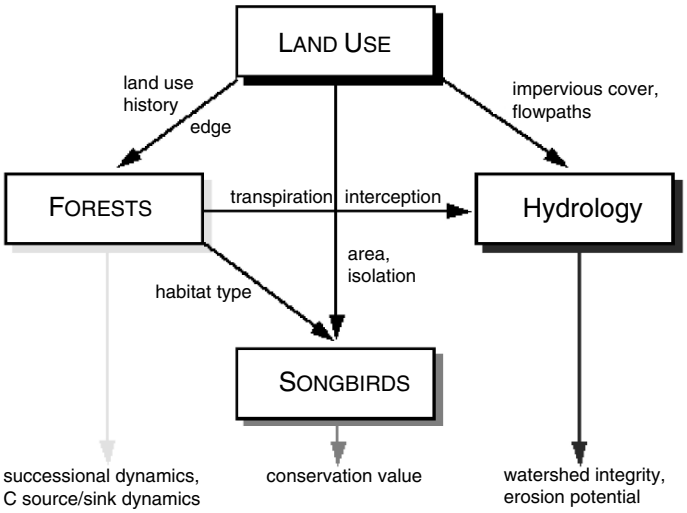


Figure 13.1. *Schematic of linkages among research themes in the Triangle Landscape Change Project.*

We believe that the Triangle Landscape Change Project is typical of many large-scale programs in integrated assessment, which increasingly rely on a shared geospatial data infrastructure and various models to interpolate field data and extrapolate the assessment to the regional scale. Two implications of this approach are that the projects within the larger program tend to be loosely coupled (i.e., studies done by people with different objectives), and that there is no single model that represents the program (i.e., there may be several models). Often, the end-users of the models are not the people who developed the models initially, as illustrated by the increasingly common use of institutionalized models such as Century (Parton et al. 1987), BASINS (US EPA 2001), and other models. These aspects of integrated studies pose some challenges when we attempt to account for uncertainty in the model projections.

Our goal in this chapter is to explore issues related to uncertainty encountered when attempting to conduct integrated, regional-scale assessments using coupled models. Specifically, we will (1) describe sources of uncertainty in scaling to regional applications with increasing reliance on remotely sensed data, and illustrate

how these sources of uncertainty are often "lost in translation" in loosely coupled applications in integrated assessments; and (2) explore methods for propagating these sources of uncertainty and communicating this information to a client audience of fellow scientists as well as city and regional planners.

We will focus on the wood thrush (*Hylocichla mustelina*) as a case study and illustration. The wood thrush is an example of an area-sensitive forest bird species, a species that seems to prefer intact woods and is sensitive to nest predation and brood parasitism by cowbirds (*Molothrus ater*) (Brittingham and Temple 1983, Roth and Johnson 1993, Hoover et al. 1995). Thrushes have recently been exhibiting a regional decline in Triangle-area Breeding Bird Censuses and are consequently a species of some concern (Sauer et al. 2002). For our present purposes, the goal of forecasting regional patterns of abundance of the wood thrush is appealing because this represents quite a stretch for our data (indeed, perhaps the worst possible case), and thus introduces a number of issues related to uncertainty and error propagation in integrated assessment.

13.2 SCALING CONSIDERATIONS AND STRATEGY

Like many of our peers, we are interested in scaling our understanding of ecological processes and patterns at small scales – the scale of field studies or detailed simulation models – to their implications at the much larger scales of resource management and policy decisions (Christensen et al. 1996). The mismatch in scale between science and management has led to a variety of scaling strategies, often aimed at interpolating and extrapolating fine-scale information over larger extent (Peters et al. 2004).

Our approach to this scaling dilemma has been to pursue an explicitly two-scaled approach to ecological modeling. At a fine scale, we develop detailed (often spatially explicit) simulators geared to interact directly with field studies; these models are often developed in a reciprocal iteration between model analysis and model-guided field studies (Urban 2000, Urban et al. 2002). To extend the understanding garnered at fine scales to a much larger spatial extent, we build new models as statistical summaries of the detailed simulators. These new models capture the essential behaviors of the simulators, but at much coarser resolution and consequently, over much larger areas. The statistical models are essentially models of the simulators, or meta-models (Acevedo et al. 1995, Urban et al. 1999). In the case of forest dynamics, the detailed simulators have been forest gap models, while the meta-models have taken various forms including semi-Markovian state transition models, stage-structured matrices, or cellular automata (Urban et al. 1999). In each case, the paired models (detailed simulator plus meta-model) provide a toolkit that allows us to work at either fine or coarse scale, while preserving a common conceptual and parametric framework.

In the case of animal metapopulations, our approach has been conceptually similar although the details vary. Our approach to forecasting regional abundance patterns for wood thrushes and other forest birds entails two linked approaches. Our ultimate goal is to develop macroscopic proxies for metapopulation dynamics, based on graph theory (Urban and Keitt 2001). Graph theory is compelling for such

applications because it can incorporate the richness of species life-history traits (habitat affinities, dispersal behavior) while still being immensely efficient computationally. This latter concern is important because our study area – like many applications in regional scale conservation practice – entails thousands of potential habitat patches and it would be logistically infeasible to address these through field studies or detailed simulation modeling. To infuse as much ecology as possible into this macroscopic approach, however, we are conducting preliminary assessments based on a very detailed individual-based simulator of metapopulation dynamics (modified from Urban and Shugart 1986). Our initial efforts suggest that a graph-theoretic model can capture the essentials of the detailed simulator, while allowing us to work with extremely large and spatially complex landscapes. At issue is the question whether this detailed simulation approach can be defended, given the very real potential that error propagation might overwhelm any insights garnered from the detailed model.

13.3 SOURCES AND SCOPE OF UNCERTAINTY

One conventional definition of sources of uncertainty identifies four components of error in ecological models (reviewed by Gardner and Urban 2003, Peters et al. 2004). These include: (1) *measurement error* (or observation error) associated with the data used to build or parameterize the model; (2) *model error* (or model misspecification) associated with the selection of state equations or the structure of the model; (3) *estimation error* associated with fitting the parameters for the model; and (4) *process error*, due to stochastic processes beyond the scope of the model (e.g., inter-annual variation in climate as an influence on recruitment or survivorship). In practice, these sources are reasonably straightforward to identify for simple models such as regressions, although even in this case estimation error is partially linked to model error and measurement error. As we will show, these definitions are not as satisfying for more complicated simulations, especially in integrated assessments where the total (cumulative) error in one model is subsumed into, or lost from, a coupled application. For example, in our case study there are complexities and associated uncertainty in land cover classification, forest characterization, and habitat modeling that are subsumed into what might be labeled "measurement error" in the metapopulation model. Similar instances will become apparent in the following illustrations.

In general, this introduces the notion of *scope* in defining model uncertainty. By scope, we refer to the tendency for sources of uncertainty or error to be recognized or ignored, depending on the specific focus of any single component of a larger integrated project. Our use of the word *scope* parallels its connotation in computer programming, to the extent that local functions might not be aware of parameters or variables elsewhere in the program; reciprocally, variables internal to a function might be invisible to the larger program. Thus, in our case the uncertainty inherent to image classification is well recognized by the colleague who was primarily responsible for performing these analyses; but these might be lost on a secondary consumer of the results of these analyses. As noted previously, the complicated result of one component of the project ("model output" in its own right) becomes

"input data" for the next person in the chain. As ecological applications increasingly are integrated efforts by large research teams, this potential for error propagation within and across projects is an especially compelling technical issue.

Given the rather fuzzy definitions of sources of uncertainty, it might be equally useful to distinguish sources that are *inside* as compared to *outside* a model. For example, once implemented as code, a model is reasonably well contained, with its assumptions and algorithms known, hence the uncertainty or potential error stemming from these is also knowable. By contrast, data used as initial or bounding conditions (e.g., a habitat map in the case of a population model) or for parameterization (e.g., demographic rates) can be gathered or estimated independent of the model itself (Botkin 1993). These data have their own intrinsic uncertainty (e.g., measurement or estimation error), but they might also induce further error as they interact with the model. In a sense, the distinction of whether a source is inside or outside a model also defines its *accessibility* to a client user of the model. For example, the specification of a dispersal algorithm in a metapopulation model might be well described, but a client user rarely has the capability to alter the algorithm to assess the uncertainty associated with that or an alternative algorithm; the end-user can only manipulate data and inputs external to the model. This is important because the conventional framework for uncertainty analysis (reviewed below) can only be extended to parameters or model components that can be freely varied by the modeler, and for which the error distribution (or range of variability) can be estimated.

One common approach to model uncertainty and sensitivity analysis uses regression as the framework (Gardner 1984, Gardner et al. 1981, Haefner 1996; also see Li and Wu, Chapter 3). In this, the model is driven by a set of input parameters x, each element x_i of which has an error distribution (typically presumed normal). The model is run in Monte Carlo fashion and, for each iteration, a stochastic set of input parameters is generated by sampling from the error distribution of each parameter in turn (i.e., drawing a random parameter value from the mean ± 1 SE). For each iteration, a selected output variable is retained along with the input parameters for that run. This is repeated for a large number of runs. The analysis consists of regressing the output variable on the input parameters. A parameter's uncertainty is indexed by its partial explanatory power in the regression (partial R^2), a direct measure of the extent to which uncertainty in the parameter maps onto uncertainty (variability) in model output. For the sake of clarity, note that this same approach is used for sensitivity analysis, but in this case each parameter is perturbed randomly by some arbitrary amount – say 10% of its nominal value. A parameter's sensitivity is indexed as its (standardized) partial regression slope: a measure of how much model output changes given a slight change in the input parameter. A parameter can have high uncertainty only if it has high sensitivity relative to its estimation error.

Clearly this regression approach to model uncertainty is awkward for model inputs that cannot be provided as a mean and standard error. It is difficult enough for Boolean or categorical variables (e.g., open- versus cavity-nesting bird species); for elements such as dispersal algorithms, the approach must be modified so that these

inputs can be assessed. This is not complicated – it merely requires that the framework be relaxed somewhat – and it still relies on Monte Carlo simulations using a variety of model configurations.

13.4 CASE STUDY: FORECASTING WOOD THRUSH ABUNDANCE PATTERNS

As part of an exercise in forecast evaluation, Minor et al. (unpublished manuscript) attempted to assess the implications of various sources of uncertainty on the precision of predictions made with an individual-based metapopulation simulator. The illustrations provided here are extracted from their larger analysis. The basic steps involved in forecasting wood thrush abundance patterns are intuitively straightforward: (1) classify land cover for the region, masking out nonforest habitats; (2) predict forest stature and gross composition to aid in predicting potential wood thrush habitat; (3) classify potential wood thrush habitat (i.e., "habitat" versus "nonhabitat"); and (4) simulate thrush metapopulation dynamics for this habitat mosaic.

As we shall illustrate, these simple steps invite a frustrating variety of potential sources of error or uncertainty. Our task is to identify and isolate these sources of uncertainty.

In this illustration, we hold to two presumptions: (1) the wood thrush project is loosely coupled to other components of the larger research agenda (i.e., this application is not conducted simultaneously with other tasks, and some information about uncertainty is lost in the chain of custody), and (2) the metapopulation model itself is not accessible to the end-user (i.e., we will not alter the code, and will need to frame the analysis in terms of elements accessible through parameterization). For purposes of illustration, we focus here on four components of uncertainty selected to represent the range of these sources and their interactions. First, we will consider alternative definitions of potential wood thrush habitat. This habitat map represents a series of analyses and models, but is provided to the metapopulation model as boundary condition "data." As model error, we will accept the gross structure of the metapopulation model as plausible, and focus instead on the implementation of bird dispersal between habitat patches. As one source of estimation error, we will consider the impacts of the uncertainty associated with our best estimate of wood thrush clutch size. We also will consider the impact of edge effects on nesting success, in effect an influence on net fecundity. Maximum dispersal range represents the third source of estimation error. Finally, we will consider process error as the sum of the main stochastic processes in the simulator: variation in clutch size, mortality, and dispersal (each implemented on a per-bird, per-event basis). We now consider the four steps to forecasting thrush populations in turn.

13.4.1 Land Cover Classification

We have compiled a time series of anniversary-dated winter/summer pairs of Landsat Thematic Mapper imagery. The images span the years 1986-2001 on

roughly a 3-year interval, depending on the availability of high-quality (cloud-free) images. Following radiometric and geometric rectification, the images were subjected to a supervised maximum-likelihood classification using high-resolution digital airphotos to identify training samples. Because the distribution of spectral values did not meet assumptions of multivariate normality, the classification was conducted using log-transformed spectral values. The classification was collapsed into 7 land cover types: (1) developed, (2) deciduous forest, (3) evergreen (pine) forest, (4) mixed forest, (5) sparse vegetation (agriculture, lawns), (6) shallow water, and (7) deep water. Because our land use change model (under development) considers land cover at the scale of a pixel (30-m cell) within the context of land use defined at the scale of the parcel, this simple classification scheme is sufficient for our purposes. The classification is reasonably robust; more importantly for our purposes, the classification provides posterior probabilities of membership in each land cover class, for each pixel of the image. Thus, we have direct estimates of the uncertainty of the land cover classification – uncertainties that might be inherited by subsequent applications that make use of the classified land cover maps. In this illustration, however, the applications are loosely coupled and what is conveyed from the land-cover classification project is simply a land cover map for further processing – the details about classification error are beyond the scope of the next stage.

13.4.2 Forest Stature and Composition

Using the land cover as a generous mask, we then predicted gross forest composition and structure as basal area of hardwoods and pines. In this, we used the winter/summer difference in greenness to separate deciduous hardwoods from evergreens (almost entirely pines in this region). Basal area estimates were derived from long-term sample quadrats archived in georeferenced form in the Duke Forest database. Basal area of hardwoods and pines was regressed separately on spectral values. The regressions were highly significant ($R^2 = 0.54$, $P < 0.001$ and $R^2 = 0.73$, $P < 0.001$, respectively). Importantly, because these predictions were by regression we can retain the prediction error for each component of forest stature. Thus, it is possible to map not only the regional extrapolation of hardwood basal area, but also the associated uncertainty. Again, however, these detailed measures of uncertainty are lost in translation; we have access to predicted maps of forest stature (basal area) for pine and hardwood components.

13.4.3 Potential Wood Thrush Habitat Classification

We should confess at this point that we have very limited field data on the habitat affinities of the wood thrush in our study area. Das (2000) used local census data and habitat measurements to attempt to discern separate effects of microhabitat (size class distribution, species composition) and landscape context (distance to edge, amount of development in the neighborhood), but small sample sizes rendered many tests nonsignificant. While we are collecting new data to build more reliable habitat

models, we have created a somewhat arbitrary habitat model for use in this illustration. The habitat model is based on predicted basal area, thresholded at values consistent with the limited field observations available. Thus, the uncertainty associated with the regressions is lost from this assignment of potential habitat, beyond the scope of this stage of the project. To provide a contrasting range of habitats, we used two thresholds. We constructed a "generous" habitat map by using 15 m^2 ha^{-1} of hardwood basal area to define "habitat" (Figure 13.2) and made a "strict" habitat map by using 20 m^2 ha^{-1} as the second threshold (Figure 13.2). The strict habitat map included 306 discrete patches (defined using an 8-neighbor rule) with a total area of 794 ha, while the generous map included 823 patches and 3230 ha of habitat. We should emphasize that these maps are intended only to illustrate the magnitude of uncertainty that might arise from uncertainty in habitat classification; the maps themselves should not be over-interpreted in terms of thrush habitat. In particular, it is reasonable to assume that we have missed significant predictive power about thrush habitat because we cannot remotely sense understory density in these forests, a component of habitat quality that is probably important to the wood thrush based on our field observations.

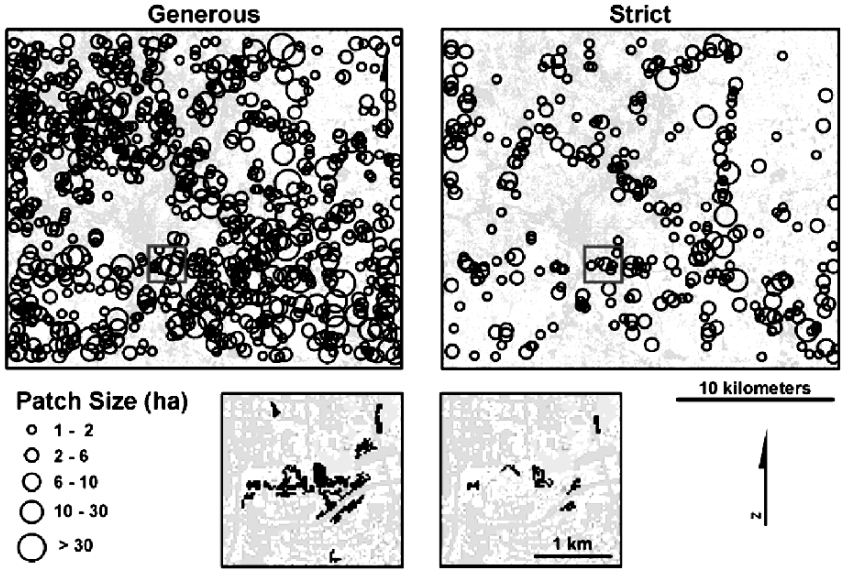


Figure 13.2. *Contrasting maps of potential wood thrush habitat, defined generously (left) and strictly (right) in terms of hardwood basal area. Because the shaded clusters of cells are difficult to resolve, patches are overlaid with circles indicating their relative sizes. Insets are expanded to highlight differences in habitat definition, below; note the generous map has more patches and patches tend to be larger.*

As with previous steps of this integrated application, issues arise that are somewhat logistical but with significant implications to the application at hand. In

mapping potential habitat, one logistical constraint is that there is a potentially overwhelming number of habitat patches if every pixel classified as "habitat" is actually retained (see Keitt et al. 1997 for a similar dilemma). In practice, some minimum patch size is selected, and smaller patches are discarded and ignored. In this case, the distribution of patch sizes is roughly negative-exponential, so the number of patches increases dramatically as the minimum patch size is reduced. For our purposes, we have retained patches larger than 1 ha (a rough estimate of territory size for the thrush; Roth et al. 1996). But this decision invites three sources of error: (1) some actual habitat is discarded and, hence, some potential thrushes are not included; (2) the apparent distances between patches are increased by removing interstitial habitat, and so landscape connectivity is decreased; and (3) if small patches are sinks (*sensu* Pulliam 1988), then their removal actually *improves* the habitat mosaic and thus produces biases predictions about the metapopulation. Importantly, each of these sources of error or bias is largely unaccountable after the decision is made to discard small patches.

13.4.4 The Metapopulation Model METAPOP1

Model overview. The metapopulation model is an individual-based simulator that tracks male birds in each patch of habitat mosaic (Urban and Shugart 1986). This version simulates a single species, although a multi-species version also exists (Urban et al. 1988). This is an "island" model, meaning that the landscape is partitioned into discrete "habitat" patches and a "nonhabitat" matrix. Each patch has a carrying capacity based on its area of preferred habitat and territory size for the species being simulated. Territorial breeders are distinguished from nonterritorial "floaters" that are assumed to occupy marginal habitats. The model works on an annual time step. Each year, birds are subjected to stochastic over-winter mortality with a probability based on expected longevity. Adults have age-independent survivorship, and juveniles (young of the year) and nonterritorial floaters have higher mortality rates than territorial adults. Survivors that occupy territories are then allowed to nest. The species has a mean clutch size (number of eggs per nest) with a standard deviation estimated from field studies. Actual clutch size is generated stochastically for each nesting attempt by each bird. The species may nest once or more per season (defined by the mean number of broods per year). For each nesting attempt, the brood may be subjected to stochastic nest predation and brood parasitism by cowbirds. Predated nests are lost entirely, while parasitism costs one fledgling of the host. Both processes are implemented as edge effects, according to functions defined in terms of the maximum rate in edges and the distance from the edge to which these processes extend into the forest. Rates and edge widths are estimated from field studies – which, we should note, vary drastically among field studies (Paton 1994, Lahti 2001). Juveniles successfully fledged are added to the pool of floaters for that patch, and then birds are dispersed.

Dispersal is implemented as the product of two species life-history parameters. A bird's dispersal *range* defines the maximum distance that it will disperse in a single episode (move). A species may move multiple times during dispersal, with the number of moves defined by its *mobility*. At each move, a bird disperses

probabilistically to a new patch. If the patch has unoccupied habitat available, the bird settles; otherwise, it moves again. This continues until the bird settles in habitat or all moves are used. Birds that do not find habitat persist as floaters in the last patch they sampled. Dispersal is also modified by site fidelity. Breeders have very high site fidelity; floaters have low fidelity. Thus, breeders tend to occupy the same site for their entire adult life, while floaters (including juveniles) tend to disperse to find new habitats. This algorithm represents a crude but efficient model of our understanding of bird dispersal (Greenwood et al. 1980, Greenwood and Harvey 1982).

At the end of each simulation year, the model updates censuses of each habitat patch and writes a variety of output statistics describing populations for each patch and for the entire landscape mosaic. The simulations are run in replicate, and the replicates are further summarized to provide means and standard deviations of patch-level and landscape-level populations.

13.4.4.1 Parameterization

The model requires a total of 20 parameters, 5 of which control a set of replicate simulations and 15 of which are species life-history parameters (Table 13.1). Of the 15 life-history parameters, four are essentially set as constants for any species, so 11 parameters must be estimated for a focal species. Previous sensitivity analyses identified clutch size and survivorship as being quite sensitive, while dispersal parameters were less so (at least in the landscapes simulated, which seemed reasonably well connected; Urban et al. 1988 and unpublished model analyses; see also Pulliam et al. 1992). For the case study illustrated here, we emphasize clutch size as a source of parameter uncertainty. This is not to deny that other parameters are important; rather, our point can be illustrated readily with clutch size.

As an additional factor related to model parameterization, we also considered nest parasitism and brood parasitism as edge effects. In this, we set the range (functional edge width) and intensity (rate or probability) of these effects as constants, and then simply toggled these effects on or off in particular simulations. These effects are essentially reductions in fecundity and would be equivalent to a commensurate reduction in clutch size, but because they occur near forest edges they have the potential to have a local rather than a global impact on model behavior. We address the influence of dispersal parameters on model uncertainty explicitly in the following section.

13.4.4.2 Alternative dispersal models

Dispersal is difficult to observe in most cases, and for metapopulations it is logistically infeasible to expect to observe dispersal sufficiently to describe the process adequately from data. For birds, there is a wide range of opinions about how dispersal operates, leading to a variety of algorithms as implemented in population models. Our implementation is intended to be quite simple. Compared to models that attempt to simulate dispersal *behavior* (e.g., McKelvey et al. 1993, Gustafson

and Gardner 1996), our model simulates the *result* of dispersal. In this, a bird has a probability of dispersal from patch *i* to patch *j* that depends only on the distance between the patches, d_{ij}, and the size of the target patch *j*. (Because the model simulates birds individually, there is also an effect of the size of the donor patch *i* due to the likelihood that larger patches tend to support more birds.) We model dispersal probability as a negative-exponential function of distance:

$$p_{ij} = exp(-\theta d_{ij}) \cdot a_j \tag{13.1}$$

where θ is an extinction coefficient estimated from the tail distance at which $p_{ij} = 0.01$, and a_j is the area of the target patch. Even this simple approach can become complicated by the way that the distances d_{ij} are defined. For purposes of illustration, we compare two such alternatives.

Table 13.1. *Parameters used in the metapopulation simulator.*

Parameter	*Definition*	*Value*[1]	*Assignment*[2]
PCCI	% of carrying capacity initialized	50%	constant
XNP	Max rate nest predation	65%	Input/constant[3]
DENP	Edge distance, predation	100 m	Input/constant[3]
XBP	Max rate, parasitism	75%	Input/constant[3]
DEBP	Edge distance, parasitism	200 m	Input/constant[3]
TS	Territory size	1 ha	input
MBD	Max breeding density	50 per 100 ha	input
Tsmin	Min occupiable territory	0.50 ha	constant
Xclutch	Mean clutch size	2.5 eggs	input
SClutch	Std Dev (clutch size)	0.5	input
NBroods	Broods per season	2	input
NT	Nest type (open/cavity)	open	input
NHt	Nest height (1=ground, 3=canopy)	2 (midcanopy)	input
XSurv	Annual survivorship	0.65	input
RSFlt	Relative survivorship of floaters	0.50	constant
SFBrd	Site fidelity of breeders	0.90	constant
SFFlt	Site fidelity of floaters	0.10	constant
Range	Dispersal range per move	1500 m	input
Mobil	Number of dispersal moves	3	input

[1]*Nominal values set for model experiments, based on literature values (detailed in Urban and Shugart 1986, Urban et al. 1988, Minor et al., unpublished manuscript).*

[2]*Values are set as: constants (typically not varied by end-user); inputs for edge effects (defined for the geographic study area and which do not vary by species); or species-level inputs, (defined for each focal species to be simulated).*

[3]*Inputs for edge effects (constant for study area).*

The first alternative estimates between-patch distances as the minimum edge-to-edge distance between the two patches. In practice, this is done by finding the minimum cell-to-cell distance between any two cells, one from each patch, and saving this as the edge-to-edge distance. This is computationally straightforward (though the recursion is sufficiently tedious in a GIS that it is easier to do the computations in a stand-alone program). For very large landscapes (N>>1000's of patches) this approach can become computationally daunting. This approach presumes implicitly that dispersing animals are sufficiently clever to find and use this minimum-distance path.

The second alternative estimates between-patch distances as least-cost paths, using optimal routing algorithms in a GIS (Bunn et al. 2000). This requires the (ultimately arbitrary) assignment of relative resistances to dispersal ("costs") for all cover types within a landscape. The routing algorithm then finds the "cheapest" path from the donor patch to the target patch. The least-cost patch method can be computationally infeasible for large sets of patches, because the analysis amounts to performing the routing solution recursively in the GIS. Various GIS-based algorithms have been devised and are available from websites such as ESRI's (for Arc/Info). Our approach has been to write a custom macro in Arc/Info by which we find least-cost paths for pairs of patches that are within a minimum Euclidean distance of each other (e.g., twice the dispersal range of the focal species), and substituting simple edge-to-edge distances for patches farther apart. This saves considerable computation time and, because dispersal between distant patches occurs only via stepping-stone paths, this does not affect the traversability of the mosaic. Note that least-cost paths assume that the dispersing organisms are actually quite clever, that is, they sample cover types locally and tend to find the easiest route from patch to patch.

This implementation of dispersal distances essentially uses different parameterizations to assess contrasting conceptual models of how birds disperse. In the case of either definition of dispersal distances, dispersal probabilities (Equation 13.1) are re-normalized in the model to account for the proximity of habitat patches (the raw probabilities typically sum to >> 1.0).

Clearly, there is a world of complexity available to us in implementing alternative dispersal models. This decision clearly influences overall model uncertainty in that the choice of dispersal model also dictates the kinds of parameters needed to implement it. This uncertainty is added to that due to the definition of habitat patches, as discussed above. Further, in many simulators, the probability of mortality increases as individuals disperse (e.g., the models described by McKelvey et al. 1993), and so assumptions about dispersal might also propagate through demographic rates or indeed to habitat definitions (Anders et al. 1998). We do not pretend to cover this full range of issues, but instead focus on two aspects of dispersal: (1) the implications of habitat resistance as implemented as least-cost paths, and (2) maximum dispersal range for the focal species. Note that because of underlying land use pattern, using least-cost paths will tend to have local rather than global effects in the model.

13.4.5 Experimental Design

As an illustration of how uncertainty can propagate through coupled regional assessments, Minor et al. (unpublished manuscript) contrived a set of model experiments to include several sources of uncertainty. In this, they considered a total of 32 cases, including:

- 2 alternative habitat maps ("strict" versus "generous"),
- 2 dispersal models (Euclidean versus least-cost distance),
- 2 dispersal ranges (1500 versus 3000 m),
- 2 clutch sizes (2.5 versus 3 eggs/clutch), and
- 2 levels of intensity of edge effects on nesting success (on/off).

Each simulation was initialized with a population at 50% of carrying capacity for the landscape, and with the initial birds randomly distributed among patches. We simulated 100 years of population dynamics, by which time populations either stabilized or the trend was well established. We ran 100 replicate simulates for each case, in which stochastic processes (dispersal, mortality, and nesting success) were free to vary on a per-bird, per-event basis. Total error or uncertainty in the forecasts is the result of all of these sources.

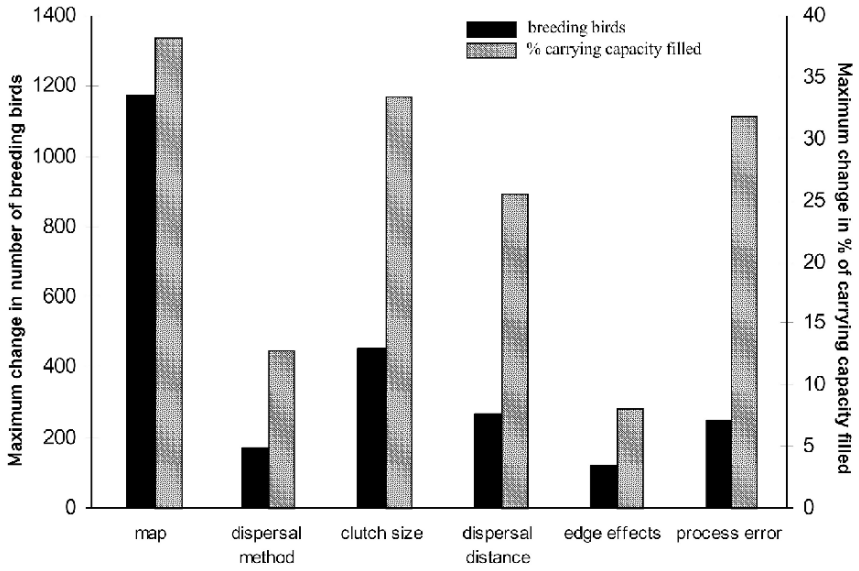


Figure 13.3. *Relative importance of each simulated factor as a source of uncertainty in the metapopulation model, based on ANOVA. Analyses were conducted separately for total population size (left-side axis) and percent of carrying capacity (right-side axis) because the two habitat maps had very different amounts of habitat. Sources of uncertainty are arrayed as input data (left), model, estimation, and process error (right).*

We analyzed the model simulations by ANOVA, using the experimental treatments as main effects. We should note that, of these factors, only clutch size and dispersal range would be easily amenable to implementation in the conventional regression framework for uncertainty analysis; that is, only these two parameters have estimates and standard errors. Process error, i.e., the result of stochastic implementations in the model, appears in the analysis not as a main effect but rather as within-treatment (replicate) variability.

13.4.6. Relative and Cumulative Effects of Uncertainty

With this range of parameters, simulations resulted in populations that ranged in size from 20-87% of carrying capacity. In the ANOVA, each main factor had a significant effect. When total population size was analyzed as the response variable, the choice of habitat map had the largest effect as expected, since the generous map includes substantially more habitat. When percent of carrying capacity was analyzed instead (to remove the effect of total habitat area), the rank order of factors was very similar (Figure 13.3).

Almost all first-order interaction effects (7 of 10) were significant as well in this analysis (Table 13.2). In particular, there was a strong interaction between choice of habitat map and dispersal factors (Euclidean/least-cost path distances as well as dispersal range). Clutch size also showed an interaction with the habitat map and with edge effects.

Table 13.2. *Results of ANOVA on sources of uncertainty in the metapopulation simulator, using percent of carrying capacity as the population response variable.*

Source of variation	*DF*	*Sum of Square*	*Mean Square*	*F value*	*Pr (F)*
Dispersal model	1	5.68	5.68	3476.25	0.00
Map	1	65.57	65.57	40150.67	0.00
Clutch size	1	60.44	60.44	37006.97	0.00
Dispersal distance	1	19.80	19.80	12121.84	0.00
Edge effects	1	1.28	1.28	786.29	0.00
Dispersal model × map	1	0.41	0.41	251.30	0.00
Dispersal model × clutch size	1	0.00	0.00	0.80	0.37
Dispersal model × dispersal distance	1	0.08	0.08	51.09	0.00
Dispersal model × edge effects	1	0.00	0.00	0.86	0.35
Map × clutch size	1	1.60	1.60	979.97	0.00
Map × dispersal distance	1	2.98	2.98	1827.02	0.00
Map × edge effects	1	0.06	0.06	34.09	0.00
Clutch size × dispersal distance	1	0.07	0.07	43.23	0.00
Clutch size × edge effects	1	0.27	0.27	166.55	0.00
Dispersal distance × edge effects	1	0.00	0.00	0.62	0.43

To illustrate the uncertainty associated with a simulation model, the convention is to use Monte Carlo methods to propagate uncertainty through the model. This approach is readily extended to include process error; indeed, this is routinely done with stochastic simulators. Likewise, it is straightforward to propagate estimation error through a model, by using stochastic parameter sets from specified distributions of the parameters. This approach becomes somewhat unwieldy when we attempt to extend it to include model error and data error of a form such as an input habitat map. In the simplest case for our metapopulation model, the simulations would be run in Monte Carlo fashion, and for each run a new set of input map, dispersal model, and parameter set would be drawn from a set of alternatives; process error would be included by default because the model is stochastic. But because the initial habitat maps and (especially) the dispersal-distance matrix are quite demanding computationally, this implies building a set of alternatives beforehand and then selecting from these for each simulation. Again, this is quite feasible in principle but rather tedious in practice. We have illustrated this approach by randomly selecting a set of 100 simulations from the full set (3200) used in our model experiment. From this set, we computed the range of population sizes as percent of carrying capacity. Because this range is not normally distributed (being bounded by 0 and carrying capacity), we index the variation as simply the central 95% quantiles of the data. For comparison, we also illustrate the amount of process error for a simulation with nominal parameter values, based on 100 replicate simulations. The differences are rather telling (Figure 13.4).

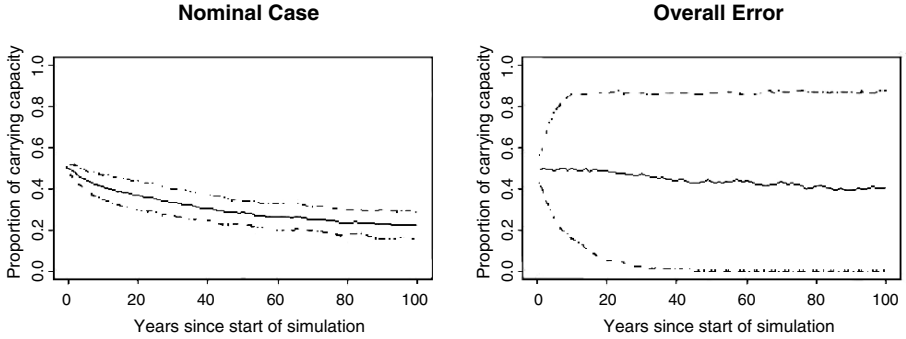


Figure 13.4. *Illustration of total uncertainty as propagated through the model, compared to the confidence limits generated by stochastic "process error" alone.*

We should emphasize that we do not propose Figure 13.4 as the true uncertainty in this modeled scenario. Rather, we offer this as an illustration of the potential magnitude of the implications of ignoring sources of error that are beyond the scope of conventional uncertainty analysis. In this case, the uncertainty associated with the underlying habitat map far outweighs the influence of details internal to the metapopulation model. Yet, even this simple illustration ignores potentially important implications of more extreme alternatives to the dispersal algorithm, as well as multiple combinations of model specifications and parameter estimation

error. On the other hand, we can also take some solace in that all sources of error are not equally likely in this case, and so if we bounded the selection of alternatives by some notion of their likelihood (i.e., via prior probabilities), we could reduce the extreme levels of uncertainty shown in Figure 13.4. For example, we do not think that the 1500 or 3000 m dispersal range is equally likely, and we could constrain our estimate of clutch size more tightly by using the available data more carefully. The use of prior probabilities to constrain error estimates in this way is only a slight generalization of the conventional approach of drawing parameter estimates from their empirical distributions. This approach is especially amenable to hierarchical Bayesian approaches to modeling (e.g., Wikle et al. 1998, Wikle 2003, Clark 2003), in which each component of the model has its prior and (fitted) posterior distributions.

13.4.7 A Patch-Level Perspective

Edge effects and differences in dispersal mediated by the "resistance" of local land cover should result in local rather than global influences on population dynamics. One example of these is illustrated in Figure 13.5, which shows the local persistence (years occupied of the last 10 years of a simulation) of selected patches. In this case the differences in local persistence are mediated by the local prevalence of developed lands, which reroute dispersal locally and thus change the pattern of dispersal subsidy that is key to patch recolonization after a chance local extinction. Because dispersal in this model is largely via short dispersal events among stepping-stone patches, local dispersal limitations have the potential to propagate within connected subregions of the habitat mosaic. This spatial error propagation would lead to strongly autocorrelated errors in model predictions.

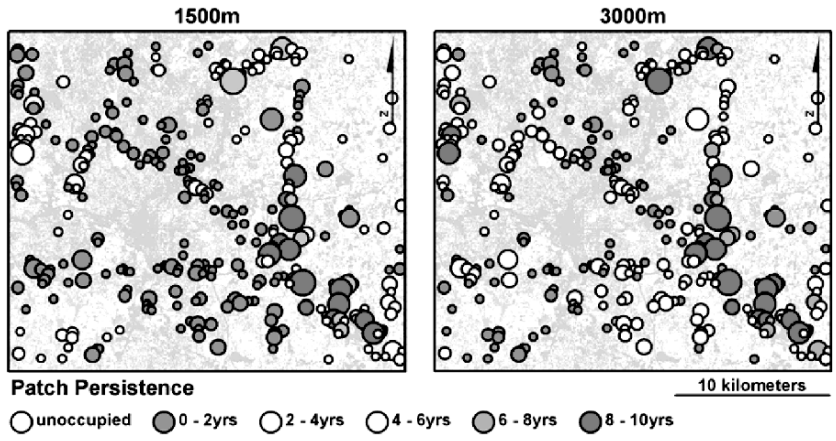


Figure 13.5. *Local effects of model uncertainty, as generated by variation in dispersal distance as modeled using least-cost paths as dispersal routes. Circles are overlaid on habitat clusters to indicate their relative size (see Figure 13.2). Effects of connectivity on population persistence are expressed locally.*

We would expect edge effects on recruitment to have a similarly local influence on population dynamics, if habitat "edginess" varied locally within the study area. That is, if patches in one region of the study area were especially edgy, then there would be some potential for strong local patterns in populations to emerge. As discussed previously, this also would invite a local interaction with dispersal, so that local source and sink patches could have an influence that propagated to other nearby patches to which they were strongly connected. The lack of a dominant edge effect in our simulations probably reflects the reality that most of the patches are equivalently edgy.

While we have not yet explored these local influences rigorously, the illustration in Figure 13.5 does provide an immediate aid to this future effort. Patches or regions that show a strong local effect as indicated by high patch-level uncertainty, clearly present themselves as compelling candidates as focal sites for follow-up field studies. This is an example of model-guided sampling design, a powerful approach for locating study sites that can provide crucial information efficiently (Urban 2000, 2002, Urban et al. 2002). In this case, the candidate sites are those exhibiting the strongest manifestation of key model uncertainties. This model-data dialogue is an added benefit of the approach to uncertainty analysis that we illustrate here: the modeling process is self-correcting if the approach admits iteration between model analysis and model-directed field studies.

13.5 CONCLUSIONS

Ecological forecasts extrapolated to regional scales invite a variety of sources of uncertainty. Performing integrated assessments across coupled applications (land use change, forest dynamics, bird communities, and watershed impacts) invites new sources that are problematic because they may not conform readily to conventional approaches to uncertainty analysis and error propagation. Moreover, if the separate components of such assessments are pursued by different research teams, participants or clients might not even be aware of these sources of uncertainty. At the least, the role of uncertainty may change dramatically across coupled project components with different specific objectives. For example, classification errors in the land cover maps might have very different implications for watershed hydrology than for thrush metapopulations. Because a client audience deserves – indeed, may demand – a full accounting of uncertainty in ecological predictions (Clark et al. 2001), it is crucial that we devise thorough but efficient methods for incorporating uncertainty in integrated regional assessments.

While an efficient analytic approach to uncertainty in integrated assessments remains a challenge, it is already quite feasible to *communicate* uncertainty in ecological extrapolations and forecasts. We do this routinely by adding error bars or confidence limits to histograms and line graphs; there is no reason why we should not include this information in maps as well. For example, we have estimated both classification error and prediction error for model predictions here, and these approaches are increasingly easy with the powerful cartographic tools available in geographic information systems. Further, we know how to translate model sensitivity and uncertainty from parameter space to geographic space (e.g., Urban

2000), and wider use of this approach can lead to a healthy model-data dialogue in which model-guided field studies are used to collect new data to improve the model most efficiently.

There remains a technical challenge of finding computationally efficient methods for incorporating a range of sources of uncertainty into simulations. This may entail generalized methods for Monte Carlo simulations, or alternative formalisms for model development (e.g., hierarchical Bayesian models). It is unlikely that we will find a convenient "one size fits all" solution, and so a healthy variety of approaches should be pursued.

REFERENCES

Acevedo, M. F., D. L. Urban, and M. Ablan. 1995. Transition and gap models of forest dynamics. Ecological Applications 5:1040-1055.

Anders, A. D., J. Faaborg, and F. R. Thompson. 1998. Postfledging dispersal, habitat use, and home-range size of juvenile wood thrushes. Auk 115:349-358

Botkin, D. B. 1993. Forest Dynamics: An Ecological Model. Oxford University Press, New York.

Brittingham, M. C., and S. A. Temple. 1983. Have cowbirds caused forest songbirds to decline? Bioscience 33:31-35.

Bunn, A. G., D. L. Urban, and T. H. Keitt. 2000. Landscape connectivity: a focal-species analysis using graph theory. Journal of Environmental Management 59:265-278.

Christensen, N. L., A. M. Bartuska, J. H. Brown, S. Carpenter, C. D'Antonio, R. Francis, J. F. Franklin, J. A. MacMahon, R. F. Noss, D. J. Parsons, C. H. Peterson, M. G. Turner, and R. G. Woodmansee. 1996. The report of the Ecological Society of America Committee on the scientific basis for ecosystem management. Ecological Applications 6:665-691.

Clark, J. S. 2003. Uncertainty and variability in demography and population growth: a hierarchical approach. Ecology 84:1370-1381.

Clark, J. S., S. Carpenter, M. Berber, S. Collins, A. Dobson, J. A. Foley, D. M. Lodge, M. Pascual, R. Pielke, W. Pizer, C. Pringle, W. V. Reid, K. A. Rose, O. E. Sala, W. H. Schlesinger, D. H. Wall, and D. Wear. 2001. Ecological forecasts: an emerging imperative. Science 293:657-660.

Das, A. A. 2000. Stand and Landscape-Level Effects on the Distribution and Abundance of Breeding Bird Species: A Case Study of the Duke Forest. Master's thesis. Duke University, Durham, NC.

Environmental Protection Agency (US EPA). 2001. Better assessment science integrating point and nonpoint sources: BASINS version 3.0. http://www.epa.gov/watersciences/BASINS.

Gardner, R. H. 1984. A unified approach to sensitivity and uncertainty analysis. *in* M. H. Hamza, editor. Applied Simulation and Modelling, Proceedings of the IASTED International Symposium. ACTA Press, Anaheim, CA.

Gardner, R. H., and D. L. Urban. 2003. Model testing and validation: past lessons and present challenges. Pages 184-203 *in* C. D. Canham, J. J. Cole, and W. K. Laurenroth, editors. Models in Ecosystem Science. Princeton University Press, Princeton.

Gardner, R. H., R. V. O'Neill, J. B. Mankin, and D. Kumar. 1981. Comparative error analysis of six predator-prey models. Ecology 61:323-332.

Greenwood, P. J. 1980. Mating systems, philopatry, and dispersal in birds and mammals. Animal Behavior 28:1140-1162.

Greenwood, P. J., and P. H. Harvey. 1982. The natal and breeding dispersal of birds. Annual Review of Ecology and Systematics 13:1-21.

Gustafson, E. J., and R. H. Gardner. 1996. The effect of landscape heterogeneity on the probability of patch colonization. Ecology 77:94-107.

Haefner, J. W. 1996. Modeling Biological Systems: Principles and Applications. Chapman & Hall, New York.

Hoover, J. P., M. C. Brittingham, and L. J. Goorich. 1995. Effects of forest patch size on nesting success of wood thrushes. Auk 112:146-155.

Keitt, T. H., D. L. Urban, and B. T. Milne. 1997. Managing fragmented landscapes: a macroscopic approach. Conservation Biology 1:4.

Lahti, D. C. 2001. The "edge effect on nest predation" hypothesis after twenty years. Biological Conservation 99:365-374.

McKelvey, K., B. R. Noon, and R. H. Lamberson. 1993. Conservation planning for species occupying fragmented landscapes: the case of the Northern potted Owl. Pages 424-450 *in* P. Kareiva, J. G. Kingsolver, and R. B. Huey, editors. Biotic Interactions and Global Change. Sinauer, Sunderland, Mass.

Minor, E. S., R. I. McDonald, E. A. Treml, and D. L. Urban. (unpublished manuscript). Uncertainty in spatially explicit metapopulation models.

Parton, W. J., D. S. Schimel, C. V. Cole, and D. S. Ojima. 1987. Analysis of factors controlling soil organic matter levels in Great Plains grasslands. Soil Science Society of America Journal 51: 1173-1179.

Paton, P. W. C. 1994. The effect of edge on avian nest success: how strong is the evidence? Conservation Biology 8:17-26.

Peters, D. P. C., J. E. Herrick, D. L. Urban, R. H. Gardner, and D. D. Breshears. 2004. Strategies for ecological extrapolation. Oikos 106:627-636.

Pulliam, H. R. 1988. Sources, sinks, and population regulation. American Naturalist 132:652-661.

Pulliam, H. R., J. B. Dunning, and J. Liu. 1992. Population dynamics in complex landscapes: a case study. Ecological Applications 2:165-177.

Roth, R. R., and R. K. Johnson. 1993. Long-term dynamics of a wood thrush population breeding in a forest fragment. Auk 110:37-48.

Roth, R. R., M. S. Johnson, and T. J. Underwood. 1996. The birds on North America. No. 246 *in* A. Poole and F. Gill, editors. The Birds of North America. Philadelphia, PA.

Sauer, J. R., J. E. Hines, and J. Fallon. 2002. The North American Breeding Bird Survey, Results and Analysis 1966-2001. Version 2002.1. USGS Patuxent Wildlife Research Center, Laurel, MD.

Urban, D. L. 2000. Using model analysis to design monitoring programs for landscape management and impact assessment. Ecological Applications 10:1820-1832.

Urban, D. L. 2002. Strategic monitoring of landscapes for natural resource management. Pages 294-311 *in* J. L. Liu and W. W. Taylor, editors. Integrating Landscape Ecology into Natural Resource Management. Cambridge University Press, Cambridge.

Urban, D. L., and T. H. Keitt. 2001. Landscape connectivity: a graph-theoretic perspective. Ecology 82:1205-1218.

Urban, D. L., and H. H. Shugart. 1986. Avian demography in mosaic landscapes: modeling paradigm and preliminary results. Pages 273-279 *in* J. Verner, M. L. Morrison, and C. J. Ralph, editors. Wildlife 2000: Modeling Habitat Relationships of Terrestrial Vertebrates. The University of Wisconsin Press, Madison, WI.

Urban, D. L., M. F. Acevedo, and S. L. Garman. 1999. Scaling fine-scale processes to large-scale patterns using models derived from models: meta-models. Pages 70-98 *in* D. J. Mladenoff and W. L. Baker, editors. Spatial Modeling of Forest Landscape Change: Approaches and Applications. Cambridge University Press, Cambridge, UK.

Urban, D. L., S. Goslee, K. Pierce, and T. Lookingbill. 2002. Extending community ecology to landscapes. Ecoscience 9:200-212.

Urban, D. L., H. H. Shugart, D. L. DeAngellis, and R. V. O'Neill. 1988. Forest Bird Demography in a Landscape Mosaic. ORNL/TM-10332, ESD Publication No. 2853. Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, TN.

Wikle, C. K. 2003. Hierarchical Bayesian models for predicting the spread of ecological processes. Ecology 84:1382-1394.

Wikle, C. K., L. M. Berliner, and N. Cressie. 1998. Hierarchical Bayesian space-time models. Environmental and Ecological Statistics 5:117-154.

CHAPTER 14

ASSESSING THE INFLUENCE OF SPATIAL SCALE ON THE RELATIONSHIP BETWEEN AVIAN NESTING SUCCESS AND FOREST FRAGMENTATION

PENN LLOYD, THOMAS E. MARTIN, ROLAND L. REDMOND, MELISSA M. HART, UTE LANGNER, AND RONALD D. BASSAR

14.1 INTRODUCTION

Ecological processes are dependent on the spatial and temporal scale at which they are viewed, and a process at any one scale may be influenced by factors at other scales. Thus, an ecological process at a broader scale may act to constrain processes at finer scales (Allen and Starr 1982, Thompson et al. 2000). Developing a full understanding of the spatial scales at which habitat conditions impinge on ecological processes therefore demands a multi-scale approach (Wiens 1989). The continuum of possible spatial scales can be broken into: (1) the space occupied by an individual, (2) the patch scale – the habitat patch occupied by many individuals and species, (3) the landscape scale – the collection of different habitat patches occupied by local populations, and (4) a biogeographic scale that encompasses different climates, vegetation formations, and assemblages of species (adapted from Wiens et al. 1986). Habitat fragmentation alters the spatial arrangement, shape and relative proportions of different habitat patches. These changes have a profound influence on ecological processes that are sensitive to alteration of the composition of environments, particularly at spatial scales 2-4 above.

Two ecological processes, nest predation and brood parasitism by the Brown-headed Cowbird (*Molothrus ater*), are the primary influences on nesting success of most North American land birds (Martin 1992). Many studies that have investigated the relationship between habitat and nesting success have focused on the question of how vegetation characteristics of the nest micro-environment influence nest success (e.g., Martin 1992, Larison et al. 2001). This focus is at the scale of the space occupied by an individual bird (Wiens et al. 1986). In contrast, studies at broader scales are mostly concerned with how variation in predator/parasite density or

J. Wu, K.B. Jones, H. Li, and O.L. Loucks (eds.),
Scaling and Uncertainty Analysis in Ecology: Methods and Applications, 259–273.

movement pattern associated with coarser-scale habitat features has an impact on nesting success. For example, at the scale of a local patch, nests situated closer to habitat edges often experience higher predation or parasitism rates, largely due to elevated predator/parasite densities at habitat edges (Andren and Angelstam 1988, Burke and Nol 2000). The relative density of predators and parasites within edge habitat may further depend on the relative proportions of different habitats at broader, landscape scales. Thus, a handful of studies have examined how nesting success varies with degree of habitat fragmentation within 1-10 km radii of study sites (Robinson et al. 1995, Tewksbury et al. 1999).

Understanding the ecological basis for scale dependence in nest predation and parasitism, requires an understanding of the distribution and abundance of predators and parasites in relation to edge, area, and biogeographic effects. These have been examined in greatest detail for the Brown-headed Cowbird. Cowbirds generally forage in open, short-grass habitats, particularly agricultural and other human-modified habitats (Lowther 1993). However, they often parasitize hosts breeding in forested habitats. Female cowbirds therefore commute daily between their foraging areas and nearby forested habitats to parasitize hosts (Donovan et al. 2000). This gives rise to an "edge effect" where parasitism is greater along the edges of forests that are closer to cowbird feeding habitat. Individual cowbirds are capable of commuting up to 15 km between foraging and breeding resources (Curson et al. 2000). In the eastern United States, however, average commuting distances are more commonly 1-3 km (Thompson 1994, Gates and Evans 1998, Raim 2000, Thompson and Dijak 2000). The distance that cowbirds will penetrate forest interiors is correlated with the local population size of cowbirds in suitable habitat surrounding the forest (Donovan et al. 1997). This local abundance of cowbirds is, to a large extent, limited by the availability of suitable feeding areas, an area effect. Thus, local cowbird abundance increases as the relative area of human-transformed, usually agricultural habitats increases (Donovan et al. 1997). Thus, edge effects at the patch scale are expected to be constrained by variation in local cowbird abundance associated with area effects at local landscape scales within as much as a 10 km radius.

At a biogeographic scale, the relative abundance of cowbirds is greatest within the Great Plains (incorporating portions of southern Canada, North and South Dakota, Nebraska, Kansas, and Oklahoma) and becomes progressively less towards the eastern and western edges of their range (Peterjohn et al. 2000). This pattern of relative abundance reflects the historical distribution of cowbirds in North America. They are believed to have been originally concentrated in the Great Plains of central North America but to have spread eastward and westward during the 19th and 20th centuries (Mayfield 1965, Rothstein 1994). Consequently, cowbird abundance decreases with increasing distance from the center of abundance in the Midwest, independently of continental patterns in land cover (Thompson et al. 2000). This is consistent with a general ecological pattern of spatial variation in abundance: density declines from the region of peak density towards the boundaries of the range (Brown 1984). Predator abundance and the risk of nest predation are hypothesized to exhibit similar scale dependence (Donovan et al. 1997, Tewksbury et al. 1999,

Thompson et al. 2002), although biogeographic variation has yet to be documented in sufficient detail.

Most studies have investigated the relationship between habitat features and nesting success at small and medium scales for individual species at usually single sites, but very few have investigated larger-scale influences. A notable exception is the study of Robinson et al. (1995), which quantified levels of nest predation and cowbird parasitism for nine species across a gradient of forest fragmentation (quantified within a 10 km radius of study sites) spanning nine landscapes within six mid-western states of the United States. Hochachka et al. (1999) also found that increased forest cover within a 10 km radius resulted in lower rates of cowbird parasitism of hosts in general. Using nest success data collected on the Ovenbird (*Seiurus aurocapillus*) collated in the Breeding Biology Research and Monitoring Database (BBIRD), we extend this approach to ask:

- what are the relationships between forest fragmentation and each of nest parasitism and nest predation?
- how do these relationships vary with the scale at which habitat fragmentation is assessed?

The Ovenbird is a Neotropical migrant songbird that nests on the ground in the interior of mature forests in the eastern United States and Canada. It is known to be sensitive to forest edges, incurring higher nest mortality in edge habitat, and is a preferred host of the Brown-headed Cowbird (Van Horn and Donovan 1994). It is therefore expected to be highly sensitive to forest fragmentation.

14.2 METHODS

Ovenbird nest success data were collated from 121 plots within 15 BBIRD sites scattered across the eastern United States (Figure 14.1). Within each site, between 3 and 20 plots, each incorporating an area of approximately 10-50 ha, were separated by distances of 1-50 km from one another. Study sites were not selected randomly. Contributing investigators selected sites and plots to meet the needs of their own research agendas, but used the standardized BBIRD protocol (see Martin and Geupel 1993) for collecting the nest data. We treated the plots of more than one investigator as a single site if these plots were less than 50 km apart.

Habitat features at a variety of spatial scales were derived from the National Land Cover Dataset (NLCD; Vogelmann et al. 2001). The NLCD recognizes 21 different land cover types mapped at a 30 m^2 pixel resolution. For our analyses, we grouped certain NLCD land cover types to recognize the following four land cover types of principal interest:

- forest – includes deciduous, evergreen and mixed Forests, and woody wetlands;
- grassland – includes grassland/herbaceous, and pasture/hay;
- cropland – includes orchards/vineyards/other, row crops, small grains, and fallow; and
- developed – includes grassland and cropland, as defined above, together with residential and commercial/industrial/transportation.

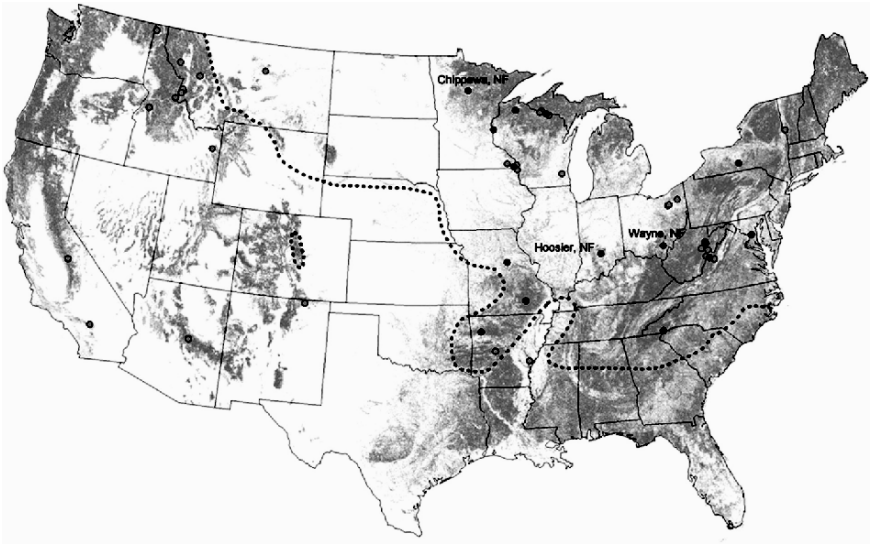


Figure 14.1. *Distribution of the 15 Ovenbird sites (solid circles) and other BBIRD sites (open circles) across the United States relative to forest cover (grey). The breeding range of the Ovenbird extends to the north and east of dashed line, with a disjunct population in Colorado (Van Horn & Donovan 1994).*

We then extracted a number of landscape-level habitat metrics using ArcInfo (ESRI 2002) and Apack (Mladenoff and DeZonia 2002) software at four principal spatial scales:

- the forest patch within which the study plot was embedded;
- within radii of 1, 5, and 10 km of each study plot center;
- within radii of 50, 100, and 150 km of each study site center (defined as the center of the minimum convex polygon connecting all plots); and
- the biogeographic region.

For analysis, we selected a suite of landscape-level habitat metrics that we thought had intuitive biological meaning with respect to the hypothesized edge, area and biogeographic influences on predator/parasite abundance. These included distance from plot center to nearest forest edge or nearest developed land cover edge (at the patch scale), percent forest, core forest, grassland, cropland and developed land cover (at the landscape scale). Using Breeding Bird Survey data downloaded from the USGS Patuxent Wildlife Research Center server (pwrcftpr.er.usgs.gov/research/sauer), we also examined the distance between site center and the nearest edge of core distribution (average count ≥30) of cowbirds in the Midwest (at the biogeographic scale). We used multiple regression (forward stepwise) to analyze and control for the effects of these predictor variables on the rates of predation and brood parasitism experienced by Ovenbirds at 15 sites. Suitable transformations were applied to all variables that were not normally distributed.

14.2.1 Comparison with Neutral Landscapes

To test (1) the representativeness of the BBIRD study sites and (2) the scaling relationship between forest area and increasing radial distance, we compared forest metrics within increasing radial distances around plot and site centers between the 15 ovenbird sites and a series of 202 sets of neutral landscapes that were randomly selected throughout the eastern US (defined as the continental land area east of 100th Meridian). The number and distribution of sites and associated plots within each randomized set was designed to mimic the total sample of forested sites in the BBIRD database east of the 100th Meridian. We had 251 plots distributed among 27 unique forested sites east of the 100th Meridian (range = 1-31 plots/site and 0-64360 m between site and plot centers). Thus, the number of random sites in each set was limited to 27; the number and distribution of plots within each site varied from 1-31; and the distance between these plots could not exceed 64.36 km. Further details of the sequential steps involved in the randomization process follow below.

Firstly, the inside perimeter of the 48 contiguous United States was buffered by 30 km. This was done to avoid or reduce the number of water or no data cells that might skew landscape statistics calculated within larger radii (50, 100, 150 km) around site centers that fell close to Canada, Mexico, or a major coastline. Secondly, the analysis was limited to the area within the buffer east of 100th Meridian. Furthermore, sampling parameters were defined such that the number and dispersion of plots at each site conformed to the pattern observed among the 27 BBIRD sites (see above). Thirdly, for each run, 27 sites were selected randomly. For a point to be selected as a random site center, two criteria had to be met: (a) it had to fall on a 30 m grid cell representing a forested cover type in the NLCD and (b) at least 10% of the cells within a 100 km radius had to represent forested cover types as well. Once a site center was suitably located, random points representing individual plot centers were selected, one at a time, and each evaluated according to the same two criteria as the site centers. After 27 random sites and their surrounding 260 plots were located in this manner, new site centers were assigned to each based on minimum convex polygons enclosing the plots at each site. This process was repeated 201 times, resulting in 5454 random sites and 52520 plots. The data were stored as two GIS point coverages, one containing all 202×27 random site centers, and the other all 202×260 random plot centers. Percent area covered by forest was calculated within 1, 5, and 10 km of all random plot centers and within 50, 100, and 150 km of all site centers. As with the landscape analyses of the actual 15 Ovenbird sites, areas of water or no data (e.g., outside the extent of NLCD coverage) were not included in calculations of land cover proportions (e.g., % forest, grassland, agriculture, etc.) around any of the random plots or sites. Finally, expecting that landscape configuration would play some role in determining the reproductive success of ovenbird populations, we examined selected metrics for the 15 Ovenbird landscapes and compared those to neutral landscapes created using Apack's percolation map output option (Mladenoff and DeZonia 2002). Neutral landscapes (Gardner et al. 1987) have randomly-assigned cell values, but maintain exactly the same land cover proportions as the input map, allowing us to ascertain how Ovenbird landscape structure differs from random with percent forest cover held equal.

14.2.2 Sources of Uncertainty

In our study, uncertainty was derived primarily from sampling error and high natural variability in nest predation and parasitism, and secondarily from other data quality issues such as measurement and database management error. Sampling error related to: (1) differences in sample size of nests from sites/plots for which predation and parasitism rates were determined; and (2) non-random selection of sites/plots within the geographical distribution of the species under study. Sample size varied from five to 241, which led to a wide range in the standard error associated with individual data points used in the regression analyses.

Measurement error may arise from: (1) differences in the interpretation of nest fate among contributors; (2) positional errors in the mapped locations of study plots; (3) variation in the distribution of plots within sites (tightly clustered at some sites, but scattered over many kilometers at others) that affects the representativeness of the site center location; (4) errors in land cover classification and image registration in the NLCD land cover database; and (5) the loss in resolution resulting from the aggregation of nest success data at a spatial scale potentially greater than that being assessed. We were unable to quantify error in nest-fate interpretation, but a detailed field protocol and annual meetings with contributors probably reduced it to an acceptable level. The coordinates for most study plot centers were estimated to be within 10 m of their true locations, although a few may be off by as much as 1 km. As a result, outputs may be affected by actual positional errors as well as errors in relation to NLCD land cover. Although the NLCD land cover database offers the most current and consistent national coverage at high resolution (30 m), it may either fail to map the appropriate habitat features for Ovenbirds, cowbirds and/or nest predators, or fail to represent them at adequate resolutions. Ideally, a distance measure to the nearest edge is required for each nest, but this was not available. As a best approximation, we aggregated data from all nests in a plot, and measured distance from plot center to the nearest edge. This may aggregate data from nests within both edge and interior habitat, resulting in a loss of resolution. Aggregating data from all plots within a site results in a further loss in resolution.

14.3 RESULTS

We found a similar scaling relationship between the increase in total forest area within increasing radii of landscape area for both the 15 Ovenbird sites (described by: $y = 1.911x + 2.439$, $R^2 = 1$ on a log-log plot) and the 5454 sites that were randomly selected throughout the eastern United States (described by: $y = 1.946x + 2.378$, $R^2 = 1$ on log-log plot). In both cases, the exponents (1.91 and 1.95) were slightly less than the isometric value of 2, indicating that forest cover tends to become slightly less abundant with increasing breadth of scale (i.e., the radial distance surrounding the analysis area). For both the random and actual sites, doubling the measurement scale from a 50 to 100 km radius around the site center results in substantially less than twice the proportional amount of forest found within the larger area ($2^{1.91} = 3.75$ times the proportional amount rather than $2^2 = 4.0$ if scaling was isometric).

Comparisons of angular second moment, aggregation index, and contagion suggest that the Ovenbird landscapes have fewer cover classes, more clumping, and fewer patches than the neutral landscapes (Table 14.1). For angular second moment and contagion, differences were strongest for Ovenbird sites with relatively high forest cover. Total edge density suggests that the Ovenbird landscapes have lower perimeter/area ratios than might be expected, meaning that patches tend to be more circular or square, and/or larger. Indeed, the average patch perimeter/area ratios for Ovenbird landscapes were lower than for neutral landscapes. Ovenbird landscapes also tended to have patches with larger perimeters based on the average perimeter of patches in landscape; this in turn suggests more complex shapes.

Among the 15 BBIRD sites, percent forest cover was highly correlated across all spatial scales with percent grassland cover ($r = -0.82$ to -0.86, $P < 0.001$), percent cropland cover ($r = -0.83$ to -0.90), and percent developed land cover ($r = -0.91$ to -0.99). Distance to nearest edge/developed edge was significantly correlated with percent forest cover at finer scales, such as within a 1 km radius ($r = 0.64–0.77$, $P = 0.01–0.001$), but not with percent forest cover at broader scales, such as within a 100 km radius ($r = 0.31–0.43$, $P = 0.11–0.26$). Cowbird abundance was significantly negatively correlated with both distance from the center of cowbird distribution in the Midwest ($r_p = -0.85$, $P < 0.001$), and with percent forest cover within a 100 km radius ($r_p = -0.91$, $P < 0.001$).

14.3.1 Effects on Nest Parasitism

Sample size (number of nests) did not affect estimates of nest parasitism rate among plots ($F_{45} = 1.58$; $P = 0.22$) or sites ($F_{14} = 1.28$; $P = 0.28$). Among sites, nest parasitism rate was positively correlated with average distance to the nearest developed edge and patterns of land cover (particularly developed land cover) across all landscape scales, but was not related to distance from the center of cowbird distribution (Table 14.2). Although autocorrelation among scales complicates comparisons of correlation coefficients across scales, the data suggest that land cover patterns at the 5-100 km landscape scales are likely the best predictors of nest parasitism rate. The relationship between nest parasitism (arcsine-square-root-transformed) and percent developed land cover within a 10 km radius (arcsine-square-root-transformed) is described by the regression: $y = 0.774x + 0.02269$ ($F = 36.9$, $P < 0.001$, $R^2 = 0.72$).

At only three sites did sufficient variation in percent forest cover among plots at one or more of the 1-10 km radii scales enable tests of the relationships between indices of forest fragmentation and each of Ovenbird nest parasitism rate and nest predation rate. At each of these sites, parasitism rate was not significantly related to distance to nearest developed edge, but generally increased as percent forest cover at landscape scales declined. This negative relationship was significant at the 10 km radius scale at Hoosier National Forest ($r = -0.81$, $P = 0.03$), and at both the 5 km ($r = -0.74$, $P = 0.002$) and 10 km ($r = -0.60$, $P = 0.02$) radii scales at Wayne National Forest (Figure 14.2). In addition, the three sites differed significantly in the levels of parasitism on Ovenbirds, independent of the effect of forest cover at the 1-10 km radii scales (Table 14.3). These differences can be attributed to differences

in percent forest cover between sites at the broader landscape scale. Thus, Ovenbirds breeding on plots with relatively low forest cover at the 100 km radius scale generally experienced considerably higher nest parasitism than those breeding on plots with relatively high forest cover within a 100 km radius, independent of the degree of forest cover at smaller spatial scales.

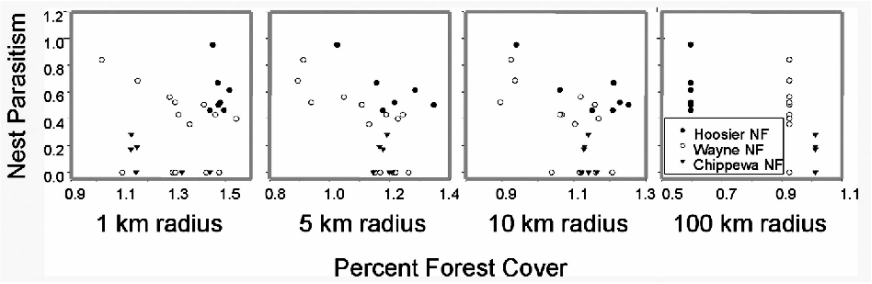


Figure 14.2. *Relationship between nest parasitism rate (arcsine transformed) and percent forest cover (arcsine transformed) at a spatial scale of 1 km, 5 km and 10 km radii of plot centers, and 100 km radius of site centers for three sites. Relative percent forest cover at the 100 km radius scale is classified as low (Hoosier NF), medium (Wayne NF) or high (Chippewa NF).*

Table 14.1. *For scales of 50 and 100 km, comparison of landscape metrics for 15 Ovenbird sites and their corresponding neutral landscapes. ASM = angular second moment; AI = aggregation index; COR = contagion - Riitters; TED = total edge density; PPAR = patch perimeter/area ratio; APP = average patch perimeter in landscape.*

Landscape Metric	*Radius (km)*	*Ovenbird Landscapes*			*Neutral Landscapes (n = 15)*
		Low Forest Cover (n = 5)	*Medium Forest Cover (n = 3)*	*High Forest Cover (n = 7)*	
		mean ± sd	*mean ± sd*	*mean ± sd*	*mean ± sd*
ASM	50	0.139 ± 0.026	0.319 ± 0.042	0.553 ± 0.139	0.221 ± 0.147
ASM	100	0.139 ± 0.015	0.294 ± 0.034	0.452 ± 0.107	0.171 ± 0.095
AI	50	0.807 ± 0.021	0.888 ± 0.021	0.926 ± 0.026	0.515 ± 0.208
AI	100	0.818 ± 0.018	0.879 ± 0.019	0.911 ± 0.020	0.464 ± 0.165
COR	50	0.358 ± 0.042	0.533 ± 0.054	0.659 ± 0.113	0.413 ± 0.167
COR	100	0.371 ± 0.053	0.512 ± 0.058	0.592 ± 0.086	0.375 ± 0.129
TED	50	0.130 ± 0.015	0.075 ± 0.014	0.050 ± 0.017	0.324 ± 0.139
TED	100	0.122 ± 0.012	0.081 ± 0.012	0.059 ± 0.013	0.358 ± 0.110
PPAR	50	1.080 ± 0.009	1.090 ± 0.039	1.108 ± 0.031	1.249 ± 0.044
PPAR	100	1.076 ± 0.008	1.088 ± 0.025	1.100 ± 0.027	1.241 ± 0.039
APP	50	0.770 ± 0.063	0.846 ± 0.187	0.721 ± 0.111	0.351 ± 0.084
APP	100	0.791 ± 0.063	0.854 ± 0.112	0.737 ± 0.093	0.363 ± 0.088

14.3.2 Effects on Nest Predation Rate

Daily nest predation rate increased with increasing sample size among sites (F_{14} = 7.33; P = 0.02), but not among plots (F_{45} = 0.07; P = 0.79). These sample-size effects are difficult to explain, largely because they are not repeated across both plots and sites. An increase in nest predation rate with sample size might be expected if sample size reflects breeding density and nest predation is positively density dependent. After controlling for the sample size effect, predation rate correlated positively with degree of forest fragmentation at the 1-100 km radii scales among sites, with percent area developed at the 5-10 km radii scales serving as the best predictors of predation on Ovenbird nests (Table 14.2). Among plots within the three sites with sufficient variation in landscape context, nest predation was significantly negatively correlated with increasing distance from the nearest edge at Chippewa National Forest (r = –0.84, P = 0.38), but not at Hoosier National Forest (r = –0.68, P = 0.09) or Wayne National Forest (r = 0.35, P = 0.2), and was not related to percent forest cover at any landscape scale.

Table 14.2. *Summary of the best predictor variables (fragmentation indices) for the relationship between each of nest parasitism rate and nest predation rate among BBIRD sites (plot averages for scales of patch and 1-10 km radii) using multiple regression analysis.*

Scale	*Independent variables*	*Dependent variables*	*Adjusted R^2*	*Partial Correlation*
Patch	Parasitism	To developed edge	0.50**	
	Predation	Sample size	0.36*	
1 km	Parasitism	Percent area developed	0.32*	
	Predation	Sample size	0.51*	0.75
		Percent area developed		0.58
5 km	Parasitism	Percent area developed	0.70***	
	Predation	Sample size	0.67***	0.83
		Percent area developed		0.74
10 km	Parasitism	Percent area developed	0.72***	
	Predation	Sample size	0.66***	0.83
		Percent area developed		0.73
50 km	Parasitism	Percent forest cover	0.72***	
	Predation	Sample size	0.54**	0.74
		Percent area developed		0.61
100 km	Parasitism	Percent area developed	0.70***	
	Predation	Sample size	0.58**	0.73
		Percent forest cover		−0.66
Biogeographic	Parasitism	To center cowbird range	−0.08	

$^{*}P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$.

14.4 DISCUSSION

The allometric scaling relationship observed between landscape measurement length (radial distance) and forest area is similar to that found by Wiens and Milne (1989) in their examination of the area of bare soil in grassland microlandscapes (fractional exponent of 1.8; see Schneider 1998). Two points are important for the interpretation of this relationship. First, it may reflect the broad extent of anthropogenic impact across the eastern US – the larger the area examined, the more people are likely to be included in the analysis, such that we would expect or predict a positive allometric relationship (exponent > 2.0) between length of landscape measurement (i.e., radial distance) and area impacted by people. Further analyses are required to explore this hypothesis. Second, BBIRD investigators selected plots that were located in forest vegetation (the better to find Ovenbird nests). Thus, the plot locations were defined at a local scale, whereas for the 202 randomization runs, the broader landscape context was considered. That is, potential sites and plots were selected only if they had at least 10% forest within 100 km. These random results did not produce a significantly different pattern in forest cover at any landscape scale from the 15 Ovenbird sites, suggesting that: (1) accounting for land cover at a broader scale had no noticeable effect on the observed relationship and (2) BBIRD sites are representative of overall forest conditions in the eastern United States.

Furthermore, we observed that the landscapes surrounding the 15 sites with Ovenbird data, particularly those with relatively high forest cover, tend to have structural characteristics in the surrounding 50 and 100 km that accord reasonably well with those preferred by Ovenbirds (Van Horn and Donovan 1994). That is, landscapes are characterized by larger patches with simpler shapes occurring in more clumped configurations than would be expected in neutral landscapes.

If patch-specific edge effects on nest parasitism and predation are constrained by landscape and biogeographic effects, then we could expect variation in nest parasitism and predation among sites to partition among spatial scales in a multivariate analysis. This was not the case. Instead, we found that univariate models were always most parsimonious. Our inability to detect edge effects on nest parasitism rate, even among plots within individual sites, may be due to a combination of inappropriate resolution due to data aggregation within plots and sites, plot location errors, and landscape area effects overwhelming edge effects. Nonetheless, studies that have focused on the local patch scale have found strong edge effects on Ovenbird nest parasitism (Donovan et al. 1995, Porneluzi and Faaborg 1999, Burke and Nol 2000, Flaspohler et al. 2001). Despite the strong evidence of biogeographic variation in cowbird abundance, which is expected to translate into biogeographic variation in nest parasitism, we found that nest parasitism rate on Ovenbird nests was unrelated to distance from the core area of cowbird abundance in the Midwest. Our results suggest, therefore, that landscape-level patterns of land cover, which are most significantly correlated with nest parasitism in our analyses, may overwhelm any biogeographic effect. Landscape area effects are not expected to extend beyond the home-range limits of cowbirds (maximum 20 km radius). Nonetheless, among BBIRD sites with similar percent forest cover at local landscape scales (within 1-10 km radius), considerable variation

in parasitism rate remains (Figure 14.2). Much of this variation is correlated with percent forest cover at the broader landscape scales (50-100 km radii of study sites), suggesting that the relative severity of parasitism among sites may be determined by the broader landscape context. As an illustrative example, the Hoosier National Forest study site averages 98%, 88% and 84% forest cover within 1 km, 5 km and 10 km radii of its study plots, i.e., it is located within the interior of a locally heavily forested area, with most of the Ovenbirds studied nesting at least 2 km from the nearest agricultural edge. Nest parasitism risk could, therefore, be expected to be relatively low. Yet, Ovenbirds breeding at this site suffered a 30% nest parasitism rate. This is likely due to the Hoosier site being surrounded by heavily fragmented, largely agricultural habitat at a broader landscape scale, which reduces the percent forest cover at the 100 km radius scale to just 31%.

Edge effects on nest predation rates were detected only in comparisons of plots within individual sites, but not in the comparisons of site averages among sites (Table 14.2). This is not surprising given that the edge-mediated increase in predation risk for Ovenbirds, mostly from edge-adapted predators such as corvids, raccoons and opossums, generally extends no farther than 100-200 m from forest edges, an effect that may be lost in aggregate data. Published daily nest predation rates on Ovenbird nests range from 4.2%-7.8% within 200 m of forest/agricultural edges (Burke and Nol 2000, Rodewald and Yahner 2001), and 3.8-5.4% within 200 m of recent clear-cut edges within forests (King et al. 1996, Flaspohler et al. 2001, Rodewald and Yahner 2001), but just 1.7-2.7% within interior forest > 200 m from the nearest edge (King et al. 1996, Burke and Nol 2000, Flaspohler et al. 2001). The among-site analyses suggest that Ovenbird nest predation is significantly correlated with landscape context within a 5 km radius (Table 14.3). The potentially strong influence of landscape context on nest predation rates was recently highlighted by a rigorously designed experiment using artificial ground nests. Donovan et al. (1997) found that in landscapes with <15% forest within a 10 km radius, predation was high in forest edge and interior; at 45-55% forest cover, predation was high in forest edge and low in forest interior; and at >90% forest cover, predation was low in both forest edge and interior.

Table 14.3. *Results from ANCOVA tests for three different radii from plot centers at three sites (Site 1 = Hoosier National Forest; site 2 = Wayne National Forest; site 3 = Chippewa National Forest). The dependent variable is nest parasitism rate, with site as a main effect and percent forest cover as a linear covariate.*

	1 km (Sites 2 and 3)		5 km (Sites 1 and 2)		10 km (Sites 1 and 2)	
Source of variation	*F*	*P*	*F*	*P*	*F*	*P*
Forest cover	2.76	0.12	19.16	<0.001	11.54	0.003
Site × forest cover	5.34	0.03	19.25	<0.001	14.32	0.001

To summarize, the relatively coarse resolution of our aggregated data is likely to downplay the influence of edge effects, particularly edge effects on nest predation,

which tend to be limited to 100-200 m of forest edges. Edge effects on nest parasitism may be limited to 100-200 m of forest edges in extensively forested landscapes (Burke and Nol 2000), or extend several kilometers into forest in highly fragmented landscapes (Robinson and Robinson 1999, Ford et al. 2001). Nonetheless, there is incontrovertible evidence of strong edge effects on Ovenbird nest parasitism and predation from a host of detailed field studies. Our results emphasize the overwhelming influence of landscape context, particularly on nest parasitism rates, in agreement with other studies and reviews (Donovan et al. 1997, Thompson et al. 2002). Although we found no strictly biogeographic effect on nest parasitism, biogeographic variation in cowbird abundance that is independent of relative feeding resource availability at landscape scales does exist, and can be expected to constrain the influence of landscape area effects on nest parasitism (Thompson et al. 2000).

Considered as a whole, the available evidence highlights the importance of considering the effects of forest fragmentation on patch-specific demography within a top-down spatial hierarchy that includes biogeographic effects exerting constraints on landscape-level effects, which, in turn, exert constraints on patch-scale edge effects, as proposed by Thompson et al. (2002). This hierarchical conceptualization of the scale-dependence of ecological processes has important implications for scaling, the extrapolation of ecological processes across scales (Wu 1999, Li and Wu, Chapter 3). Using knowledge of patch-specific edge effects of parasitism and predation on nesting success to estimate average nesting success within the patch as a whole may require only information on the edge-perimeter to core area ratio for that patch, assuming that the influence of the landscape is uniform for the patch as a whole. However, landscape context will likely vary among patches as landscape extent increases, and biogeographic context will vary among landscapes. Thus, scaling up an ecological process, such as nest predation or parasitism, from a patch-scale model to a model for a larger spatial extent requires consideration of the changes in the functional representation (i.e., transmutation) of that ecological process at each of the recognized levels of the spatial hierarchy (King et al. 1991). BBIRD data are derived from point locations within patches. We describe the relationships between these point data and habitat variables at different scales around these points, but we do not attempt to construct scaling relations to scale up from these data, as they lack sufficient resolution at an important level in the hierarchy, that of the relationship between nest success and distance to edge. With sufficient data resolution at the patch, landscape and biogeographic scales, multilevel statistical models (as described by Berk and de Leeuw, Chapter 4) could be used to derive scaling functions to describe the transmutation of processes such as nest predation and parasitism across these spatial scales.

14.5 CONCLUSIONS

Ovenbird nest parasitism rate and daily nest predation rate correlated positively with the extent of forest fragmentation at landscape scales, particularly within a 5-10 km radius of study plots. Our results emphasize the overwhelming influence of landscape context over edge effects, especially on nest parasitism rates. Given the

extensive published evidence of edge effects on nest predation in particular, we highlight the importance of considering the effects of forest fragmentation on patch-specific demography within a top-down spatial hierarchy.

ACKNOWLEDGEMENTS

We extend our grateful thanks to all contributors to the BBIRD database, especially those investigators whose sites yielded sufficient Ovenbird data for analysis, including David Andersen, Vanessa Artman, Jaime Collazo, Robert Cooper, Randy Dettmers, Therese Donovan, John Faaborg, David Flaspohler, Melinda Knutson, Lisa Petit, Vicki Piaskowski, Charles Smith, Don Whitehead, and Petra Wood. We also thank those at MTCWRU who have worked on aspects of this project, past and present, including Mark Clark, Wes Hochachka, Leslie Jette, Shane Mason and Jim Schumacher. Finally, our thanks to Barry DeZonia for his assistance with questions about APACK software. We are grateful for funding support from USGS BRD Global Change Program and USDA Forest Service.

REFERENCES

Allen, T. F. H., and T. B. Starr. 1982. Hierarchy: Perspectives for Ecological Complexity. University of Chicago Press, Chicago.

Andren, H., and P. Angelstam. 1988. Elevated predation rates as an edge effect in habitat islands: experimental evidence. Ecology 69:544-547.

Brown, J. H. 1984. On the relationship between abundance and distribution of species. American Naturalist 124: 255-279.

Burke, D. M., and E. Nol. 2000. Landscape and fragment size effects on reproductive success of forest-breeding birds in Ontario. Ecological Applications 10:1749-1761.

Curson, D. R., C. B. Goguen, and N. E. Mathews. 2000. Long-distance commuting by Brown-headed Cowbirds in New Mexico. Auk 117:795-799.

Donovan, T. M., P. W. Jones, and F. R. Thompson. 1997. Variation in local-scale edge effects: mechanisms and landscape context. Ecology 78:2064.

Donovan, T. M., F. R. Thompson, and J. Faaborg. 2000. Cowbird distribution at different scales of fragmentation: tradeoffs between breeding and feeding opportunities. Pages 255-264 *in* J. N. M. Smith, T. L. Cook, S. I. Rothstein, S. K. Robinson, and S. G. Sealy, editors. Ecology and Management of Cowbirds and their Hosts. University of Texas Press, Austin, TX.

Donovan, T. M., F. R. Thompson, J. Faaborg, and J. R. Probst. 1995. Reproductive success of migratory birds in habitat sources and sinks. Conservation Biology 9:1380-1395.

ESRI. 2002. ArcInfo 8.2. Environmental Systems Research Institute, Inc., Redlands, CA.

Flaspohler, D. J., S. A. Temple, and R. N. Rosenfield. 2001. Species-specific edge effects on nest success and breeding bird density in a forested landscape. Ecological Applications 11:32-46.

Ford, T. B., D. E. Winslow, D. R. Whitehead, and M. A. Koukol. 2001. Reproductive success of forest-dependent songbirds near an agricultural corridor in south-central Indiana. Auk 118:864-873.

Gardner, R. H., B. T. Milne, M. G. Turner, and R. V. O'Neill. 1987. Neutral models for the analysis of broad-scale landscape pattern. Landscape Ecology 1:19-28.

Gates, J. E., and D. R. Evans. 1998. Cowbirds breeding in the central Appalachians: spatial and temporal patterns and habitat selection. Ecological Applications 8:27-40.

Hochachka, W. M., T. E. Martin, V. Artman, C. R. Smith, S. J. Hejl, D. E. Andersen, D. Curson, L. Petit, N. Mathews, T. Donovan, E. E. Klaas, P. B. Wood, J. C. Manolis, K. P. McFarland, J. V. Nichols, J. C. Bednarz, D. M. Evans, J. P. Duguay, S. Garner, J. Tewksbury, K. L. Purcell, J. Faaborg, C. B. Goguen, C. Rimmer, R. Dettmers, M. Knutson, J. A. Collazo, L. Garner, D. Whitehead, and G. Geupel. 1999. Scale dependence in the effects of forest coverage on parasitization by Brown-headed Cowbirds. Studies in Avian Biology 18:80-88.

King, A. W., A. R. Johnson, and R. V. O'Neill. 1991. Transmutation and functional representation of heterogeneous landscapes. Landscape Ecology 5:239-253.

King, D. I., C. R. Griffin, and R. M. Degraaf. 1996. Effects of clearcutting on habitat use and reproductive success of the Ovenbird in forested landscapes. Conservation Biology 10:1380-1386.

Larison, B., S. A. Laymon, P. L. Williams, and T. B. Smith. 2001. Avian responses to restoration: nest-site selection and reproductive success in Song Sparrows. Auk 118:432-442.

Lowther, P. E. 1993. Brown-headed Cowbird. *in* A. Poole and F. Gill, editors. The Birds of North America. Academy of Natural Sciences, Philadelphia, and American Ornithologists' Union, Washington, D.C.

Martin, T. E. 1992. Breeding productivity considerations: what are the appropriate habitat features for management? Pages 455-473 *in* J. M. Hagan and D. W. Johnston, editors. Ecology and Conservation of Neotropical Migrants. Smithsonian Institution Press, Washington, D.C.

Martin, T. E., and G. E. Geupel. 1993. Nest-monitoring plots: methods for locating nests and monitoring success. Journal of Field Ornithology 64:507-519.

Mayfield, H. F. 1965. The Brown-headed Cowbird with old and new hosts. Living Bird 4:13-28.

Mladenoff, D. J., and B. DeZonia. 2002. APACK 2.22 Analysis Software Users Guide. Department of Forest Ecology and Management, University of Wisconsin, Madison, WI.

Peterjohn, B. G., J. R. Sauer, and S. Schwarz. 2000. Temporal and geographic patterns in population trends of brown-headed Cowbirds. Pages 21-34 *in* J. N. M. Smith, T. L. Cook, S. I. Rothstein, S. K. Robinson, and S. G. Sealy, editors. Ecology and Management of Cowbirds and Their Hosts. University of Texas Press, Austin, TX.

Porneluzi, P. A., and J. Faaborg. 1999. Season-long fecundity, survival, and viability of Ovenbirds in fragmented and unfragmented landscapes. Conservation Biology 13:1151-1161.

Raim, A. 2000. Spatial patterns of breeding female Brown-headed Cowbirds on an Illinois site. *in* J. N. M. Smith, T. L. Cook, S. I. Rothstein, S. K. Robinson, and S. G. Sealy, editors. Ecology and Management of Cowbirds and Their Hosts. University of Texas Press, Austin, TX.

Robinson, S. K., and W. D. Robinson. 1999. Avian nesting success in a selectively harvested north temperate deciduous forest. Conservation Biology 15:1763-1771.

Robinson, S. K., and D. S. Wilcove. 1994. Forest fragmentation in the temperate zone and its effects on migratory songbirds. Bird Conservation International 4:233-249.

Robinson, S. K., F. R. Thompson, T. M. Donovan, D. R. Whitehead, and J. Faaborg. 1995. Regional forest fragmentation and the nesting success of migratory birds. Science 267:1987-1989.

Rodewald, A. D., and R. H. Yahner. 2001. Avian nesting success in forested landscapes: influence of landscape composition, stand and nest-patch microhabitat, and biotic interactions. Auk 118: 1018-1028.

Rothstein, S. I. 1994. The cowbird's invasion of the far west: History, causes, and consequences experienced by host species. Studies in Avian Biology 15:301-315.

Schneider, D. C. 1998. Applied scaling theory. Pages 253-270 *in* D. L. Peterson and V. T. Parker, editors. Ecological Scale: Theory and Applications. Columbia University Press, New York.

Tewksbury, J. J., T. E. Martin, S. J. Heil, T. S. Redman, and F. J. Wheeler. 1999. Cowbirds in a western valley: effects of landscape structure, vegetation and host density. Studies in Avian Biology 18: 23-33.

Thompson, F. R. 1994. Temporal and spatial patterns of breeding Brown-headed Cowbirds in the Midwestern United States. Auk 111:979-990.

Thompson, F. R., and W. D. Dijak. 2000. Differences in movements, home range, and habitat preferences of female Brown-headed Cowbirds in three Midwestern landscapes. *in* J. N. M. Smith, T. L. Cook, S. I. Rothstein, S. K. Robinson, and S. G. Sealy, editors. Ecology and Management of Cowbirds and Their Hosts. University of Texas Press, Austin, TX.

Thompson, F. R., T. M. Donovan, R. M. DeGraaf, J. Faaborg, and S. K. Robinson. 2002. A multi-scale perspective of the effects of forest fragmentation on birds in eastern forests. Studies in Avian Biology 25:8-19.

Thompson, F. R., S. K. Robinson, T. M. Donovan, J. R. Faaborg, D. R. Whitehead, and D. R. Larsen. 2000. Biogeographic, landscape, and local factors affecting cowbird abundance and host parasitism levels. Pages 271-279 *in* J. N. M. Smith, T. L. Cook, S. I. Rothstein, S. K. Robinson, and S. G. Sealy, editors. Ecology and Management of Cowbirds and Their Hosts. University of Texas Press, Austin, TX.

Van Horn, M. A., and T. M. Donovan. 1994. Ovenbird *(Seiurus aurocapillus),* No. 88. *in* A. Poole and F. Gill, editors. The Birds of North America. Academy of Natural Sciences, Philadelphia, and American Ornithologists' Union, Washington, D.C.

Vogelmann, J. E., S. M. Howard, L. Yang, C. R. Larson, B. K. Wylie, and N. V. Direl. 2001. Completion of the 1990s national land cover data set for the conterminous United States from Landsat Thematic Mapper data and ancillary data sources. Photogrammetric Engineering and Remote Sensing 67: 650-662.

Wiens, J. A., and B. T. Milne. 1989. Scaling of 'landscape' in landscape ecology, or, landscape ecology from a beetle's perspective. Landscape Ecology 3:87-96.

Wiens, J. A., J. F. Addicott, T. J. Case, and J. Diamond. 1986. Overview: the importance of spatial and temporal scale in ecological investigations. Pages 145-153 *in* J. Diamond and T. J. Case, editors. Community Ecology. Harper and Row, New York.

Wu, J. 1999. Hierarchy and scaling: extrapolating information along a scaling ladder. Canadian Journal of Remote Sensing 25:367-380.

CHAPTER 15

SCALING ISSUES IN MAPPING RIPARIAN ZONES WITH REMOTE SENSING DATA:

Quantifying Errors and Sources of Uncertainty

THOMAS P. HOLLENHORST, GEORGE E. HOST, AND LUCINDA B. JOHNSON

15.1 INTRODUCTION

Riparian zones are the ecotones or transition areas between upland and aquatic ecosystems, located at the margins of rivers, lakes, ponds and wetlands. Their boundaries are defined by changes in soil, moisture, and vegetation (Naiman 2000, Décamps 1996, Gregory et al. 1991). Although these ecosystems may be small relative to the aquatic systems they abut, they perform many important ecosystem services, including shading (thus buffering air and water temperature), retaining nutrients and/or sediments, stabilizing stream banks and littoral zones, and providing organic material (leaves, wood) and critical habitat for a diverse community of plant and animal species (Malanson 1993). Riparian zones are highly variable systems whose structure and composition are shaped by geomorphology, vegetation patterns, disturbance regimes (Décamps 1996), as well as current (Erickson and DeYoung 1993) and perhaps historic land use practices (Foster et al. 2003). Processes that operate over a large range of temporal and spatial scales control these structuring factors. At one end of the time/space continuum are processes such as tectonics, volcanism, glaciation, and climate change. At intermediate spatial and temporal scales are historic land use practices (e.g., burning regimes implemented by native peoples, permanent land cover conversion) and catastrophic flooding. At small scales, localized flooding and land management practices influence the structure and function of riparian zones. Some processes occur over multiple scales and their effects may also vary by scale.

The fine-scale variation of vegetative cover resulting from moisture and soil gradients around streams and wetlands has posed a challenge to research scientists and land managers (Muller 1997, Congalton et al. 2002). One of the current

J. Wu, K.B. Jones, H. Li, and O.L. Loucks (eds.),
Scaling and Uncertainty Analysis in Ecology: Methods and Applications, 275–295.

challenges in watershed science is the development of economical and accurate methods for mapping riparian areas. The problem of quantifying riparian zones is particularly difficult in developed areas (e.g., urban and agricultural) and arid landscapes, where riparian zones are frequently narrower than the resolution of the standard land cover data sets often used for watershed characterization.

The current widespread availability of inexpensive and user-friendly geographic information systems (GIS), and a larger selection of spatial data layers with global to regional coverage has expanded the use of these tools worldwide. A negative effect of the expanded use of these tools and data is that users occasionally overlook the inherent limitations, particularly with respect to the scale (grain and extent) of the data relative to the question being asked. The recent availability of an updated USGS National Land Cover Dataset (NLCD; Vogelmann et al. 2001) has enabled researchers, managers and policy makers to explore many new questions. Of particular interest is the possible use of the NLCD for addressing water quality questions from a watershed perspective.

The specific requirements for mapping vegetation communities from satellite imagery were assessed by Woodcock and Strahler (1987) and Marceau (1994). Both papers concluded that satellite remote sensing was more appropriate for mapping broad vegetative categories rather than distinct vegetative communities. Although higher resolution imagery is now available, the high cost of acquiring and processing these data precludes their use for large regions. Yet, high-resolution data derived from aerial photography also is constrained by the high cost of acquiring, interpreting, and digitizing data. Therefore, highly resolved maps of riparian areas are often only available for portions of the streams, as opposed to the entire contributing watershed. Muller (1997, p. 419) states, “The scale factor is of prior importance for studying riparian vegetation,” yet this factor is seldom taken into account in ecological assessments of riparian zone function. Our objectives in this chapter are to (1) quantify the potential errors and sources of uncertainty in the use of the NLCD data for mapping land use and cover in riparian zones in agricultural landscapes, (2) examine the scale effects with respect to changing grain and extent in these mapping exercises, (3) compare these phenomena in stream versus wetland riparian ecosystems, and (4) assess these differences in landscapes whose matrix is dominated by agriculture versus those with a more diverse matrix.

Quantifying scale effects, hierarchical linkages, and resulting uncertainty associated with mapping and analyzing riparian zones is necessary for understanding the structure and composition of riparian areas along streams and wetlands. An understanding of the underlying processes that influence riparian zones and control water quality is of fundamental interest not only to researchers and resource managers, but also to policy makers and private landowners who must evaluate the potential consequences of land management practices under different scenarios.

15.2 METHODS

We quantified the effects of data source grain and extent associated with the analysis of stream and wetland riparian areas by comparing land cover summaries from two commonly used sources: photo-interpreted land cover from large-scale aerial

photography (0.5 m resolution) and Landsat-based land cover (30 m resolution) from the NLCD. The stream analyses were based on a set of 12 watersheds in southeastern Minnesota; a parallel analysis was conducted on 36 central Minnesota wetlands (Figure 15.1). Stream sites were located in southeastern Minnesota, a region dominated by intensive rowcrop agriculture; wetland sites were located in central Minnesota, where land use is more diverse. Streams ranged from the first to the third order with baseflow width of 3.4 to 11.2 m with a mean of 5.4 m. The total length of streams within watersheds ranged from 8.3 to 40.4 km with a mean of 26.6 km. Sites were chosen to reflect a range of land uses within the region as part of a larger study to identify indicators of stream condition. Wetlands types were palustrine emergent (Cowardin 1979), and ranged in size from 0.3 to 9 ha with a mean of 1.9 ha. These sites also were chosen to reflect the gradient of land uses, as part of another study identifying indicators of wetland condition.

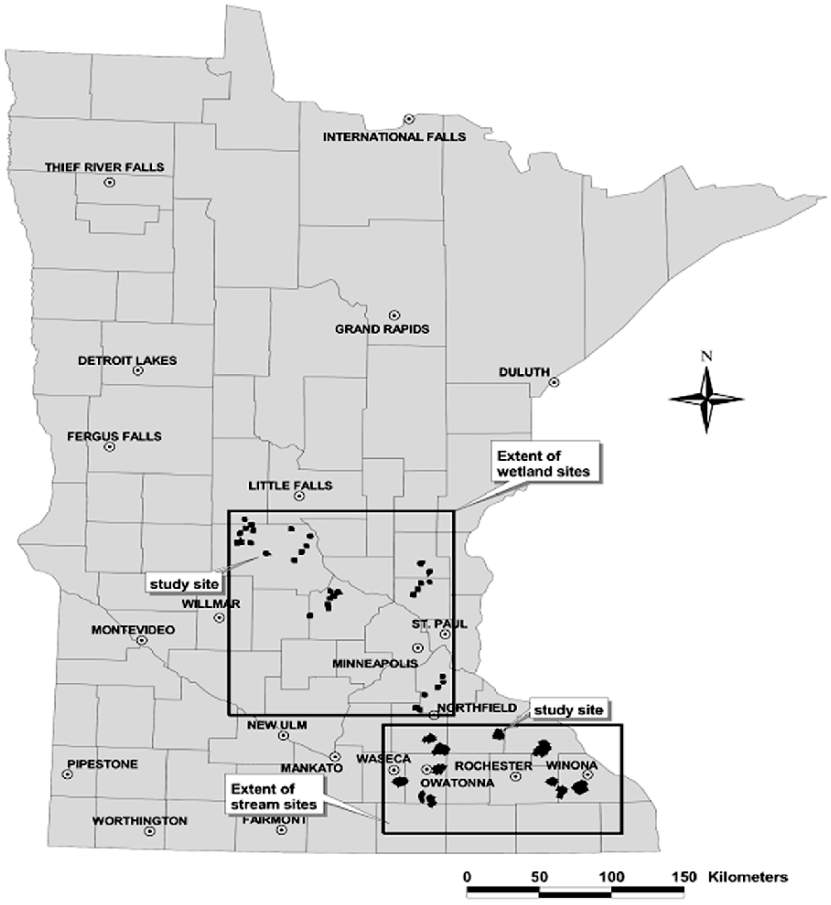


Figure 15.1. *Distribution of stream and wetland study areas.*

Streams and wetlands were buffered to 500 m, and land cover within this buffer was visually interpreted and digitized on-screen based on rectified aerial photo mosaics. The aerial photos were medium scale (1:15,880) color infrared (CIR) imagery obtained from the Minnesota DNR Department of Forestry. Most of the CIR photography was taken during the fall, and ranged in year from 1994-1997. Photos were scanned at 800 dpi and georeferenced to USGS Digital Orthophotos. Stereo pairs were used for map interpretation, and map polygons were screen-digitized using ArcView GIS. Polygons as narrow as 2 m were identified with a minimum map unit of 100 m^2. Polygons were classified according to a modified Anderson Level II classification (Anderson et al. 1976) also used by the NLCD. A subset of interpreted map polygons was visited in the field to assess the consistency of the classification.

Land cover from the NLCD data was clipped to the 500 m map extent and resampled from 30 m to 0.5 m to match the resolution of the air-photo data. This allowed us to compare the two data sets in a spatially explicit manner. For this part of Minnesota, the NLCD was classified from Landsat imagery collected between 1989 and 1994. Land cover classes for both the NLCD data and the air-photo land cover were aggregated to provide a common 9-class land cover scheme (Table 15.1). The two sets of land cover data were then summarized for a series of incremental buffers (5, 10, 30, 60, 90, 120, 150, 200, 300, 400, and 500 m) around each stream or wetland.

To understand bias and scale effects associated with mapping riparian land cover data, we compared proportions from the NLCD and photo-interpreted data for the cover classes within each buffer width**.** Land cover proportions were plotted across the range of buffer widths to quantify the type and magnitude of differences between the data types as well as the effect of changing spatial extent with increasing distance from the stream or wetland. Comparisons between the photo and the NLCD data were performed by calculating the relative proportional differences between photo and NLCD data for each buffer width (e.g., [proportion of LU photo – proportion of LU NLCD]/ proportion of LU photo). To assess the potential effects of differences in the composition of the landscape matrix we performed these analyses independently for: (1) three predominately agricultural watersheds (“Most-Ag” – defined as the three streams with the highest proportion of agricultural land within the 500 m buffer, ranging from 69-85%), (2) three watersheds with the lowest proportion of agricultural land use in the 500 m buffer (“Least-Ag” – ranging from 36-40%), and (3) all 12 watersheds combined. A similar stratification was conducted for wetland sites (“Most-Ag” – ranging from 57-73%; “Least-Ag” – ranging from 0-5%). To quantify errors of commission (i.e., erroneously classifying a given cover type, e.g., classifying a pixel as agriculture when it is grassland) and errors of omission (i.e., failure to classify a given cover type, e.g., failure to classify residential land use properly) we performed pixel-to-pixel comparisons of the two data sets. Error estimates for 5, 30, 90, and 500 m buffers were calculated separately for the 12 streams and 36 wetlands to understand how the NLCD varied from the more highly resolved photo data, and how these differences varied with spatial extent of the analysis. For the purpose of our analyses we consider the photos to be

the more accurate data source, or ground-truth, recognizing that there is underlying error in both data sets and possible errors due to registration.

Table 15.1. *Cross walk of land cover classifications with Anderson classification.*

Photo Class	*NLCD Class*	*Common Class / Anderson*	*Anderson Level*
Contour Crop	Row Crops	Agriculture (ARC)	I
Row Crop	Row Crops	Agriculture (ARC)	I
High Intensity	Commercial/Industrial/ Transport	Commercial (COM)	II
Low Intensity	Commercial/Industrial/ Transport	Commercial (COM)	II
Large Farm Operation	Commercial/Industrial/ Transport	Commercial (COM)	II
Highway (Paved)	Commercial/Industrial/ Transport	Commercial (COM)	II
Livestock	Commercial/Industrial/ Transport	Commercial (COM)	II
Poultry	Commercial/Industrial/ Transport	Commercial (COM)	II
Cattle	Commercial/Industrial/ Transport	Commercial (COM)	II
Confined Feeding	Commercial/Industrial/ Transport	Commercial (COM))	II
Gravel Pit/Mine	Quarries/Strip Mines/Gravel Pi	Commercial (COM)	I
Coniferous	Evergreen Forest	Forest (FO)	I
Deciduous	Deciduous Forest	Forest (FO)	I
Mixed Forest	Mixed Forest	Forest (FO)	I
Shrub	Shrubland	Forest (FO)	I
Grass	Grasslands/Herbaceous	Grass (FR)	I
Forage Crop	Pasture/Hay	Grass (GR)	I
Hay	Pasture/Hay	Grass(GR)	I
Pasture	Pasture/Hay	Grass (GR)	I
Golf Course	Urban/Recreational Grasses	Grass (GR)	I
Right Of Way	Urban/Recreational Grasses	Grass (GR)	I
Grave Yard	Urban/Recreational Grasses	Grass (GR)	I

Table 15.1 (contd.)

Photo Class	*NLCD Class*	*Common Class / Anderson*	*Anderson Level*
Low Intensity	Low Intensity Residential	Low Intensity Residential (RLI)	II
Small Farm Stead	Low Intensity Residential	Low Intensity Residential (RLI)	II
Secondary Road (Paved)	Low Intensity Residential	Low Intensity Residential (RLI)	II
Secondary Road (Un-Paved)	Low Intensity Residential	Low Intensity Residential (RLI)	II
Local Road (Paved)	Low Intensity Residential	Low Intensity Residential (RLI)	II
Local Road (Un-Paved)	Low Intensity Residential	Low Intensity Residential (RLI)	II
Access Road (Paved)	Low Intensity Residential	Low Intensity Residential (RLI)	II
Access Road (Un-Paved)	Low Intensity Residential	Low Intensity Residential (RLI)	II
Rail Road	Low Intensity Residential	Low Intensity Residential (RLI)	II
Robust Emergent	Emergent Herbaceous Wetlands	Emergent Wetland (PEM)	II
Non-Persistent Emergent	Emergent Herbaceous Wetlands	Emergent Wetland (PEM)	II
Bog/Fen	Emergent Herbaceous Wetlands	Emergent Wetland (PEM)	II
Aquatic Bed	Emergent Herbaceous Wetlands	Emergent Wetland (PEM)	II
Wet Meadow	Emergent Herbaceous Wetlands	Emergent Wetland (PEM)	II
Pasture	Emergent Herbaceous Wetlands	Emergent Wetland (PEM)	II
Reed Canary Grass	Emergent Herbaceous Wetlands	Emergent Wetland (PEM)	II
Scrub-Shrub	Woody Wetlands	Forested Wetland (PFO)	II
Forested	Woody Wetlands	Forested Wetland (PFO)	II
Pond, Small Stream	Open Water	Open Water (POW)	I
Lake	Open Water	Open Water (POW)	I
River	Open Water	Open Water (POW)	I
High Intensity	High Intensity Residential	High Intensity Residential (RHI)	II

To determine if there were scale effects related to grid cell size with respect to estimates of land cover proportions, we also resampled the photo classification at a series of cell sizes ranging from 0.5 to 120 m and recalculated the land cover

proportions. This analysis was repeated for all buffer widths to determine if any scale effects related to grain sizes were sensitive to the spatial extent of the data.

15.3 RESULTS

15.3.1 NLCD to Air-Photo Comparisons – Streams

In the Most-Ag watersheds the NLCD classification essentially did not recognize the stream or its buffer (Figure 15.2A); thus, the classified land use for the 5-10 m buffers did not differ from the 100-500 m buffers. The NLCD data reported 85% rowcrop agriculture (ARC) and approximately 10% grassland (GR) at all buffer widths. In contrast, the air-photo classification reported 10-16% agriculture at the 5-10 m buffer widths, increasing to a maximum of 78% in the 500 m buffer (Figure 15.2B). As a result, when compared to the photointerpreted data, the NLCD data over-represented rowcrop agriculture by between 40-74% within one pixel width (~30 m) of the stream, diminishing to a difference of only 7% at 500 m (Figure 15.3A). Commercial land use (COM) comprises a very small proportion of the landscape; however, the mapping bias for this land use was very similar to that of agricultural land use (Figure 15.3A).

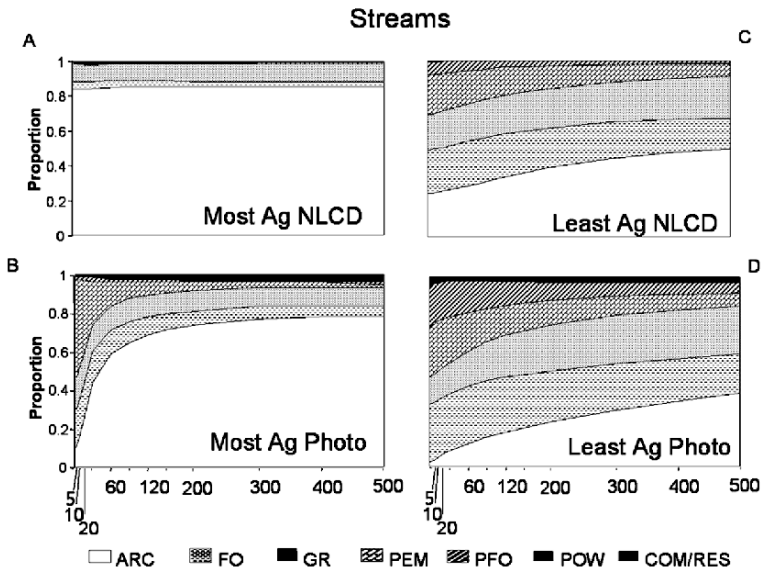


Figure 15.2. *Land cover proportions surrounding streams in the 3 most and 3 least agricultural areas (RLI and RHI have been combined with COM for this figure). (A) Most Ag NLCD; (B) Most-Ag Photo; (C) Least-Ag NLCD; (D) Least-Ag Photo.*

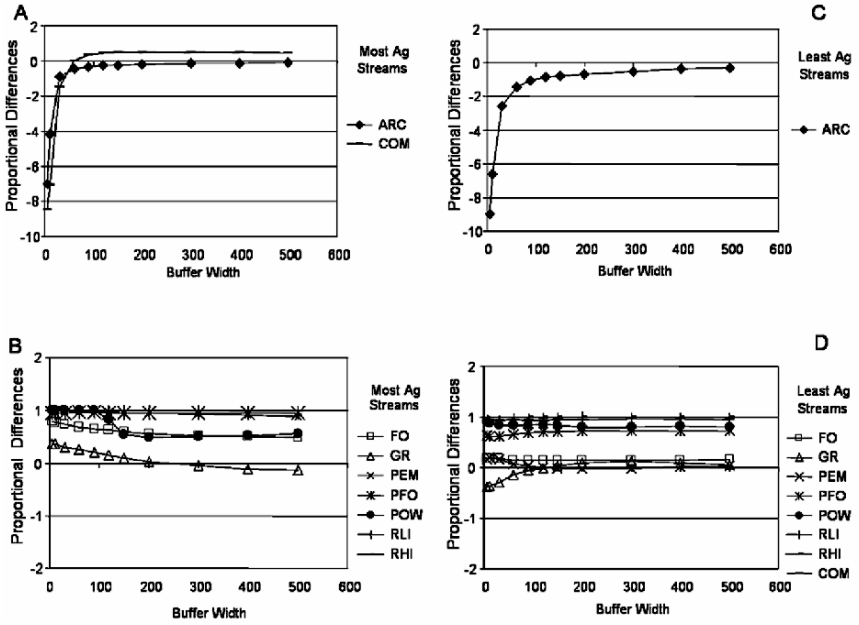


Figure 15.3. *Relative proportional differences between air photo and NLCD within Most-Ag and Least-Ag streams (see Methods for calculations). (A) Most-Ag streams; (B) Most-Ag streams; (C) Least-Ag streams; (D) Least-Ag streams.*

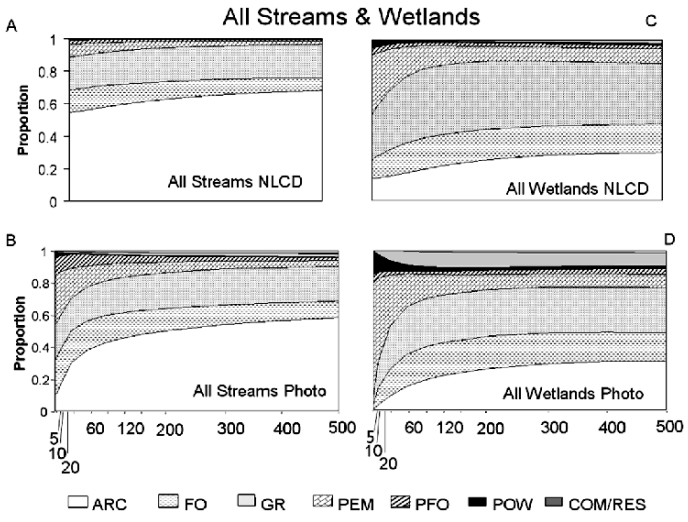


Figure 15.4. *Land cover proportions surrounding all streams and all wetlands (RLI and RHI have been combined with COM for this figure). (A) All Streams NLCD; (B) All Streams Photo; (C) All Wetlands NLCD; (D) All Wetlands Photo.*

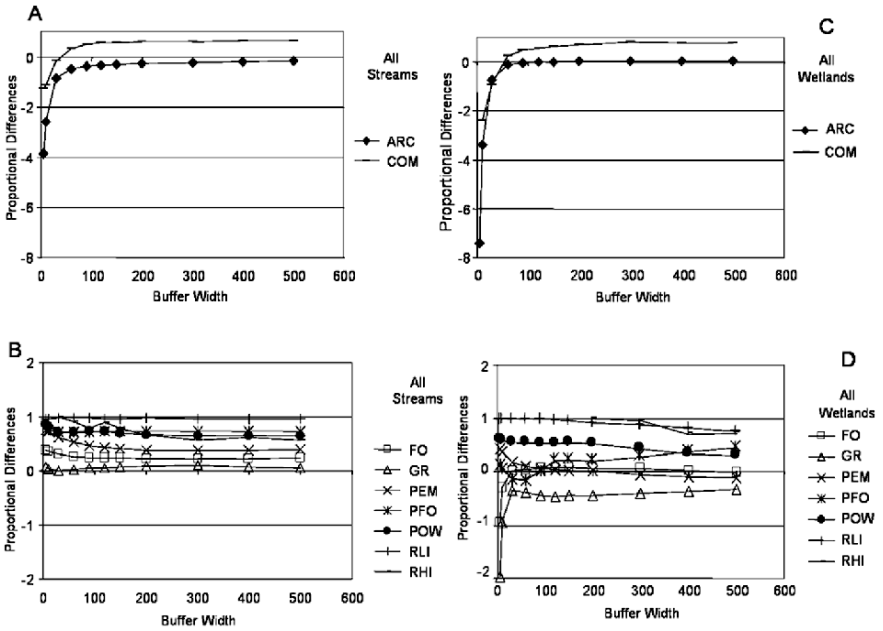


Figure 15.5. *Relative proportional differences between air photo and NLCD within all streams and wetlands (see methods for calculations). (A) Streams; (B) Streams; (C) Wetlands; (D) Wetlands.*

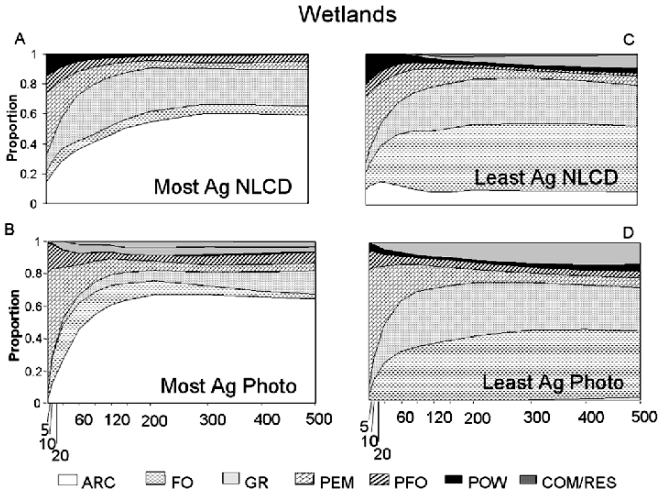


Figure 15.6. *Land cover proportions surrounding wetlands in the 3 most and least agricultural areas (RLI and RHI have been combined with COM for this figure). (A) Most-Ag NLCD; (B) Most-Ag Photo; (C) Least-Ag NLCD; (D) Least-Ag Photo.*

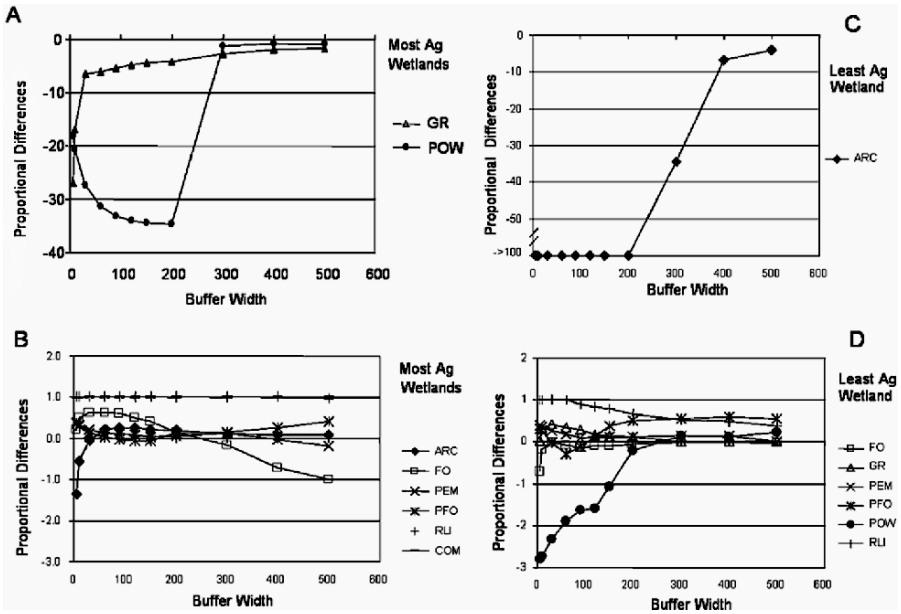


Figure 15.7. *Relative proportional differences within most and least agricultural wetlands (see methods for calculations). (A) Most-Ag wetlands: Grass, open water; (B) Most-Ag wetlands: agriculture, forest, grassland, emergent wetland, forested wetland, open water, low-density residential, high-density residential; (C) Least-Ag wetlands: agricultural land use; (D) Least-Ag wetlands: forest, grassland, emergent wetland, forested wetland, open water, low-density residential, high-density residential.*

All other land use classes were under-represented in the NLCD relative to the photo in these Most-Ag watersheds. In the air-photo classification, slightly more than 50% of the riparian land cover within the 10 m buffer was classified as a permanent wetland (PEM, PFO) (Figure 15.2B), but less than 1% of that buffer was classified as wetland by the NLCD (Figure 15.2A). The proportion of grassland was fairly similar in the air-photo and NLCD classification (10% versus 16%; Figure 15.2A, B), but the air-photo classification reported more forest (FO), particularly within the first 30 m (15-20%; Figure 15.2B) compared to the NLCD (4%; Figure 15.2A).

In Least-Ag watersheds, where no single land use type dominated the matrix, the NLCD overestimated the amount of agricultural land, by more than 8 times (Figure 15.3C), classifying the 10 m buffer as ~25% agricultural (Figure 15.2C), compared with approximately 3% in the air-photo data (Figure 15.2D). Although low-density residential (RLI) and commercial areas represent a very small proportion of the watershed as a whole, the NLCD data completely missed these land uses (Figure 15.2C, 2D, 3D). Unlike the Most-Ag watersheds, the NLCD and photo classifications of permanent wetlands in Least-Ag watersheds were fairly similar.

Proportions of forest and grasslands for both classifications also were within 5% of one another across all buffer widths (Figure 15.3D).

Across all twelve streams, differences between the photo and the NLCD were dampened; the greatest differences were observed within the first 10 m (Figure 15.4A, 4B), with agriculture being over-represented in the NLCD by 2-4 times relative to the photo (Figure 15.5A). Grasslands, followed by forests, exhibit the smallest differences between the two data sets across all buffer widths (Figure 15.5B).

15.3.2 NLCD to Air-Photo Comparisons – Wetlands

The wetlands study area had slightly less agricultural landuse than the stream study area at the scale of the 500 m buffers (65% versus 78% measured in photos). In wetland riparian zones within Most-Ag areas, grassland and open water (POW) classes displayed the largest discrepancies between the two data sources (Figure 15.6A, 6B). Grassland was greatly over-represented in the NLCD within the first pixel width (~30 m), while open water was over-represented out to 200 m and then converged with NLCD at buffers greater than 200 m (Figure 15.7A). Agricultural land use within the 5-10 m buffers was also over-represented, but to a much smaller extent than in streams, and the proportions converged beyond 30 m**.** Forest cover was generally under-represented close to the wetland, and over-represented at buffers greater than 300 m from the wetland (Figure 15.7B). Low-density residential and commercial land uses were consistently under-represented by the NLCD across all buffer widths.

In the Least-Ag wetlands essentially no agricultural land use was observed in the air-photo classification, at buffer widths less than 300 m (Figure 15.6D), while the NLCD showed approximately 10-15% agricultural land cover across all buffer widths (Figure 15.6C). Open water also was over-represented in the NLCD relative to the photo within the first 200 m of the wetland (Figure 15.6C, 6D, 7D). Low-density residential land was not detected in the NLCD data in the first 90 m surrounding the wetlands, and thereafter was under-represented, relative to the photo data (Figure 15.7D). Forest cover was also under-represented in the NLCD data within the first 5 m of the wetland, but converged with the photo data beyond that distance.

Across all 36 study wetlands, the NLCD data were relatively similar to the photos (Figure 15.4C, 4D), with the exception that NLCD over-represented agriculture by approximately 7 times within the 5 m buffer, decreasing to 3 times within the 10m buffer (Figure 15.5C). Grassland also was over-represented in the NLCD for the first 10 m surrounding the wetlands (Figure 15.5D).

15.3.3 Effects of Changing Spatial Extent

Differences between the photo and NLCD classifications for most land use types diminish with increasing buffer distances, albeit to different magnitudes for the two ecosystem types and landscape matrix types (Figures 15.3, 15.5, 15.7). A notable

exception is open water and forest cover in Most-Ag wetlands (Figure 15.7A, 7B). Emergent wetlands are underestimated by NLCD at approximately the same rate as forested wetlands within the first 30 m, but beyond that emergent wetlands converge with NLCD, while forested wetlands remain under-represented across all buffer extents.

In Most-Ag streams, areal measures of agricultural land use for both data sets converge at around 60 m, although nearly a 10% difference (relative to the photo data) persists beyond that point (Figure 15.2A, 2B). In the Least-Ag watersheds differences between the two data sets diminish rapidly within the first 60 m, but do not converge to the same extent (Figure 15.2C, 2D, 3C). Relative to the photo, the NLCD continues to over-estimate agricultural land use by 30% at the largest spatial extent (500 m). Responses across the 12 study watersheds are intermediate; relative to photos, NLCD measurements of agricultural land use is over-estimated by 17% at the largest spatial extent (Figures 15.4A, 4B, 5A).

15.3.4 Classification Error

Classification concordance between the photo and NLCD were low across all buffer widths and both ecosystem types (Figure 15.8). Not surprisingly, classification differences were greatest at the sub-pixel level, but persisted to the 500 m buffers. Overall errors were similar for wetlands and streams at the 5 m buffer width, but at wider buffers error rates were higher for wetlands than streams (Figure 15.8).

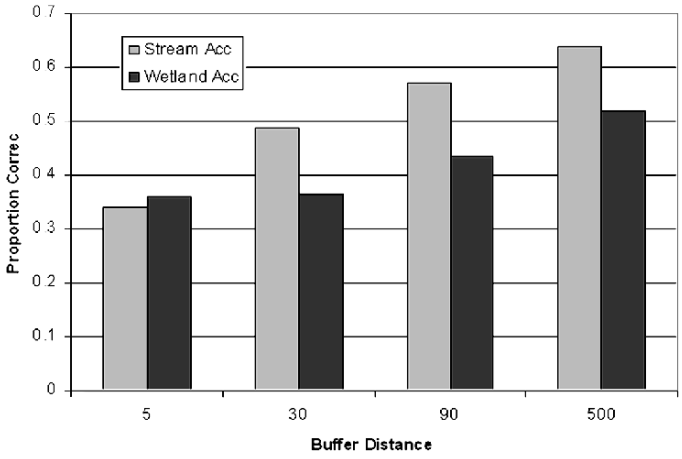


Figure 15.8. *Overall mapping accuracy of NLCD data for 30- and 500- m buffers for all stream and wetland study sites, based on a comparison with airphoto data.*

In the stream analyses, classification differences were high across all scales, but highest at narrow (30 m or less) buffer distances (Table 15.2). A pixel-by-pixel comparison of classification errors shows that, while 81% of pixels that were agricultural in the air-photos were correctly classified in the NCLD data, 37-55% of

pixels that were commercial, emergent wetland, grassland, and low-density residential, and 100% of high-density residential pixels were erroneously classified as agriculture by NLCD (Table 15.2). This effect persisted at greater spatial extents; in the 500 m buffer, 40-50% of grassland, emergent wetland, high-density residential (RHI) and commercial were still misclassified as agriculture (Table 15.2). Not surprisingly, agriculture and grassland are most frequently confused with one another in the NLCD.

Table 15.2. *Error analysis. Column values represent errors of omission (a land cover type classified as something else); row values represent errors of commission (other classes classified as that particular land cover type). Values on the diagonal represent the proportion that is correctly classified. Abbreviations can be found in Table 15.1.*

30 meter buffer - Streams

NLCD	ARC	FO	GR	PEM	PFO	POW	RLI	RHI	COM	Com-mission
ARC	**0.81**	0.31	0.43	0.55	0.31	0.32	0.49	1.00	0.37	**3.76**
FO	0.03	**0.42**	0.05	0.05	0.37	0.12	0.07	0.00	0.09	**0.77**
GR	0.14	0.22	**0.49**	0.11	0.10	0.20	0.32	0.00	0.47	**1.57**
PEM	0.01	0.01	0.02	**0.26**	0.08	0.22	0.01	0.00	0.00	**0.33**
PFO	0.00	0.04	0.00	0.02	**0.15**	0.01	0.01	0.00	0.00	**0.08**
POW	0.00	0.00	0.00	0.00	0.00	**0.13**	0.00	0.00	0.00	**0.01**
RLI	0.00	0.00	0.00	0.00	0.00	0.00	**0.01**	0.00	0.00	**0.00**
RHI	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**0.00**	0.00	**0.00**
COM	0.00	0.00	0.01	0.00	0.00	0.00	0.09	0.00	**0.08**	**0.11**
Omis-sion	**0.19**	**0.58**	**0.51**	**0.74**	**0.85**	**0.87**	**0.99**	**1.00**	**0.92**	

500 meter buffer - Streams

	ARC	FO	GR	PEM	PFO	POW	RLI	RHI	COM	Com-mission
ARC	**0.83**	0.27	0.52	0.47	0.27	0.27	0.48	0.04	0.40	**2.71**
FO	0.02	**0.47**	0.04	0.06	0.35	0.14	0.06	0.01	0.05	**0.72**
GR	0.14	0.24	**0.42**	0.15	0.13	0.20	0.37	0.09	0.47	**1.79**
PEM	0.00	0.01	0.01	**0.30**	0.14	0.13	0.00	0.00	0.00	**0.30**
PFO	0.00	0.01	0.00	0.01	**0.11**	0.01	0.00	0.00	0.00	**0.04**
POW	0.00	0.00	0.00	0.00	0.00	**0.26**	0.00	0.00	0.00	**0.00**
RLI	0.00	0.00	0.00	0.00	0.00	0.00	**0.03**	0.43	0.00	**0.44**
RHI	0.00	0.00	0.00	0.00	0.00	0.00	0.01	**0.26**	0.00	**0.01**
COM	0.00	0.00	0.01	0.00	0.00	0.01	0.04	0.17	**0.07**	**0.24**
Omis-sion	**0.17**	**0.53**	**0.58**	**0.70**	**0.89**	**0.74**	**0.97**	**0.74**	**0.93**	

Table 15.2 (contd.)

30 meter buffer - Wetlands

	ARC	FO	GR	PEM	PFO	POW	RLI	RHI	COM	Com-mission
ARC	**0.30**	0.07	0.22	0.12	0.13	0.14	0.17	-	0.00	**0.85**
FO	0.07	**0.46**	0.10	0.11	0.19	0.05	0.15	-	0.04	**0.71**
GR	0.38	0.25	**0.44**	0.22	0.17	0.31	0.53	-	0.96	**2.82**
PEM	0.18	0.17	0.22	**0.45**	0.28	0.45	0.14	-	0.00	**1.44**
PFO	0.06	0.04	0.01	0.06	**0.17**	0.01	0.01	-	0.00	**0.19**
POW	0.01	0.01	0.00	0.05	0.06	**0.04**	0.00	-	0.00	**0.13**
RLI	0.00	0.00	0.00	0.00	0.00	0.00	**0.00**	-	0.00	**0.00**
RHI	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-	0.00	**0.00**
COM	0.00	0.00	0.00	0.00	0.00	0.00	0.01	-	**0.00**	**0.01**
Omis-sion	**0.70**	**0.54**	**0.56**	**0.55**	**0.83**	**0.96**	**1.00**	-	**1.00**	

500 meter buffer - Wetlands

	ARC	FO	GR	PEM	PFO	POW	RLI	RHI	COM	Com-mission
ARC	**0.62**	0.06	0.29	0.10	0.09	0.06	0.13	0.01	0.17	**0.89**
FO	0.04	**0.61**	0.08	0.11	0.33	0.09	0.13	0.11	0.16	**1.05**
GR	0.31	0.23	**0.54**	0.18	0.16	0.10	0.57	0.03	0.63	**2.21**
PEM	0.03	0.04	0.07	**0.54**	0.18	0.19	0.03	0.00	0.02	**0.56**
PFO	0.00	0.02	0.01	0.04	**0.23**	0.03	0.01	0.00	0.00	**0.12**
POW	0.00	0.00	0.00	0.02	0.01	**0.53**	0.00	0.08	0.00	**0.12**
RLI	0.00	0.03	0.01	0.00	0.00	0.00	**0.12**	0.70	0.00	**0.75**
RHI	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**0.07**	0.00	**0.00**
COM	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	**0.02**	**0.02**
Omis-sion	**0.38**	**0.39**	**0.46**	**0.46**	**0.77**	**0.47**	**0.88**	**0.93**	**0.98**	

The NLCD classification around wetlands was less concordant with the air-photo classification than around streams. Only emergent wetlands had higher classification concordance in wetlands than streams, for both the 5 m and the 500 m buffer widths (Table 15.2). At 30 m, 44-46% of forest cover, grassland, and emergent wetland were classified similarly between the photo and NLCD. Within the 30 m buffer, agriculture was correctly classified by the NLCD only 30% of the time. The NLCD under-represented the amount of agriculture in this narrow buffer, confusing these areas with grasslands 38% of the time. At 500 m, classification concordance improved substantially for all classes except emergent wetlands and commercial land. Correct classifications ranged from 53% to about 62% for agricultural land use, forest, grassland, emergent wetland and open water. Agricultural land use was misclassified as grassland 29% of the time by NLCD.

15.3.5 Grain Size Effects

Although these differences between the NLCD data and the photo data might be attributed to scaling up to a larger grain size, we don't believe they are. After recalculating land cover proportions for the resampled land cover maps (essentially scaling up the photo data), we found that, in general, the estimates of proportional land cover were quite insensitive to grid cell size – there was very little change in the proportions of land cover, even between the extremes of grid cell sizes. This insensitivity to grain size was consistent across all buffer distances (Figure 15.9). We did observe a few instances in which classes such as light residential, which were initially present in low abundance (<1% of the landscape), "disappeared" as the aggregation process caused these types to be subsumed into more dominant classes.

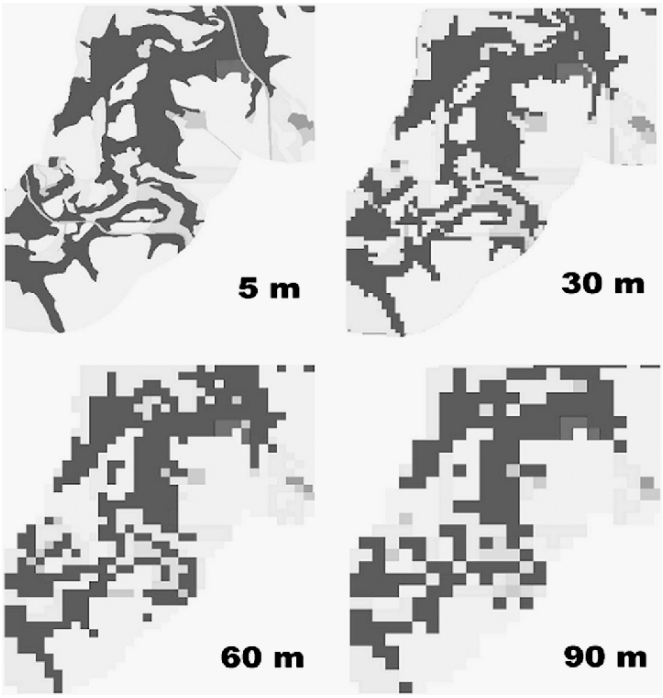


Figure 15.9. *Visual effects of resampling aerial photo derived land cover data from 0.5 m to 30, 60 and 90 m.*

15.4 DISCUSSION

Many ecological assessments employ remotely sensed data such as Landsat or other space-borne sensors to quantify anthropogenic stress factors, such as land use, human-induced changes in landscape pattern, or alteration of riparian zones. These data cover broad geographic extents, are collected with relative frequency, and have

been shown to be good predictors of ecological health or impairment. It is uncommon, however, to consider the uncertainty resulting from classification errors or spatial resolution as part of the assessments, or in the use of these data as inputs to predictive models (Li and Wu, Chapter 3). Even more rarely are these errors analyzed in a spatially explicit manner, as we did here. These errors appear to be particularly problematic in quantifying the composition and spatial structure of riparian ecosystems, which have unique attributes that make them ecologically important but difficult to quantify with traditional Landsat-based classifications such as the NLCD.

In this study we developed spatially explicit assessments to determine if accuracy varies spatially according to specific features in the landscape, particularly wetlands, streams and the riparian zones near them. Our results show that, in the Midwest, accuracy does indeed vary spatially in relation to these important landscape features. This uncertainty results from challenges inherent in the classification of satellite imagery as well as landscape context and structure of the area under consideration.

15.4.1 Classification Issues

Although state level NLCD classification accuracies are not yet available for Minnesota, accuracy assessments are available for the eastern United States (Stehman et al. 2003). Overall, the classifications generally meets the standards, established by Anderson et al. (1976), at Level I, but are inadequate for Level II. Nearly half of the land cover classes we used were Level I, which had significantly better accuracies (Table 15.1).

While the 30 m resolution data has proven to be effective for watershed-scale analyses (Johnson and Gage 1997, Hernandez et al. 2000, Jones et al. 2001), the present analyses suggest that it is relatively ineffective for quantifying fine-scale features such as riparian vegetation. Congalton et al. (2002) also found that there were large discrepancies between air-photo and Landsat TM classifications of riparian vegetation structure that were attributed to the inherent diversity of riparian vegetation, the linear arrangement of riparian zones, and the displacement of stream channels between their two data sources. Essentially, these discrepancies result from differences in grain size (resolution), map registration, temporal differences, classification method (manual vs. statistical), classification level, and classification error.

Riparian zones are highly diverse with respect to soil, moisture and topography and have diverse vegetative communities that vary at sub-pixel scales. Superimposing a raster grid over these narrow features potentially causes the stream itself, riparian vegetation and upland vegetation to be included in an individual pixel, resulting in classification error. This effect may be compounded by registration errors. In addition, supervised or unsupervised image classifications operate on an individual pixel basis. While post-classification filtering is often used to reduce the number of isolated pixels, satellite classification procedures are fundamentally different than manual interpretation of air photos, in which the

interpreter can use adjacency, contextual and other visual cues to generate polygon-based classifications.

Potential annual and inter-annual differences in wetland extent, as well as differences in both the width and location of a stream channel between the photo and the NLCD dates may also contribute to classification differences. Finally, similarities in the spectral response of certain land cover types (e.g., grasslands, crops) present a technical challenge that is independent of the resolution of the data.

Using a larger range of grid cell sizes for testing grain size effects, Turner et al. (1989) reported that any land cover present in less than 50% would ultimately be lost at coarser resolutions. At the range of grid cell sizes in our study, however, this scale effect was quite weak. A key point made by Turner et al. (1989) and also by Saura (2002) is that the rate at which land cover types are lost with increasing grain size is controlled by the spatial arrangement of cover types – specifically, features with a high level of contagion disappear at a slower rate than more disaggregated types. The continuous linear nature of riverine wetlands and other riparian features may contribute to the persistence of these types with increasing grain size.

Other landscape metrics, such as dominance and contagion (Turner et al. 1989) or mean patch size and landscape shape indices (Wu 2004) do show predictable sensitivities to grid cell size. Turner et al. (1989) reported a decrease in dominance and contagion with increasing grid cell size, and that this response was a stair-step function that was strongly related to changes in the numbers of land cover classes. Our resampling exercise would likely show similar trends in patch shape, edge and other metrics (Figure 15.9), but it is interesting to note that proportional abundances of riparian cover types were preserved when grid cell sizes were changed over approximately three orders of magnitude. This relative insensitivity to grain size implies that classification accuracy and map extent are more important in modeling riparian attributes than map grain.

15.4.2 Landscape Context

In a largely agricultural region in the Midwestern US, we found that there were large differences in the ability of the NLCD and air-photos to classify land use in riparian zones, but these differences were exacerbated when the landscape matrix was dominated by rowcrop agriculture. The ecosystem type (wetland versus stream) around which the riparian zone was examined also influenced the direction and magnitude of the discrepancies between the NLCD and the air-photos.

Riparian zones are inherently variable ecosystems, due to the gradients in soils, moisture, and topography that control vegetation patterns, as well as disturbance type and frequency (Gregory et al. 1991, Décamps 1996, Naiman et al. 2000). As a result, riparian vegetation communities are very diverse and therefore difficult to map accurately using satellite image classification methods (Muller 1997, Congalton et al. 2002). There is also a regional component underlying the inability of Landsat imagery to accurately map riparian land use and cover. In the Midwestern US, the headwaters of small streams (not originating in wetlands) are highly accessible to mechanized equipment and are therefore vulnerable to anthropogenic disturbances including channelization and vegetation removal (L. Johnson, personal observations). Intensive

land use practices in these agricultural areas frequently result in removal of woody riparian vegetation and conversion to herbaceous plant communities or narrow wooded buffer strips on stream banks with row crops planted to the edge of the bank. The intact patches of forest vegetation that remain vary in size and composition depending on stewardship practices of private landowners. As a result, many riparian buffer strips are much narrower than the 30 m resolution of Landsat imagery.

Primarily as a result of the narrow configuration of the riparian buffer strips, the magnitude of classification error was striking. Since confusion exists among both land uses that might ameliorate the problems associated with agricultural land use (emergent wetlands, grasslands and forest), and those that exacerbate it (high-intensity residential and commercial), these classification errors have the potential to generate additional uncertainty when modeling ecosystem processes.

The fact that commonly accessible satellite imagery does not accurately map riparian zones has large implications for researchers and managers. Ecological processes across the riparian zone occur at a number of spatial scales, many of which are finer than the spatial resolution of Landsat sensors. The immediate stream/land interface is characterized by a number of material transfer processes, including inputs of groundwater and overland flow, associated transports of sediments and nutrients, and delivery of fine organic matter. Expanding this interface to the 5-10 m buffer includes processes such as shading, delivery of coarse organic matter, entrainment of particulates, and use of the buffer as a migration corridor. Clearly, misrepresentation of these immediate features due to classification error will increase the uncertainty in predictions of how stream biological, chemical, and physical properties respond to land use, as well as subsequent management policies created to sustain riparian function. Based on our analyses, NLCD data would overestimate the amount of rowcrop agriculture and underestimated other land cover classes in this heavily agricultural landscape. As a result, models predicting the potential effect of rowcrop agriculture on in-stream processes, such as primary production or eutrophication, would not accurately reflect the effects of land use and land cover, since riparian vegetation that could potentially provide shade, or ameliorate nutrient and sediment inputs would not have been evident.

Land cover types within riparian zones have adjacency relationships that are dissimilar to those in other parts of the watershed. These relationships are due to both landscape physiognomy (e.g., topographic conditions favor the development of wetlands along streams) and human activities (farmers may be less likely to clear riparian forests for agriculture). The present study indicates that the NLCD scale of classification does not effectively capture these relationships. Since there are specific ecological or hydrological processes associated with these adjacent systems, such as the wetland's role in retaining sediments, not accounting for these elements would likely result in overestimating the amount of environmental stress on the system. This could then lead to overestimating the ability of a stream or wetland ecosystem to recover from that stress.

15.4.3 Landscape Structure

The discrepancies shown above are due in part to the spatial structure of these land cover types. The riparian systems we studied were embedded in an agricultural region where row crop agriculture constituted 56% and 49% of the study areas for streams and wetlands, respectively. The boundaries of agricultural fields are typically defined in two ways – as topographic breaks, which often correspond to boundaries of riparian systems, and ownership or road boundaries. For economic and logistic reasons, patches in agricultural land use tend to be large and have fairly simple shapes. The mean sizes of agricultural patches were 12.2 and 28.6 ha in the stream and wetland systems, respectively; mean patch sizes for all other types combined ranged from 0.4 to 5.9 ha. The lowest perimeter-area ratios, an index of shape complexity, occurred in the agricultural, commercial, and high-intensity residential land use types – all these were less than 0.06 m m^{-2}, compared with the range of 0.25 to 0.46 m m^{-2} for forests, permanent wetlands, and grasslands. The raster nature and resolution of TM satellite imagery makes it poorly suited to identifying small features with complex boundaries (Benson and MacKenzie 1995, Moody and Woodcock 1995), which characterize wetlands and forest systems within riparian zones. As the proportion of agriculture increases as it did from the Least- to Most-Ag study sites, these patterns should become even more pronounced.

Many anthropogenically-modified landscapes exhibit a certain level of self-similarity across some range of scales. Krummel et al. (1987), for example, found that human activities structure the landscape at grain sizes up to 73 ha; forest patches in the range of 28-73 ha tend to have rectangular or simple shapes related to land use patterns or survey and township divisions. Above this range, topographic and hydrological patterns produce patches that have higher fractal dimensions (i.e., more complex shapes). Our results indicate that the topography and hydrology generate complex patterns at finer scales as well – the immediate patterns observed adjacent to stream and wetland systems. These riparian systems are fundamentally different from the agricultural matrix in both composition and spatial structure. This is related again to the interactions of physiognomy and human land use within riparian zones, which promotes the formation of narrow features with complex boundaries oriented in parallel to the stream channel or concentrically around open-water wetlands.

15.5 CONCLUSIONS: IMPLICATIONS FOR SCALING

Remotely sensed data from aerial photography and space-borne sensors have had wide applications in watershed assessments. In previous studies, we have used 1-ha resolution air-photo data to quantify variation in water chemistry (Johnson et al. 1997), stream macroinvertebrates and habitat in agricultural landscapes (Richards et al. 1996), and 16 ha resolution Land Use/Land Cover (LUDA) data for quantifying variation in stream macroinvertebrates and habitat characteristics in heavily forested landscapes (Richards and Host 1994). Numerous researchers have conducted similar analysis. In all these studies, however, there were significant amounts of unexplained variation in the response variables, which can be attributed to both the natural variation in the variable, as well as not incorporating other causal

variables. A chronic and well-documented shortcoming in all these analyses is the lack of fine-scale information on riparian condition (Richards et al. 1996).

The NLCD data used here is no exception. The National Land Cover Characterization Project has provided 30 m coverage of 21 land use classes over the conterminous United States. This ambitious project has created a publicly available dataset with wide variety of applications, including environmental inventories, watershed assessments, and fire risk assessment (Vogelmann et al. 2001). It is important to consider, in any modeling, analysis, or planning effort, the appropriate scale of the source data relative to the questions being asked. In this study, we have shown that there are limitations to the NLCD (and to other Landsat or coarser) scale data that relate to data resolution, spatial extent, and classification resolution for fine-scale landscape features, in particular, riparian communities along streams and wetlands. The scale of ecological processes at the interface of terrestrial and aquatic systems is well below the spatial resolution of Thematic Mapper data.

In spite of these limitations, remotely sensed data remains a valuable tool for ecological assessment. As higher resolution data, in the form of SPOT, IKONOS, Radarsat and Quickbird imagery receive more widespread use, we will have an increased ability to incorporate causal information at finer spatial scales. These new data sources will improve both our analytical and modeling efforts, as well as the management interpretations that stem from them.

ACKNOWLEDGEMENTS

Funding for this project was provided by a grant from the U.S. Environmental Protection Agency to C. Richards, L. Johnson and G. Host (STAR Grant #R825870-01-0). We also wish to acknowledge the contributions of Brian Allen and Connie Host, who assembled many of the GIS data sets used. This is contribution number 350 of the Center for Water and the Environment at the Natural Resources Research Institute. Although the research described in this article has been funded wholly or in part by the United States Environmental Protection, it has not been subjected to the Agency's required peer and policy review and therefore does not necessarily reflect the views of the Agency and no official endorsement should be inferred.

REFERENCES

Anderson, K. A., T. J. Kelly, R. M. Sushak, C. A. Hagley, D. A. Jensen, and G. M. Kreag. 1999. Summary Report on Public Perceptions of the Impacts, Use, and Future of Minnesota Lakes: Results of the 1998 Minnesota Lakes Survey. University of Minnesota Sea Grant and Minnesota Department of Natural Resources, Office of Management and Budget Services, Washingto, D.C.

Benson, B. J., and M. D. Mackenzie. 1995. Effects of sensor spatial resolution on landscape structure parameters. Landscape Ecology 10:113-120.

Congalton, R. G., K. Birch, R. Jones, and J. Schreiver. 2002. Evaluating remotely sensed techniques for mapping riparian vegetation. Computers and Electronics in Agriculture 37:113-126.

Cowardin, L. M., V. Carter, F. Golet, and E. LaRoe. 1979. Classification of Wetlands and Deepwater Habitats of the United States. U.S. Fish and Wildlife Service.

Décamps, H. 1996. The renewal of floodplain forests along rivers: a landscape perspective. Verhandlungen der Internationale Vereinigung Fur Limnology 26:35-59.

Erickson, D., and R. De Young. 1993. Management of farm woodlots and windbreaks: some psychological and landscape patterns. Journal of Environmental Systems 22:233-247.

Foster, D., F. J. Swanson, J. D. Aber, I. Burke, N. Brokaw, D. Tilman, and A. Knapp. 2003. The importance of land-use legacies to ecology and conservation. Bioscience 53:77-88.

Gregory, S. V., F. J. Swanson, and W. A. McKee. 1991. An ecosystem perspective of riparian zones. Bioscience 40:540-551.

Hernandez, M., S. N. Miller, D. C. Goodrich, B. F. Goff, W. G. Kepner, C. M. Edmunds, and K. B. Jones. 2000. Modeling runoff response to land cover and rainfall: spatial variability in semi-arid watersheds. Journal of Environment and Monitoring and Assessment 64:285-298.

Johnson, L. B., and S. H. Gage. 1997. Landscape approaches to the analysis of aquatic ecosystems. Freshwater Biology 37:113-132.

Johnson, L. B., C. Richards, G. E. Host, and J. W. Arthur. 1997. Landscape influences on water chemistry in Midwestern stream ecosystems. Freshwater Biology 37:193-208.

Jones, K. B., A. C. Neale, M. S. Nash, R. D. V. Remortal, J. D. Wickham, K. H. Riitters, and R. V. O'Neill. 2001a. Predicting nutrient and sediment loadings to streams from landscape metrics: a multiple watershed study from the United States Mid-Atlantic Region. Landscape Ecology 16: 301-312.

Krummel, J. R., R. H. Gardner, G. Sugihara, R. V. O'Neill, and P. R. Coleman. 1987. Landscape patterns in a disturbed environment. Oikos 48:321-324.

Malanson, G. P. 1993. Riparian Landscapes. Cambridge University Press, Cambridge.

Marceau, D. J., D. J. Gratton, R. A. Fournier, and J. P. Fortin. 1994. Remote sensing and the measurement of geographical entities in a forested environment. 2. The optimal spatial resolution. Remote Sensing of Environment 49:105-117.

Moody, A., and C. E. Woodcock. 1995. The influence of scale and the spatial characteristics of landscapes on land-cover mapping using remote sensing. Landscape Ecology 10:363-379.

Muller, E. 1997. Mapping riparian vegetation along rivers: old concepts and new methods. Aquatic Botany 58:411-437.

Naiman, R. J., R. E. Bilby, and P. A. Bisson. 2000. Riparian ecology and management in the Pacific coastal rain forest. Bioscience 50:996-1112.

Richards, C. and G. Host. 1994. Examining land use influences on stream habitats and macroinvertebrates: a GIS approach. Water Resources Bulletin 30:729-738.

Richards, C., L. B. Johnson, and G. E. Host. 1996. Landscape scale influences on stream habitats and biota. Canadian Journal of Fisheries and Aquatic Sciences 53 (Suppl. 1):295-311.

Saura, S. 2002. Effects of minimum mapping unit on land cover data spatial configuration and composition. International Journal of Remote Sensing 23:4853-4880.

Stehman, S. V., J. D. Wickham, J. H. Smith, and L. Yang. 2003. Thematic accuracy of the 1992 National Land-Cover Data for the eastern United States: statistical methodology and regional results. Remote Sensing of Environment 86:500-516.

Turner, M. G., R. V. O'Neill, R. H. Gardner, and B. T. Milne. 1989. Effects of changing spatial scale on the analysis of landscape pattern. Landscape Ecology 3:153-162.

Vogelmann, J. E., S. M. Howard, L. Yang, C. R. Larson, B. K. Wylie, and N. V. Direl. 2001. Completion of the 1990s national land cover data set for the conterminous United States from Landsat Thematic Mapper data and ancillary data sources. Photogrammetric Engineering and Remote Sensing 67: 650-662.

Woodcock, C. E., and A. H. Strahler. 1987. The factor of scale in remote sensing. Remote sensing of Environment 21:311-332.

Wu, J. 2004. Effects of changing scale on landscape pattern analysis: scaling relations. Landscape Ecology 19:125-138.

CHAPTER 16

SCALE ISSUES IN LAKE-WATERSHED INTERACTIONS:

Assessing Shoreline Development Impacts on Water Clarity

CAROL A. JOHNSTON AND BORIS A. SHMAGIN

16.1 INTRODUCTION

Lakes are valued for their recreational benefits and aesthetics, clearer lakes being more desirable than murky lakes. Lake water clarity is affected by the production of microscopic algae suspended in the water column, which is in turn affected by nutrient inputs to lakes. Certain human land uses within lake watersheds can increase waterborne nutrient fluxes to lakes, so regulations that restrict land use to protect lake integrity are widely accepted. Maine, Minnesota, and Wisconsin have implemented land use zoning laws within the shorelands surrounding lakes statewide. Lakeshore zoning programs also exist for large regions within Michigan (Great Lakes Shorelands Management Program) and New York (Adirondack Park, New York City Watershed).

Shoreland zoning limits human activities within the lake/land interface, with the underlying assumption that land perturbations proximal to the lake have the greatest effect on water quality. However, more distal portions of watersheds may have a greater effect if streams provide efficient pathways of connection between upland and lake (see Wickham et al., Chapter 12, Jones et al., Chapter 11). Subsurface ground water flow into lakes often influences lake water and chemical budgets to a greater extent than surface water inflow, and the spatial extent of regional ground-water systems may be considerably different than the spatial extent of land surface draining into lakes (Winter et al. 1998). Furthermore, within-lake processes may override the influence of incoming materials from the watershed. Understanding of the relative importance of proximal vs. distal, surface water vs. ground water, and watershed vs. in-lake processes is essential to enlightened lake management.

In Minnesota, shoreland regulation was enacted in 1989 (Minnesota Regulations Parts 6120.2500-6120.3900). Statewide standards, implemented by county and city governments, set guidelines for the use and development of shoreland property. Minnesota defines lake shoreland as lands within 1000' from the ordinary high water

J. Wu, K.B. Jones, H. Li, and O.L. Loucks (eds.),
Scaling and Uncertainty Analysis in Ecology: Methods and Applications, 297–313.

level of lakes larger than 25 acres (10 acres in municipalities). The regulations are implemented by a hierarchy of political entities operating over different spatial scales: (1) the *state* sets minimum standards, (2) *counties* and *municipalities* implement zoning ordinances, (3) *lake area* differentiates larger lakes subject to regulation from small lakes that are not, and (4) *shoreland* is the area within which land use activities are regulated. Shoreland regulation is not mandated at the federal level, but Section 305(b) of the Federal Water Pollution Control Act provides impetus to states to improve lake water quality (EPA 2002).

The *ecoregion* scale has also been embraced by state agencies responsible for protecting Minnesota's lakes, because the biology and chemistry of Minnesota's lakes was regionally variable even before the arrival of European settlers. Studies of reference lakes, deemed to be representative and minimally impacted by humans, have shown large differences among Minnesota's four major ecoregions in transparency, phosphorus, and chlorophyll concentrations, measures of lake water quality (Heiskary and Wilson 1989). These studies resulted in different phosphorus target concentrations for lakes in these ecoregions (Heiskary and Walker 1988).

The Minnesota Department of Natural Resources (MDNR) funded this and a companion study to investigate the relationship between development and the condition of lakes within the Northern Lakes and Forests and North Central Hardwood Forest ecoregions. The research had the specific goal of creating a methodology to assess the cumulative effects of development on lakes by relating indicators of human activity to indicators of environmental condition. The methods used in the research were statistical analysis of landscape empirical relationships using a very large sample size (Johnston and Shmagin 1999).

The analysis of scaling effects was not an *a priori* goal of this research. By specifying two separate ecoregions, the MDNR demonstrated an awareness of the concept that spatial extent could be delimited so as to minimize spatial heterogeneity due to underlying natural variability. Apart from this initial definition of extent, however, addressing scaling issues was not a specific goal of the management agency. Despite the relative lack of scaling design, a number of scaling issues emerged from the research as unanticipated findings, as discussed at the conclusion of the paper.

16.2 METHODS

The overall approach used in this study was to statistically analyze empirical data for multiple lakes and their surrounding landscapes. Data consisted of three types: (1) GIS databases defining the spatial extents of lakes, shorelands, and watersheds, (2) databases of lake transparency data from field measurements, and (3) databases of lake, shoreland and watershed characteristics, primarily GIS-derived. Only existing databases and GIS-derived data were used; no new field data were collected for the study due to funding limitations.

16.2.1 Study Region

The Northern Lakes and Forests (NLF) ecoregion (Omernik 1987) covers 68,243 km^2 in northeastern Minnesota, nearly 1/3 of the state. Split by three major drainage divides, the waters of the NLF region ultimately flow to the St. Lawrence Seaway, Hudson Bay, and the Gulf of Mexico. As its name implies, the ecoregion is primarily covered by forests (75% of ecoregion area) and lakes and marsh (11% of ecoregion area), with relatively little agriculture (5% cultivated land, 7% pasture & open land: Heiskary and Wilson 1989). Population of the shoreland zone is quite variable, however, with human populations ranging from 0 to 5,754 people in the shoreland zone (Johnston and Shmagin 1999). The abundance of lakes, lack of agriculture, and wide range of shoreland population make the NLF ecoregion an ideal region to study the effects of lakeshore development on lake quality.

Lakes in the NLF ecoregion are some of the cleanest in the state (MPCA 1999). Fully 26% of the lakes in the NLF ecoregion are oligotrophic (i.e., nutrient depauperate), in comparison with only 4% of the lakes in the Central Hardwood Forests ecoregion. In recognition of the exceptional water quality in the ecoregion's lakes, the Minnesota Pollution Control Agency (MPCA) has set lower thresholds for phosphorus concentrations that threaten swimmable use: 30 $\mu g\ L^{-1}$ in the NLF ecoregion vs. 70 $\mu g\ L^{-1}$ in the Western Corn Belt Plains and Northern Glaciated Plains ecoregions. These thresholds take into account regional differences in citizen user perceptions and values from minimally impacted lakes.

16.2.2 Defining Extents of Spatial Entities

16.2.2.1 Lakes

The unique identification of lakes within the study region was crucial to ensure that indicators of lake condition (derived primarily from EPA's STORET database) were linked to the appropriate GIS-derived landscape data. A coding system for uniquely numbering the thousands of lakes in Minnesota, commonly referred to as the Division of Waters (DOW) lake identification system, was used by Minnesota state agencies that submit lake condition data to STORET (Minnesota Conservation Department 1968). Lake identifiers used by other federal agencies (e.g., for lakes in Chippewa and Superior National Forests) had to be converted to the DOW system. We obtained, edited, and merged county-wide USGS Digital Line Graph hydrography databases to which DOW lake numbers had been added, and clipped the result with the mapped boundary of the NLF ecoregion. The resulting database contained 5,408 lakes with area >10 ha, hereafter referred to as the NLF Lakes Database (Johnston and Shmagin 1999).

16.2.2.2 Shorelands

The shoreland zone was determined by using a GIS to establish a 1000' buffer around the edge of each study lake, consistent with the regulatory definition of shoreland zone. Because lakes were often closer than 1000' from each other, the

shoreland zone of one lake often intersected the shoreland zones of other lakes. Therefore, it was necessary to iteratively generate a shoreland zone buffer for each study lake, and use it to clip out shoreland land use/land cover data for each individual lake. This process was accomplished by an ARC/INFO AML program that repeatedly selected a lake, generated its buffer, measured buffer area, and clipped out and saved shoreland data.

16.2.2.3 Watersheds

Although MDNR is in the process of mapping the watersheds for individual lakes in the state, that work was not completed when this study was done. Alternatively, we used an existing digital database of stream watersheds (average area = 40 km^2/watershed) that had been manually delineated and digitized by the Minnesota Department of Natural Resources, referred to as "minor watersheds" (http://lucy.lmic.state.mn.us/metadata/wshed95.html). This minor watershed coverage was intersected with the NLF Lakes database to determine lake occurrence within the minor watersheds. Of the 1189 minor watersheds in the NLF ecoregion that contained lakes, there was a median of three lakes per watershed. All lakes occurring within a minor watershed were assigned the same watershed attributes.

16.2.2.4 Ecoregion

Original boundaries of the NLF ecoregion, developed for a national aquatic ecoregion map by the U.S. Environmental Protection Agency (Omernik 1987), were recompiled by the Minnesota Pollution Control Agency to match MDNR minor watershed boundaries. The "Minnesota Ecological Classification System" (http://www.dnr.state.mn.us/ecs/index.html) is a finer scale eco-regionalization system developed by the MDNR Division of Forestry which aided in the interpretation of results.

16.2.3 Dependent Variable: Secchi Transparency

Secchi transparency was used as the measure of lake condition. Secchi depth is a simple measurement whereby a 20 cm black and white disk is lowered into a lake until it is no longer visible by an observer at the surface. This depth of disappearance is a measure of the transparency of the water, and is usually related to the trophic state of the lake (Carlson 1977). Lakes that are oligotrophic are clear, nutrient poor, and low in chlorophyll A due to lack of algal production. "Eutrophication" occurs when the addition of nutrients (particularly phosphorus) increases the primary productivity of algae suspended in lake water, thereby decreasing water clarity (Smith 1979). Lakeshore development can accelerate eutrophication by contributing nutrients to lake water from leaking septic systems, lawn fertilization, runoff from impervious surfaces, and eroded phosphorus-rich soil particles (Anderson et al. 1999, Garn 2002).

Table 16.1. *Shoreland-scale land use and population characteristics used in the analysis.*

Shoreland Zone Variable	*GIS Database*		*Variable Code*		*Variable Description*
	1977-79	1994-96	1977-79	1994-96	
Water, % of area	USGS Land Use/Land Cover	Minn. Census of Land	BUFWAT	BUFWAT	Areas of water (usually adjacent lakes) within the study lake's shoreland zone
Agriculture, % of area	USGS Land Use/Land Cover	Minn. Census of Land	BUFAGR	BUFCULT BUFGRASS	Cultivated cropland Hay/pasture/grassland
Forest, % of area	USGS Land Use/Land Cover	Minn. Census of Land	BUFFOR	BUFFOR BUFBRUSH	Forest (trees > 2/3 of total canopy) Shrubs, grass, scattered trees
Wetland, % of area	USGS Land Use/Land Cover	Minn. Census of Land	BUFWET	BUFWET	Bog/marsh/fen (1977 also includes forested wetland)
Mining, % of area	USGS Land Use/Land Cover	Minn. Census of Land	BUFMIN	BUFMIN	Mine pits and tailings piles; includes rock outcrops and sand dunes
Urban and rural development, % of area	USGS Land Use/Land Cover	Minn. Census of Land	BUFURB	BUFURB	Urban and rural development
Shoreland homes	1982 Minn. Shoreland Development Survey	–	SEASONHM PERMANHM SEASONLK PERMANLK SEASONBUF PERMANBUF	–	Number of seasonal homes Number of permanent homes Seasonal homes/lake area (#/km^2) Permanent homes/lake area (#/km^2) Seaonal homes/buffer area (#/km^2) Perman. homes/buffer area (#/km^2)
Population	–	U.S. Census	–	PERSONS PERSONLK PERSONBUF	Number of people in buffer area People/lake area (#/km^2) People/buffer area (#/km^2)

Secchi transparency is usually interpreted as an indicator of eutrophication, but it can also be affected by factors unrelated to eutrophication, such as dissolved organic compounds or suspended sediments in the water column. Data on lake Secchi depths were retrieved from the U.S. Environmental Protection Agency's STORET database and from the Minnesota Department of Natural Resources Fisheries Data Warehouse (FDW), a repository of lake survey data collected by the Fish and Wildlife Division (Scidmore 1970). Secchi transparency data were matched in time to correspond with the time periods represented by available land use/land cover data: 1977-1979 and 1994-1996 (see *Shoreland Land Use and Population Characteristics* section below).

16.2.4 Independent Variables: Lake, Shoreland and Watershed Characteristics

Given the objective of the study, to investigate the relationship between development and the condition of lakes, we utilized a number of independent variables quantifying shoreland land use and population characteristics (Table 16.1). Natural as well as anthropogenic factors are known to affect lake condition, so we additionally characterized lake morphometry and watershed physical and soil characteristics. Because land use and population are subject to change over time, shoreland zone characteristics were period-specific, but lake and watershed characteristics were not.

16.2.4.1 Lake morphometry

Lake area and perimeter were measured by GIS using our NLF Lakes Database, and lake buffer area was measured by GIS as the area of the shoreland zone, the derivation of which is described below. These three variables were log-transformed to normalize their data distribution prior to statistical analysis. Maximum lake depth data were obtained from three state agencies: (1) the FDW, (2) lake assessments conducted by the Minnesota Pollution Control Agency pursuant to Section 305(b) of the Federal Water Pollution Control Act, and (3) the SWIM Lake Summary Database maintained by the Land Management Information Center of Minnesota Planning. In all cases, maximum lake depth was greater than Secchi transparency depth (i.e., the Secchi disk didn't hit bottom before it disappeared from view).

16.2.4.2 Watershed characteristics

Watershed physical and soil characteristics were obtained primarily from GIS databases included in the Minnesota Land Management Information Center's MGC 100, a statewide 100 meter cell size raster database series (http://lucy.lmic.state.mn.us/). Minor watershed boundaries (see "Defining Extents of Spatial Entities") were intersected with MGC 100 files for topography, soil characteristics, public land ownership, and proximity to roads (Table 16.2). Topographic and soil properties were selected due to their potential relationship to nonpoint-source pollution: topography drives water flow, surface and vadose soil textures can influence water runoff, the P content and erodibility of watershed soils

can influence P and suspended solids contributions to lakes, and watershed soil pH can influence lake pH.

Table 16.2. *Watershed-scale physical and chemical characteristics derived from Minnesota Land Management Information Center MGC 100 databases.*

Database	*Description*	*Processing*
Elevation	elevation in meters above mean sea level, derived from USGS 1-Degree DEM 3- x 3-arc-second data spacing	computed average watershed elevation (m), elevation range within watershed (m), and average slope of non-lake watershed area (%)
Public Ownership	ownership as of 1983 for land under federal, state, or county jurisdiction	original 62 categories grouped into 3 categories: wilderness, non-wilderness public, and private; as % of non-lake watershed area
Proximity to Highways	40-acre parcels adjacent to or containing roads and intersections	original 26 categories grouped into 3 categories: paved roads, unpaved roads, and no highway access; as % of non-lake watershed area
MPCA Soil Materials	surface soil texture	original 9 categories grouped into 5 categories; as % of non-lake watershed area
MPCA Vadose Zone Materials	soil texture in the vadose zone	original 9 categories grouped into 5 categories; as % of non-lake watershed area
Available P in the Rooting Zone	interpreted from Minnesota Soil Atlas	4 categories: low, medium, high, variable from low to high; as % of non-lake watershed area
Soil pH	interpreted from Minnesota Soil Atlas	11 original pH categories grouped into 6 categories: <5.5, 5.6-6.0, 6.1-7.3, < 6.0, > 6.0, no_rating
Surface K-factor	soil erodibility factor (K) in the Universal Soil Loss Equation, interpreted from Minnesota Soil Atlas based on surface soil texture	33 original K-factor categories grouped into 7 categories: 1-7, 8-19, 20-24, 25-27, 28, 29-36, no_rating
Lake Area, % of Watershed (LAKE:SHED)	water area from MPCA Soil Materials file	as % of total watershed area

16.2.4.3 Shoreland land use and population characteristics

The boundary of each shoreland zone was used to clip data from existing land use/land cover databases representing two different time periods: 1977-1979 and 1994-1996 (Table 16.1). USGS Land Use/Land Cover had been mapped by interpretation of aerial photos taken in 1977-79, whereas the Minnesota 1990s Census of the Land had been derived from Landsat Thematic Mapper imagery. The two data sets have classification systems that are similar but not identical, the main

difference being that the Census of the Land distinguishes two agricultural classes ("cultivated" and "grassland"), and separates brushland from forest.

A Shoreland Development Survey conducted in 1982 by MDNR was used as a source of housing data to correspond with the 1977-79 land use data. The original data, expressed as counts per lake of seasonal and permanent homes, were divided by GIS-derived lake and lake buffer areas to compute densities. The 1990 U.S. Census was used as a source of population data to correspond with the 1994-96 land use data. The number of people within the shoreland zone was computed by multiplying population per census block by the proportion of block area within the GIS-derived shoreland zone, and summing the results for the entire shoreland zone of each lake. A t-test for independent samples was used to compare Secchi transparency and maximum depth of lakes with no human population vs. lakes with one or more persons in the shoreland zone.

16.2.5 Data Sets and Statistical Analysis

In order to establish relationships between dependent and independent variables, we compiled data matrices containing both types of data. To be considered for analysis, a lake had to have Secchi transparency measured within 1977-79 or 1994-96, and a maximum depth measurement. Relationships between dependent and independent variables were thus established on three data sets: the 163 lakes sampled in 1977-79 for which maximum depth and 1982 DNR Shoreland Development Survey data were available ("Data Set I"), the 321 lakes sampled in 1977-79 for which maximum depth data were available, which included the 163 lakes in the first data set ("Data Set II"), and the 575 lakes sampled in 1994-96 for which maximum depth data were available ("Data Set III").

Statistical analysis to assess the effects of environmental characteristics on Secchi transparency consisted of an exploratory phase and a multiple regression phase. In the exploratory phase, descriptive statistics were generated for each variable, the normality of the data was examined, and $\log_{10}$ transformations were performed on variables where necessary to stabilize the variance (Snedecor and Cochran 1980). In the multiple regression phase, we performed stepwise multiple regressions between Secchi transparency (dependent variable), and environmental characteristics. Independent variables were selected for inclusion in regression equations based on the magnitude of partial correlations with Secchi transparency. Selected variables were then included or rejected, using an F-test threshold of 4.0. All variables retained in the stepwise multiple regressions were significant at $p < 0.05$.

We hypothesized that lake clarity would decrease with increasing shoreland development due to increased lake eutrophication associated with that development. Specifically, we hypothesized that Secchi transparency would be inversely related to the following measures of shoreland development (described in Table 16.1): BUFURB, SEASONHM, PERMANHM, SEASONLK, PERMANLK, SEASONBUF, PERMANBUF, PERSONS, PERSONLK, and PERSONBUF. We also hypothesized that Secchi transparency would be inversely related to the proportion of mining land

use in the shoreland (BUFMIN) due to increased turbidity caused by suspended sediments associated with mining activities.

16.3 RESULTS

16.3.1 Univariate Relationships between Development and Secchi Transparency

Secchi transparency ranged widely within the NLF ecoregion, from 0.6 to nearly 12 m. Lakes also varied widely in degree of development, with human populations ranging from 0 to 5,754 people in the shoreland zone. However, there were no significant correlations between development-related variables and lake transparency, with the exception of mining in the shoreland zone (BUFMIN), which was associated with clearer lakes in data set III ($r = 0.181$) but not data set II ($r = 0.030$). This relationship was affected by the fact that three of the clearest lakes in data set III were mine pit lakes, with transparencies exceeding 7.8 m: Tioga Mine Pit, Sabin Lake, and the Judson Mine Pit. These lakes were not part of the 1977-79 data set.

An unexpected result was that lakes with people living around them were significantly *clearer* than those without ($t = 1.98$, $P = 0.048$). Mean Secchi depth for lakes with people ($n = 417$) was 3.26 m, and mean Secchi depth for those without ($n = 172$) was 2.97 m. Maximum lake depth was also significantly greater for populated than unpopulated lakes, averaging 11.1 m for unpopulated lakes and 14.5 m for populated lakes ($t = 3.82$, $P = 0.000$). This difference in maximum lake depth, rather than the presence or absence of people *per se*, may explain why the populated lakes were clearer than unpopulated lakes. There is also the possibility of self-selection, in that people prefer to live on clearer lakes.

16.3.2 Relationships between Environmental Setting and Secchi Depth

In the stepwise regressions, maximum lake depth was the first explanatory variable selected for all three data sets, with deeper Secchi transparencies associated with deeper lakes (i.e., deeper lakes were clearer). Maximum lake depth provided 47 to 75% of the explanatory power of the stepwise multiple regressions.

The LAKE:SHED ratio was negatively related to transparency in the stepwise regressions for all three data sets, meaning that lakes in watersheds with a large proportion of lake area tended to be less transparent than lakes embedded in watersheds with a small proportion of lake area. Although lakes with a high value for LAKE:SHED tended to be large (the nine lakes with the highest LAKE:SHED ratios were among the 20 largest lakes in the state: Table 16.3), not all large lakes had high LAKE:SHED values. The LAKE:SHED variable seemed to be an indicator of the size of a lake's littoral zone (i.e., lake areas with a depth < 4.6 m), rather than lake area *per se*. Littoral zone area was not used as an independent variable in this study because these data are not widely available for Minnesota lakes, but most of the nine lakes with the highest LAKE:SHED ratios had large littoral zones, ranging from 20% to 47% of total lake area. These large littoral zones were a subaqueous

continuation of the gradually sloping, glaciated land surface surrounding the lakes. For example, Big Sandy Lake in Aitkin County is embedded in the relatively subdued topography of the Culver End Moraine (Hobbs and Goebel 1982), and the maximum elevation in its watershed is only 44 m higher than the lake surface (Figure 16.1). The lake area variable (LGLKAREA) was not selected as an explanatory variable in any of the stepwise multiple regressions.

Table 16.3. *Lakes with the largest LAKE:SHED values, sorted by transparency. Littoral zone = lake area with depth < 15 feet, computed as a proportion of total lake area.*

Lake Name	*Secchi depth, m*	*Lake area, ha*	*LAKE:SHED, %*	*Maximum lake depth, m*	*Littoral zone, %*
Cass	4.42	6429	45%	36.6	20%
Whitefish	3.96	3149	49%	42.1	37%
Gull	2.74	3961	46%	24.4	30%
Vermilion	2.66	15825	44%	23.2	37%
Mille Lacs	2.35	51748	76%	12.8	25%
Winnibigoshish	2.07	22976	52%	21.3	32%
Bemidji	1.89	2617	42%	23.2	29%
Island Lake Res.	1.68	3333	41%	28.7	44%
Big Sandy	1.55	2943	60%	25.6	47%

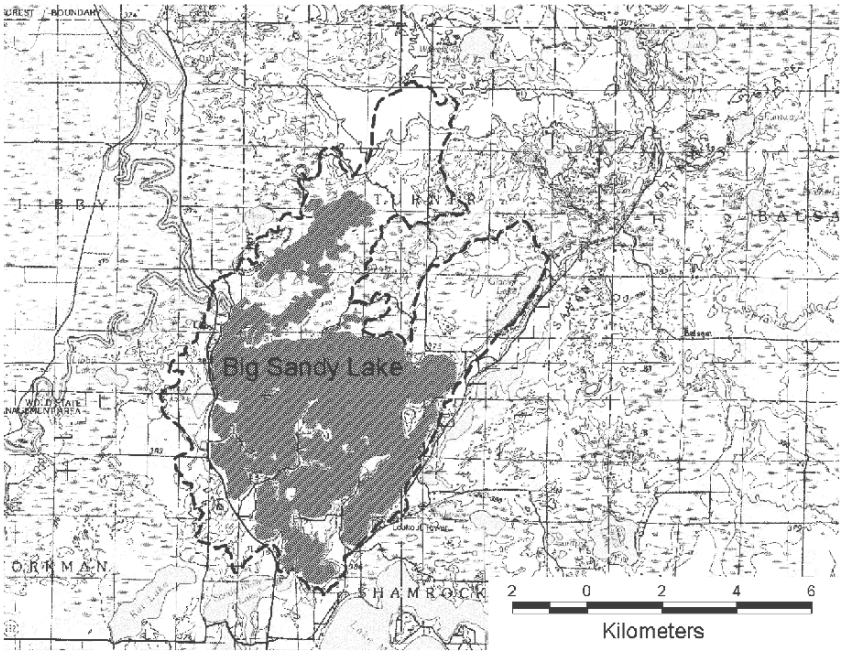


Figure 16.1. *Map of Big Sandy Lake (2,943 ha) and its watershed (dashed line), illustrating high LAKE:SHED ratio and gentle terrain. Topographic contour interval = 10 m. Elevation of Big Sandy Lake = 370 m.*

Average watershed slope was positively related to transparency in the stepwise regressions, steeper terrain being associated with deeper Secchi transparency. Examination of the slope map for the NLF region revealed a concentration of steeply sloping lands along the northernmost edge of the ecoregion (Figure 16.2). This region constitutes the Border Lakes Subsection of the Laurentian Mixed Forest Province, as defined by the Minnesota Ecological Classification System (http://www.dnr.state.mn.us/ecs/index.html):

> This subsection consists of scoured bedrock uplands or shallow soils on bedrock, with large numbers of lakes. Glacial ice moved from west to east across the subsection, deepening stream valleys in the bedrock. Long, east-west oriented lakes now occupy these enlarged valleys. Topography is dominantly rolling with irregular slopes and many craggy outcrops of bedrock.

Given our emphasis on potential human impacts on lake condition (Table 16.1), it was surprising that few variables related to human activity were selected in the stepwise regressions. No human activity variables were selected for data sets I and II. In data set III, the number of people per unit lake area (PERSONLK) was inversely related to lake clarity, and BUFMIN was positively related to Secchi transparency, because very clear mine pit lakes were associated with this land use.

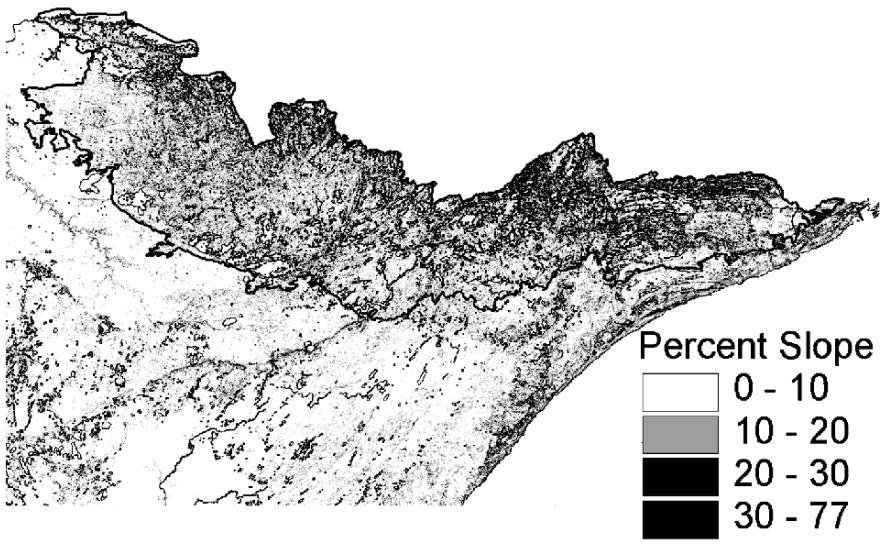


Figure 16.2. *Slope map for the northern edge of the Northern Lakes and Forests Ecoregion, with outline of the Border Lakes Subsection of the Laurentian Mixed Forest Province, as defined by the Minnesota Ecological Classification System.*

16.4 DISCUSSION

16.4.1 Influence of Lake Morphometry

Maximum lake depth was consistently the single best predictor of Secchi transparency: deeper lakes were clearer. Lake depth provided 47 to 75% of the explanatory power of the stepwise multiple regressions, much more than any variable related to development. Lake depth is not a causative variable in and of itself, but is related to internal cycling of phosphorus (within-lake scale) and ground-water fluxes (regional scale). Deeper lakes have a greater tendency to thermally stratify, such that the layer at the surface which receives the most light (the epilimnion) is effectively separated from the phosphorus-rich waters at depth except during times of lake water mixing, usually during the spring and fall (Nurnberg 1998). The phytoplankton that can decrease lake transparency require *both* light and phosphorus, so this spatial segregation of essential resources limits primary production during periods of stratification, thus keeping the lakes more transparent (Fee 1979). Very shallow lakes are less likely to stratify, and those that do may have more frequent wind-driven mixing events.

Lake mixing may also explain the perplexing result that Secchi transparency *decreased* with an increase in the LAKE:SHED ratio. Based on Schindler's (1971) work on Canadian lakes, we expected that the greater the contributing area of watershed to a given volume of lake, the greater the external loading of nutrients to the lake. Our study showed that lakes in watersheds dominated by very large lakes (high LAKE:SHED ratio) tended to be less transparent than lakes embedded in watersheds with a smaller proportion of lake area (low LAKE:SHED ratio), the opposite of Schindler's prediction. We believe that this is due to the complicating effect of internal phosphorus release from mixing events, because the lakes with the highest LAKE:SHED ratios were large and had large littoral zones. Internal phosphorus release depends on contact of the profundal sediment with lake water, which is enhanced by deep summertime mixing events in shallow lakes (Larsen et al. 1981, Jacoby et al. 1982, Patalas 1984). In a study of 96 lakes in the Twin Cities, Minnesota metropolitan area, Osgood (1988) found that the ratio of mean lake depth to the square root of lake surface area was a good indicator of a lake's internal phosphorus loading due to frequent mixing. Similarly, Nurnberg's (1998) model of phosphorus retention was based on mean lake depth and annual water residence time. Thus, the influence of the LAKE:SHED variable in this study is probably due to a lake-scale phenomenon (mixing), rather than the watershed scale.

Deeper lakes are also more likely to intercept regional aquifers, thereby receiving a larger percentage of their water inputs from ground water (Winter et al. 1998). Ground water is typically low in nitrogen and phosphorus, so that lakes receiving a substantial portion of their water volume from deep regional aquifers are likely to be nutrient-limited than those that receive water primarily from surface water. For example, ground water provides 40% of the annual water volume to White Clay Lake in Wisconsin, but only 26% of the nitrogen and 21% of the phosphorus because of the dilute concentrations of those nutrients in ground water (Madison and Peterson 1976). This input of nutrient-poor ground water dilutes

nutrient inputs from other water sources, thereby decreasing potential phytoplankton production.

16.4.2 Influence of Shoreland Development

Although the shoreland is the scale of greatest interest to land managers, few shoreland-scale variables were found to have a significant influence on lake quality by this analysis. We expected lakes with greater development and human populations to be less transparent, due to nutrient inputs from failing septic systems and lawn fertilizers (Anderson et al. 1999, Garn et al. 2002). However, we found that only small, shallow lakes in the most urban settings had demonstrably diminished transparency relative to other lakes in the ecoregion. We believe that lake water clarity in the NLF ecoregion is controlled primarily by processes operating at different scales than the shoreland scale, specifically in-lake nutrient cycling and regional geomorphology and groundwater inputs (see below), and that any nutrient contributions from shoreland development are negligible as a driver of lake eutrophication except when that development exceeds extreme thresholds. This finding is consistent with a study reconstructing water quality from fossil diatoms in 20 NLF lakes, which concluded that the phosphorus concentration and rates of sediment accumulation in lakes in the NLF Ecoregion have increased only slightly from pre-European times and that little change is apparent over the last 20 years (Heiskary and Swain 2002).

Mining in the shoreland was associated with clearer lakes, a result that was initially counterintuitive. Mine pit lakes receive nearly all of their water volume from ground water, which explains their observed clarity. When surface mines are abandoned, pits dug deep into the earth fill with nutrient-poor ground water that sustains minimal primary production. Such ground-water dominated lakes are hydrologically isolated from the land surface surrounding them. Thus, the positive relationship between mining in the buffer zone (BUFMIN) and Secchi transparency is probably unrelated to any actual activity within the shoreland, but rather a landscape scale effect due to the source of the water delivered to the hole in the ground created by mining.

16.4.3 Watershed to Regional Scale Influences

Given that topography drives surface water runoff and soils on steeper slopes are usually more erosive than comparable soils on level slopes, we expected that lakes in watersheds with steeper terrain would be less transparent. In fact, we found the opposite relationship: lakes in steep terrain were clearer. In the NLF, average watershed steepness is greatest in the bedrock-controlled Border Lakes area (Figure 16.2). This region contains the Boundary Waters Canoe Area Wilderness, and comprises undulating Precambrian bedrock of the Laurentian Shield with small watersheds and abundant lakes. Thus, although measured at the watershed scale, average watershed slope is really a surrogate indicator for a larger underlying geomorphic region. We expect that this relationship between average watershed

slope and lake transparency would not hold in subsections of the state where topography is not bedrock-controlled.

16.4.4 Variables Measured at One Scale May Reflect Influences at Another Scale

As the results demonstrate, measurements of independent variables made at one scale may actually indicate processes occurring at a different scale. For example:

- LAKE:SHED ratio is a watershed-scale measurement, but appeared to be an indicator of within-lake mixing
- Lake depth is a lake-scale measurement, but can be an indicator of regional ground-water inputs
- BUFMIN is a shoreland-zone measurement, but appeared to be an indicator of regional ground-water inputs
- Average watershed slope is a watershed-scale measurement, but in northeastern Minnesota it appeared to be an indicator of regional geomorphology.

Discovery of these relationships required critical examination of the results with a knowledge of lake-scale limnological processes and regional-scale hydrology and geology. Such a critical examination might not have occurred if the results had been less counterintuitive. Because correlative results do not determine cause and effect, the results from landscape empirical studies such as this one should always be interpreted within a broader scientific context.

The primary dependent variable, Secchi transparency, also responds to phenomena at a variety of scales: within-lake phosphorus cycling, people in the shoreland zone, regional ground water and geomorphology. This complicates the interpretation of Secchi transparency as a response variable.

16.4.5 Lakes as Complex Systems

The rationale behind shoreland zoning derives from the first law of geography – "Everything is related to everything else, but near things are more related than distant things" (sensu Tobler 1970). However, the relationship between lakes and landscapes may also be governed by the second law of geography (sensu Arbia et al. 1996): "Everything is related to everything else, but things observed at a coarse spatial resolution are more related than things observed at a finer resolution." In this case study, for example, the condition of lakes within the Border Lakes Subsection (Figure 16.2) appeared to be governed more by their common regional geomorphology than by any finer-scale land uses superimposed upon that geomorphic foundation.

This case study illustrates the problem of scale multiplicity in ecological pattern and process (King et al. 1991, Wu 1999, 2004). Real environments have hierarchical structure, such that each patch contains smaller patches of component subhabitats, which in turn contain finer scale subhabitats, and so on (MacArthur 1972, Urban et al. 1987). A point observation such as Secchi transparency integrates the effects on lake eutrophication of ecological drivers operating at multiple scales, making it

difficult to parse out the most influential scale of influence. A lake is thus an example of "organized complexity" (Weaver 1948, Allen and Starr 1982), in which systems have more components than can be characterized deterministically by analytical mathematics but an insufficient number of random components to be characterized adequately by traditional statistical methods (Wu 1999).

16.4.6 Consider the Scale of Regulatory Reality

Public policy often dictates the scales at which management agencies operate (Loucks et al., Chapter 17). The limits of the shoreland zone defined by different states probably had more to do with the political feasibility of implementing land use controls than any ecological analysis. There is no ecological reason why the lakeshore regulated in Maine is 250' wide, whereas the lakeshore regulated in Minnesota and Wisconsin is 1000' wide. In fact, one study has shown that pollutant-contributing areas do not follow a pattern defined by an equidistant corridor around a stream, but are irregularly shaped (Levine et al. 1993). Other studies have shown that correlations between land use measured at the watershed scale is as good or better as a predictor of stream water quality than is land use measured at within an equidistant corridor around the stream (Omernik et al. 1981, Hunsaker et al. 1992). The challenge is to define a shoreland zone that is both politically and ecologically defensible.

State agencies charged with managing natural resources generally implement statutes uniformly statewide, because to do otherwise might imply political favoritism. However, when there is a compelling scientific reason to regionalize states into smaller areas for regulatory purposes, management agencies are becoming increasingly willing to do so. Minnesota was an early proponent of the regionalization of phosphorus criteria for lakes in different ecoregions of the state. Our results suggest that the next logical step after regionalization is to differentiate lakes into depth classes, using the scale of the lakes themselves to apply different standards.

ACKNOWLEDGEMENTS

This material is based upon work supported by the Minnesota Department of Natural Resources (Charles H. Welling, Contracting Officer), by the National Science Foundation (EPSCoR Grant EPS-0091948) and by the State of South Dakota. We thank Suzanne Maeder of the Minnesota Land Management Information Center, Louise Hotka and Steve Heiskary of the Minnesota Pollution Control Agency, and Karen Schroeder of the Minnesota Department of Natural Resources for assistance with databases used in this study. Dave Wright of the Minnesota Pollution Control Agency provided helpful comments on the draft manuscript. Paul Meysembourg, Brian Allen, Michael Bourdaghs, and Amy Trauger contributed to data assembly, GIS analysis, and figure production.

REFERENCES

Allen, T. F. H., and T. B. Starr. 1982. Hierarchy: Perspectives for Ecological Complexity. University of Chicago Press, Chicago.

Anderson, K. A., T. J. Kelly, R. M. Sushak, C. A. Hagley, D. A. Jensen, and G. M. Kreag. 1999. Summary Report on Public Perceptions of the Impacts, Use, and Future of Minnesota Lakes. University of Minnesota Sea Grant and Minnesota Department of Natural Resources, Office of Management and Budget Services.

Arbia, G., R. Benedetti, and G. Espa. 1996. Effects of the MAUP on image classification. Geographic Systems 3:123-141.

Carlson, R. E. 1977. A trophic state index for lakes. Limnology and Oceanography 22:361-369.

Environmental Protection Agency (EPA). 2002. National Water Quality Inventory, 2000 Report. Office of Water, U.S. Environmental Protection Agency, Washington, D.C.

Fee, E. J. 1979. A relation between lake morphometry and primary productivity and its use in interpreting whole-lake eutrophication experiments. Limnology and Oceanography 24:401-416.

Garn, H. S. 2002. Effects of Lawn Fertilizer on Nutrient Concentration in Runoff from Lakeshore Lawns, Lauderdale Lakes, Wisconsin. U.S. Geological Survey, Middleton.

Heiskary, S. A., and E. B. Swain. 2002. Water quality reconstruction from fossil diatoms: applications for trend assessment, model verification, and development of nutrient criteria for lakes in Minnesota, USA. Water Quality Division, Minnesota Pollution Control Agency, St. Paul, MN.

Heiskary, S. A., and W. W. Walker. 1988. Developing phosphorus criteria for Minnesota lakes. Lake and Reservoir Management 4:1-10.

Heiskary, S. A., and C. B. Wilson. 1989. The regional nature of lake water quality across Minnesota: an analysis for improving resource management. Journal of the Minnesota Academy of Science 55: 71-77.

Hobbs, H. C., and J. E. Goebel. 1982. Geologic Map of Minnesota: Quaternary Geology (1:500,000 map). State Map Series S-1. Minnesota Geological Survey, University of Minnesota, St. Paul, MN.

Hunsaker, C. T., D. A. Levine, S. P. Timmons, B. L. Jackson, and R. V. O'Neill. 1992. Landscape characterization for assessing regional water quality. Pages 997-1006 *in* D. McKenzie, E. Hyatt, and J. McDonald, editors. Proceedings of International Symposium on Ecological Indicators. Elsevier, Ft. Lauderdale, FL.

Jacoby, J. M., D. D. Lynch, E. B. Welch, and M. A. Perkins. 1982. Internal phosphorus loading in a shallow eutrophic lake. Water Resources 16:911-919.

Johnston, C. A., and B. A. Shmagin. 1999. Relationship between development and condition of lakes in Minnesota's northern lakes and forests ecoregion. Natural Resources Research Institute Technical Report NRRI/TR-99/34.

King, A. W., A. R. Johnson, and R. V. O'Neill. 1991. Transmutation and functional representation of heterogeneous landscapes. Landscape Ecology 5:239-253.

Larson, D. P., D. W. Schults, and K.W. Malueg. 1981. Summer internal phosphorus supplies in Shagawa Lake, Minnesota. Limnology and Oceanography 26:740-753.

Levine, D. A., C. T. Hunsaker, S. P. Timmins, and J. J. Beauchamp. 1993. A geographic information system approach to modeling nutrient and sediment transport. ORNL-6736. Environmental Sciences Division Publication No. 3993. Oak Ridge National Laboratory, Oak Ridge, TN.

MacArthur, R. H. 1972. Geographical Ecology: Patterns in the Distribution of Species. Princeton University Press, Princeton, NJ.

Madison, F. W., and J. O. Peterson. 1976. White Clay Lake Demonstration Project, Final Report to the U.S. Environmental Protection Agency. Department of Soil Science, University of Wisconsin, Madison, WI.

Minnesota Conservation Department. 1968. An inventory of Minnesota lakes. Bulletin No. 25, Division of Waters, Soils, and Minerals, Minnesota Conservation Department, St. Paul, MN.

Minnesota Pollution Control Agency (MPCA). 1999. Minnesota Lake Water Quality Assessment Data: 1998. An update to data presented in the Minnesota Lake Water Quality Assessment Report: 1990. Water Quality Division, Minnesota Pollution Control Agency, St. Paul, MN.

Nurnberg, G. K. 1998. Prediction of annual and seasonal phosphorus concentrations in stratified and polymictic lakes. Limnology and Oceanography 43:1544-1552.

Omernik, J. M. 1987. Ecoregions of the coterminous United States. Annals of the Association of American Geographers 771:118-125.

Omernik, J. M., A. R. Abernathy, and L. M. Male. 1981. Stream nutrient levels and proximity of agricultural and forest land to streams: some relationships. Journal of Soil and Water Conservation 36:227-231.

Osgood, R. A. 1988. Lake mixis and internal phosphorus dynamics. Archiv Fuer Hydrobiologie 113: 629-638.

Patalas, K. 1984. Mid-summer mixing depths of lakes of different latitudes. Internationale Vereinigung Fuer Theoretische und Angewandte Limnologie Verhandlungen 22:97-102.

Schindler, D. W. 1971. A hypothesis to explain differences and similarities among lakes in the Experimental Lakes Area, Northwestern Ontario. Journal of the Fisheries Research Board of Canada 28:295-301.

Scidmore, W. J. 1970. Manual of Instructions for Lake Survey. Minnesota Department of Conservation, Special Publication 1 (revised), St. Paul, MN.

Smith, V. H. 1979. Nutrition dependence of primary productivity in lakes. Limnology and Oceanography 24:1051-1064.

Snedecor, G. W., and W. G. Cochrane. 1980. Statistical Methods, 7th edition. Iowa State University Press, Ames, IA.

Tobler, W. 1970. A computer movie simulating urban growth in the Detroit region. Economical Geography (Suppl.) 46:234-240.

Urban, D. L., R. V. O'Neill, and H. H. Shugart. 1987. Landscape ecology: a hierarchical perspective can help scientists understand spatial patterns. Bioscience 37:119-127.

Weaver, W. 1948. Science and complexity. American Scientist 36:536-544.

Winter, T. C., J. W. Harvey, O. L. Franke, and W. M. Alley. 1998. Ground Water and Surface Water. A Single Resource. USGS Circular 1139. U.S. Geological Survey, Denver, CO.

Wu, J. 1999. Hierarchy and scaling: extrapolating information along a scaling ladder. Canadian Journal of Remote Sensing 25:367-380.

Wu, J. 2004. Effects of changing scale on landscape pattern analysis: scaling relations. Landscape Ecology 19:125-138.

CHAPTER 17

SCALING AND UNCERTAINTY IN REGION-WIDE WATER QUALITY DECISION-MAKING

ORIE L. LOUCKS, HARRY J. STONE, AND BRUCE M. KAHN

17.1 INTRODUCTION

Quite interesting questions about information flow across scale arise when we try to extend knowledge about ecological scaling to the corresponding scale questions arising when human societies make natural resource policy choices. In this paper we examine scaling in relation to planning and management (see Wu and Li, Chapter 1) for water quality at national, regional, state, and local levels (Stone 2004). However, the geopolitical scales at which institutional decisions are made do not often match spatially or temporally with the scale and dynamics of natural resource systems (Peterson 2000). Unlike ecological systems, state, federal and local regulatory decisions are bounded by historical decisions and legal jurisdiction or issues associated with the responsibilities of each level of government. In scaling for decision-making, we need to understand how such bounds change with time, as well as with spatial scale (Ostrom 1990).

In water quality regulation, limitations to the flow of information (scaling) associated with different levels in a hierarchy of decision-making authority, from federal to state to the local level, can also lead to uncertainty over who has responsibility at a specific level for specific management outcomes on the ground. The challenge for environmental policy analysts, therefore, is to understand how scaling influences, first, the debate about policy options, and secondly, the decisions after a policy has been determined, across the hierarchy of scale. Such understanding must include knowledge about the functioning of ecological systems, the sources of power or authority in policy systems, an ability to deconstruct the network of decision-making nodes, and the relationships associated with transfer of information or control across scale (Wu and Li, Chapter 1).

Here we follow the definition of scaling provided in Wu and Li (Chapter 1), and want to elaborate on what is implicit in use of the term "policy scale," which is listed along with intrinsic scale, process scale and observational scale. These last three scales can all be seen as linked to potential information transfer across levels in

J. Wu, K.B. Jones, H. Li, and O.L. Loucks (eds.),
Scaling and Uncertainty Analysis in Ecology: Methods and Applications, 315–325.

an ecological hierarchy. Policy relevant scaling for resource planning and management, however, must also be framed in relation to levels of government and the diverse geopolitical and non-government institutions that laws and regulations create. One can ask at what scale should we focus, but the answer noted by Wu and Li (Chapter 1) allows a focus on the role of information related to hierarchical system functioning, or the control mechanisms operating across a continuum of scale. Such information, in either ecological or policy systems, can transfer upward or downward through the levels of organization. The information flow can be followed through the hierarchy (or network) of public institutions just as well as through an ecological foodweb. Therefore, we do not initially focus on any one scale as being more important than others, but on scaling as the understanding and analysis of the flow of information, influence, and control across a range of scales, in ecological systems as well as among institutions and public policy systems.

Differences over regulatory jurisdiction lead to troubling questions about scale when the issues are presented in terms of a "conflicting power structure" (Stone 2004). Various authors have noted how conflict may arise from three dimensions of power implicit in the scale at which power operates in decision-making (Peterson 2000). Overt power occurs during brief periods of time and in a specific place. Covert power controls the type of decisions that can be made, and occurs at larger institutional scales. Structural power, the broadest dimension, works through the manipulation of culture (Luke 1973). These three power considerations frame the way information from various scales influences resource management decisions. According to Gaventa (1980), the power element may be broadly diffuse at some times, but at other times a majority of sustainable development decisions could rest in the hands of the elite.

Unique problems arise for policy-related scaling, however, in suburban or rural areas where water quality decisions implemented in one local political jurisdiction can have effects in other jurisdictions downstream. Upstream rural landowners may be seen as a kind of minority, sometimes neither powerful politically nor powerful economically. Their non-participation in framing choices can lead to a further loss of power (Stone 1980). Such powerlessness is structural, but local, although it may be reinforced by institutions that potentially control regional policy system outcomes. Local officials, when confronted by local rural landowner interests, often have to heed an elite from a larger regional scale, after considering aspects of their own economic self-interest. Based on economic, associational and social elements of scale (Stone 1980), an elite can exert great influence on information transfer across scale when making choices, often disregarding the less concentrated or less powerful interests (Crenson 1971, Stone 1982, 2004).

The questions we ask in this paper, therefore, relate to understanding how information or influence from the largest scale (federal legislation) either controls, becomes diffused, or is made uncertain, as it passes to the States for implementation and then to local government for decisions on the ground. Our hypothesis is that hierarchical flows of information affect policy and decision outcomes by modifying personal or institutional power structure at local to national scales, thereby influencing institutional responses at the next lower or next higher level. For example, an individual landowner, low in the hierarchy, may be influenced to make

a choice, or decision, by the institutionally broader considerations of the county or state scale. However, the landowner can join with many others to resist and modify regulatory decisions being passed down from the state. We will use two case studies to consider such scale-determined processes for water quality regulation and the uncertainty that is generated across the full range of scale.

17.2 A FRAMEWORK

Just as our understanding of the nature and strength of ecosystem functioning is framed by hierarchically-linked levels of organization within ecosystems, with interaction across scales (Allen and Starr 1982), so also must our understanding of scaling in institutional decision-making be framed. Individual human behaviors are guided by values, attitudes, education and experience (Heberlein 1974), and similarly, the scope and connectedness of institutional behavior can be influenced by the attitudes and information made available through local, regional and national group leadership and their governing bodies (Edleman 1964). Societal factors link individual human responses, as well as institutional responses, to water quality problems across a range in scales. The responses on the part of various interest groups include education, the use of media, technologies available, changing value systems, and local to national power structures (Heberlein 1974). Economic factors at local, regional or national scales also may determine how cross-scale information transfer can affect water quality outcomes.

Therefore, the critical variables of interest for this paper, across a range of scales, are, first, various water quality measures, and secondly, the array of local to national policies and decisions already made to protect water quality. The federal Clean Water Act of 1972 established broad national goals and made the early policy choices, including the "fishable, swimmable" targets. However, local point-source dischargers, and even non-point contributors to pollution, often seek exemptions from regulation at some level due to considerations of personal hardship. In addition, on-the-ground water quality outcomes have elements that are physical and relatively unmanageable (e.g., consequences from seasonal low flow), and ecological (e.g., shifts in species composition). Thus, the ongoing process for making policy choices about water quality has had to consider information imbedded at the various scales of pollutant inputs and the physical and biological responses of ecosystems. Although local publics (or their institutions) often seek to moderate local control, the downstream consequences of choices made upstream influence ecological outcomes at scales extending far beyond the local domain. Information on large-scale downstream effects from upstream decisions should, theoretically, be adjudicated at some even larger scale (probably national) that considers rules for a level playing field across all such watershed problems. The result has been downscaling from federal regulations under the Clean Water Act, to the state and local level, with enforcement largely by mid-level institutions, the state agencies, but with judicial appeal of decisions at all scales from local to national.

A key feature of scaling is illustrated in the principle of "command and control," or top-down decision making. Our most common experience with such hierarchies (and associated information scaling) is conveyed to us through institutions where we

work or live. One example is the structuring (across a range of scales) of federal agency budget and planning processes (Table 17.1; Sample 1990). Scale here ranges from thousands of people at the largest level, to just a few individuals at the finest scale. It is interesting to note the counter-scaling in budget and planning, where the top-down directives have to be responded to from the bottom, all the way back to the top. We need to be aware of the appropriate application of power and authority at all the levels involved. Interestingly, clear assignment of responsibility and accountability in financial matters generally avoids any substantial uncertainty for budget-related decision processes across these scales.

Two policy scholars, Sabatier (1999) and Ostrom (1990), consider how scale operates during the political choice process in a democracy at any scale. They represent relationships and information flow as structured from the bottom up, recognizing grass roots public needs that then generate a typical up-scaling phenomenon. They see individuals as operating to make choices at each level, from local through area-wide "collective choice" situations, to "state variables in the world." People affected by choices made in this way then have the potential to influence the downscaling processes discussed above in relation to individuals wanting to overturn regulations applied in local situations. Uncertainty in outcome exists here because of diffuse "community" inputs, often reflecting subtle shifts in societal values at various temporal and spatial scales, shifts that are rarely predictable.

Linkages between physical or ecological system scales, and institutional or management scales, have been described by Gunderson and Holling (2002). The framework they suggest for considering scale in decision-making by institutions requires that linkages be understood and specified at each scale, with due consideration of up-scaling versus downscaling processes. Two other components of scale are considered by Gunderson and Holling (2002, citing Sabatier and Ostrom papers) as being central to public choice processes: broad scale in temporal inputs and outcomes, and a very large range in the numbers of people involved (from 10^0 to 10^8). Both up-scaling and downscaling processes can be present in the expectations from choices affecting these scales, as is implicit in the concept of "panarchy" used by Gunderson and Holling (2002).

Finally, before we come to some case studies, we should consider both institutional hierarchies of scale and associated sources of power as forces for change in the dynamic "adaptive cycle" theory presented by Holling in several papers (see Gunderson and Holling 2002, and Stone 2004). During the relatively slow "conservation" phase of an adaptive cycle, institutional decisions (from both downscaling and up scaling), and the application of concentrated institutional interest or power, tend to sustain the institution's previous practice, even if it is out of date. However, when external circumstances change substantially, often at a large scale (following a federal court reversal, or election of new leadership in relation to an issue), a quick "release phase" is generated. In this phase, new information can be transferred from other scales and becomes a part of adopting new policies and related decisions. Most decisions are highly predictable during the conservation phase, but when external conditions change and create a "release phase," the use of new knowledge may make outcomes relatively more reasonable, but also

unpredictable for a time (Gunderson and Holling 2002). In the case studies that follow, we should expect to see uncertainty deriving from both ecological system scaling and from policy system scaling. We may also see patterns of change in the institutions managing resources, a pattern that could derive as well from a change in knowledge-use across scales.

Table 17.1. *Downscaling in the budget development and planning process of the U.S. Forest Service (after Sample 1990).*

Office/Level	*Advice/Directive/Response*		
OMB	Initial Budget Advice June	Initial Review by budget Examiner for Forest Service Budget	Tentative Budget for Forest Service "passback" for Secretary Review
Secretary's Office		Department Allowance (review by Office of Budget & Program Analysis October)	Secretary's Review of "passback"
Washington Office Forest Service	PD &: Budget Instructions (September)	Forest Service Agency Request (September)	
Regional Offices	Regional Budget Instructions	Regional Budget Proposal (June)	
Supervisors Offices	Request for Budget Information	National Forest Level Budget Proposal (June)	
Ranger Districts			

17.3 CASE STUDY I: ECOLOGICAL RISK AND CONSERVATION DECISIONS IN BIG DARBY CREEK

We have conducted research for several years in the Big Darby Creek watershed of central Ohio to examine how U.S. EPA's Ecological Risk Assessment process can be applied to species conservation and water quality improvement (Erekson et al. 2005). The Big Darby is recognized regionally for its scenic beauty, generally good water quality, the unique biological diversity it supports, and the recreational opportunities it affords residents of the nearby city of Columbus, Ohio. The studies have focused on the portion of the Big Darby landscape that faces likely suburban development over the next 20 years, and on the fact that this stream is home to a unique assemblage of rare and endangered freshwater mussel and fish species. Stream conditions are being monitored in terms of the Index of Biotic Integrity (Cormier and Smith 1996). Results show that at least three tributary streams, those influenced by the spreading out of Columbus, have experienced degradation of water quality and associated loss of biotic diversity.

The risks to resources in the Big Darby landscape derive from local agricultural practice (which is gradually improving), as well as from the growth of suburban land use. These changes are taking place over an extended time scale, but in combination they add gradually to the risk of water quality impairment and species loss in the main Big Darby stream. The risks could be reduced through good land management practices for both urban and agricultural land (Cormier and Smith 1996, Cormier et al. 2000), practices that are required generally by federal law, but which must be implemented and enforced through state and local governments. Agricultural use currently occupies 92.4% of the land, with an estimated 1,170 farms in the watershed. Under the present pattern of development, the small towns and rural-area industries do not have greater impacts on water quality than agriculture, but substantial urbanization is occurring in eastern portions of the watershed adjacent to the city of Columbus. To understand the risks to water quality in the main stem, and in selected tributaries, we need to consider both the ecological scale at which effects are expressed within the watershed, and the scale (and information flow) through which decisions are made to protect (or degrade) water quality in this multi-county area.

We investigated the public's willingness to pay to prevent risks to both water and biological diversity (Erekson et al. 2005), and in the course of that work considered the results in relation to hierarchical policy systems and decision-making in the Darby watershed. Reducing the non-point source (NPS) pollution from farmland has been a federally legislated goal, but implementation is to be accomplished through local zoning and agricultural incentive programs, largely a responsibility of the states and local government. Like other states, however, Ohio state agencies must listen to local expressions of concern, including the erosion of private property rights and the increasing costs to farmers being imposed by regulation. Accordingly, responsibility for decisions to minimize non-point source pollution devolves into a web of institutional relationships and incentives with three or four levels of government involved.

At the most local level, responsibility for certain land-use zoning decisions can be seen as the responsibility of township trustees. The several scales at which stakeholders function and decisions get made are outlined implicitly in Table 17.2. Clearly, many more local stakeholder groups are involved than there are higher order interests, considering both the kinds of individuals and the responsibilities of institutions listed in Table 17.2. Similarly, hierarchy theory tells us that many more local decision-makers should be recognized at the lower levels than at the state or federal level, as is expected in any hierarchical structure. Because the federal water quality legislation was passed many years ago, we also see an element of temporal scale here, as over time, more and more local institutions seek "balanced" implementation of higher order regulations.

Table 17.2. *Elements of scale and authority evident among stakeholders and decision makers concerned with water quality in the Big Darby Creek Watershed of Ohio.*

Stakeholders	*Decision-Makers*
Farm owners and families	Township Trustees
Rural businesses	Town Councils
Non-farm rural residents	County Planning Agencies
Incorporated towns	County Commissioners
Developers and builders	Metro Columbus Sewer District
Sport fishers and hunters	Hog and Egg Producers Associations
Recreational visitors	State and County Health Agency & Inspectors
Columbus Metro Parks (& users)	Bank Loan Officers
Local Conservation NGO's (& members)	State and County Courts
Federal Non-regulatory Conservation Agencies	Ohio EPA Enforcement Agencies Ohio EPA Division of Surface Water U.S. EPA Region V U.S. EPA Office of Water

In the face of varied interpretations of the intent of legislation and associated regulations, concern about unfunded mandates, the relative dearth of known consequences for downstream owners, and the diffuse political influence of those owners, Ohio has modified its management decision process in several ways. The most important change has been in designating a number of multi-county watersheds in the State where water quality targets have been established through public hearings, and strict review of local decisions (followed by enforcement) has been implemented by state agencies. The Big Darby watershed is one such priority site, requiring township and county permitting authorities to consider downstream consequences, subject to state review. As shown in Table 17.2, the state retains authority in these watersheds to overturn local upstream discharge or runoff decisions that could degrade downstream

water quality. This is recognition that information transferred across ecological scale (in the stream) must influence the decisions across equivalent scales among decision-making institutions. Related federal, state and conservation NGO programs also work in parallel with the state regulatory agencies at this large watershed scale. Review by state courts, as shown in Table 17.2, is a crosscutting decision scale that has overturned some aspects of the county and state watershed decision process.

The pattern seen in Table 17.2 captures a variety of scaling elements, beginning with direct transfers of authority and information from federal agencies to state agencies, and then to the county commissions or their designated agency heads. As we consider these transfers of information (and responsibility), we should also consider an emerging array of new decisions, evident in the capacity (or disinclination) of each scale to carry out enforcement, including investigation and prosecution of violators. Clearly, the lowest scale, with what may be the most severe budget limitations, has tended to show the least prosecutorial capacity. The indeterminate transfer of only implicit authority across scale illustrates a kind of diffusion of responsibility that adds to uncertainty about policy choices and decision-making at these fine scales. We also need to consider the evidence of up-scaling, as local residents and farm owners challenge top-down decisions and avoid compliance. Often they are supported in their decisions by mid-scale farm produce trade associations and local-scale township trustees or county commissioners. Indeed, during 2002, state legislation was passed in Ohio authorizing the State Department of Agriculture to take over from the Ohio Environmental Protection Agency all the permitting of wastewater discharges from industrial hog and egg producers, an indication of a state-wide aggregated response at mid-scale, adding further uncertainty as to who has long-term authority and power at this scale.

This Big Darby case study, therefore, demonstrates the principle that a hierarchy in ecological scale, seen in watershed systems is associated with a hierarchy in decision-making and concomitant power. However, this hierarchy of decision-making is not the simple command and control relationship often implicit in the idea of a decision-maker. Scale effects in water-related decisions are seen to involve information flow upward across scales, some of the time, and downward across scales at other times, contributing further to diffusion of authority and uncertainty.

17.4 CASE STUDY II: TOTAL MAXIMUM DAILY LOAD REGULATION IN MILL CREEK, CINCINNATI

For situations found when conventional control of permitted sources of pollution fail to achieve desired water quality standards, the U.S. Clean Water Act has other provisions to reduce all pollutant sources. A plan is mandated through Total Maximum Daily Load (TMDL) regulations to reduce all sources of pollutants to within a specified daily limit. Mill Creek is a stream in the greater Cincinnati area of Ohio that flows from its semi-rural headwaters through suburbia and into downtown Cincinnati (Stone 2004). It has been identified by federal and state authorities as having failed conventional improvement and now requires a TMDL to reduce current and future pollutant loads. For Ohio, and therefore for the Mill Creek stream, the water quality goal mandated by the federal Clean Water Act has been defined

biologically as meeting a threshold level of the Index of Biotic Integrity (IBI). An IBI score of 40 or higher (Division of Surface Water 2001) would meet Ohio's minimal requirement for water quality in warm water streams. IBI is calculated based on the abundances and variety of fish sampled, and ranges from 12 to 60. Because IBI encompasses both the top and the bottom of in-stream food chains, it can be viewed as an indicator across scale of many environmental stressors and resultant biological conditions (Karr and Chu 2000, Stone 2004).

Conceptual and practical problems of scale and uncertainty arise in preparing any part of a TMDL implementation plan. The problems mostly arise from the upstream to downstream effect gradient, described for the Big Darby. The effects are primarily associated with cumulative land use impacts along the length of the stream. The land use impacts flow from the cumulative decisions regarding land development, storm runoff and wastewater discharges, all made by local municipalities with certain state and federal agencies reviewing them. Local decision processes tend to promote local growth, often at the expense of downstream water quality. Given the presence of 24 municipal jurisdictions within the Mill Creek watershed, TMDL planning and the considerations of scaling inputs and outcomes become extremely complex.

Conflict arises when water quality objectives that are set nationally by the Clean Water Act cannot be met except through what is perceived to be great (and unequal) costs to local communities. When a TDML plan requires all pollutant discharges to the stream to be reduced, even modestly, stakeholders who once held a right to non-point source discharges, or held pre-existing discharge permits, feel imposed upon, and challenge the TMDL through up-scaling processes. What is interesting is that, despite very costly measures to retrofit and restore modest levels of IBI in the downtown main stem of Mill Creek, other municipalities upstream are making plans to increase the release of storm-water flows and associated pollutants, potentially degrading IBI in the midsections of the watershed.

As seen in the Big Darby case study, therefore, an apparent loss of power and authority, along with increased uncertainty, characterizes the transient from large-scale to local scale control and back again. Tension is expressed between competing communities, as well as across scale, as local-scale stakeholders resist measures introduced at the larger scale, and federal representatives resist initiatives for local exceptions. The result is an interaction among up-scaling and down-scaling processes, as decision responsibility is passed across levels of scale, greatly increasing uncertainty in some decisions. The diffusion of authority during downscaling may result from cross-scale information processes directly, or may result from the strength of resistance and up-scaling passed back to the state and federal agencies. In the Mill Creek case, the up-scaling component is driven, in substantial part, by scaling elements implicit in the wastewater treatment costs necessary to achieve compliance with the TMDL plan (Stone 2004).

17.5 DISCUSSION AND CONCLUSIONS

As noted in Wu and Li's discussion (Chapter 1) of kinds of scale, environmental planning and management frame what they identify as "policy scale." The Mill

Creek and Big Darby case studies presented here provide field models of how scaling, as information-flow hierarchically, functions within ecological and sociopolitical systems. Both the variables (or aggregated variables) and the processes in these decision-making models illustrate functions for information flow and sources of uncertainty. Information is needed at each scale and decisions are expected from institutions at each scale (state, local and national). Formalizing the relationships among variables and processes (especially enforcement, for example) would aid in understanding scaling effects directly, but jurisdictional rights and prerogatives inevitably add to uncertainty from one geopolitical area to another. Also, as suggested by Gunderson and Holling (2002) and by Stone (2004), decision outcomes can be very different depending on whether the institutional dynamics (from local to national scales) are in the conservation phase or the release phase of the adaptive cycle (Stone 2004).

We have looked at the theory underlying scaling for public policy, and at two water-quality case studies for decisions implementing a policy. We find that down-scaling relationships are at least as important in water quality as up-scaling is for landscape and other ecological scaling applications discussed in previous chapters. Uncertainty is inherently prominent in decision making, of course, but the uncertainty is not due to pixel resolution, imprecise measurements, or from an inadequate model. Rather, the sources of uncertainty in scaling for policy and decision-making appear to derive from the blurring of responsibilities across scale, or diffuse information flows across scale. The uncertainty at fine-grained scales of decision-making for water quality cannot be expressed numerically, but can be understood and is predictable much of the time.

Both the theory and the case studies presented here show that describing the way that information flow is affected by scale can greatly help our participation in decision-making processes. Describing and modeling the sources of uncertainty in information flow across scale has a potential for improving public dialog with respect to the decisions being made. Understanding of the scaling among ecological relationships within watersheds, together with the scaling for decision-processes, is a pressing need. Probably most important, this review shows that an appreciation of who benefits and who loses among federal, state and local stakeholders after certain decisions is critical for evaluating outcomes. Understanding of legal and institutional authority, hierarchically, should complement our knowledge of interaction and complexity across scale (from biophysical to social/political). It is this knowledge that will help most in achieving compatibility between social, economic, and ecological systems that are now seen as being at odds with one another, partly due to scale and uncertainty effects.

REFERENCES

Allen, T. F. H., and T. B. Starr. 1982. Hierarchy: Perspectives for Ecological Complexity. University of Chicago Press, Chicago.

Erekson, O. H, O. L. Loucks, S. R. Elliott, D. S. McCollum, M. Smith, and R. J. F. Bruins. 2005. Evaluating development alternatives for a high-quality stream threatened by urbanization: Big Darby Creek watershed. Pages 227-252 *in* R. J. F. Bruins, and M. T. Heberling, editors. Economics and Ecological Risk Assessment: Applications to Watershed Management. CRC Press, Boca Raton, FL.

Gaventa, J. 1980. Power and Powerlessness: Quiescence and Rebellion in an Appalachian Valley. Clarendon Press, Oxford.

Heberlein, T. A. 1974. The three fixes: technological, cognitive and structural. Pages 279-296 *in* D. Field, J. C. Barren, and B. F. Long, editors. Water and Community Development: Social and Economic Perspectives. Ann Arbor Science Publishers, Inc., Ann Arbor, MI.

Holling, C. S., L. H. Gunderson, and G. D. Peterson, editors. 2002. Panarchy: Understanding Transformations in Human and Natural Systems. Island Press, Washington, D.C.

Karr, J. R., and E. W. Chu. 1999. Restoring Life in Running Waters, Better Biological Monitoring. Island Press, Washington, D.C.

Ohio Environmental Protection Agency. 2001. Total maximum daily loads for the Mill Creek in Butler and Hamilton Counties. Division of Surface Water, Ohio Environmental Protection Agency.

Ostrom, E. 1990. Governing the Commons: The Evolution of Institutions for Collective Action. Cambridge University Press, New York.

Peterson, G. 2000. Political ecology and ecological resilience: an integration of human and ecological dynamics. Ecological Economics 35:323-336.

Sabatier, P. A. 1999. Theories of the Policy Process. Westview Press, Boulder, CO.

Sample, V. A. 1990. The Impact of the Federal Budget Process on National Forest Planning. Greenwood Press, New York.

Stone, C. N. 1980. Systemic Power in community decision-making: a restatement of stratification theory. American Political Science Review 74:978-990.

Stone, H. J. 2004. The Political Economy of Ecological Research: Analyzing the "Total Maximum Daily Land Process" in the Mill Creek (Cincinnati). Ph.D. Dissertation, Miami University, Oxford, OH.

PART III

SYNTHESIS

CHAPTER 18

SCALING WITH KNOWN UNCERTAINTY:

A Synthesis

JIANGUO WU, HARBIN LI, K. BRUCE JONES,
AND ORIE L. LOUCKS

18.1 INTRODUCTION

Scale is a fundamental concept in ecology and all sciences (Levin 1992, Wu and Loucks 1995, Barenblatt 1996), which has received increasing attention in recent years. The previous chapters have demonstrated an immerse diversity of scaling issues present in different areas of ecology, covering species distribution, population dynamics, ecosystem processes, and environmental assessment. Scale issues occur in every facet of ecological research, including study design, data collection, experimentation, statistical analysis, and modeling. The scales of observations and outcomes in the case studies range from plots, ecosystems, landscapes, to regions.

Readers will surely ask then, what new synthesis can be achieved from these and other recent contributions to the literature on scale? We see several overarching themes evident in the theory, methods, and case studies presented here, not necessarily in every chapter, but from the body of work as a whole. The following themes are illustrative: novel ideas for integrating diverse scaling perspectives, distinctions among sources of uncertainty, advances in the quantification of scaling error, improved applications of scaling principles, improved recognition of the phenomenon of scale effects (especially for cross-scale material exchange of chemicals, gases, etc.), and advances in the use of scale-related understandings for public policy and decision-making.

Taken together these themes can be understood and organized by thinking through three closely related scale issues: identifying characteristic scales, understanding scale effects, and developing methods for scaling and quantifying sources of error in relation to uncertainties. In this last chapter of the book, we attempt to build from the richness of the methods and case studies toward an integration of the entire volume. To do this we briefly recapitulate scale and scaling concepts, summarize how different kinds of scale issues are dealt with in the

J. Wu, K.B. Jones, H. Li, and O.L. Loucks (eds.),
Scaling and Uncertainty Analysis in Ecology: Methods and Applications, 329–346.

chapters, and present a synthesis in the form of a pluralistic scaling paradigm. In the end, we conclude with some general guidelines for scaling.

18.2 WHAT HAVE WE LEARNT ABOUT SCALE AND SCALING?

In the past two decades, scale and scaling have become a central issue in biological and earth sciences. While many concepts exist, a comprehensive conceptual framework of scale and scaling is still lacking. To address this gap we first need to answer the question, what do authors really mean by "scale" and "scaling?"

Diversity of concepts is not necessarily a problem in development of a new area of science or discipline, but divergence of concepts without addressing a common set of key questions can be a profound problem (Wu and Hobbs 2002). The issues may be manageable when the same terms have been used with only small differences in connotations across disciplines, but major problems arise when the terms are used without clear definitions. To achieve a comprehensive understanding of scale issues, therefore, the full range of concepts relating to scale and scaling in ecology need to be compared and contrasted in a coherent framework. This has not always been accomplished in the chapters of this book, but it is one of the main objectives of the book.

Chapter 1 introduced the definitions of scale and scaling used in disciplines ranging from physical to social sciences, and proposed a three-tiered conceptual framework: dimensions, kinds, and components of scale. Space, time, and levels of organization are the three common dimensions in discussion of scale issues, evident in many of the proceeding chapters. Although there are general scaling rules common to the three dimensions, the behavior of one phenomenon across scales may differ significantly when examined in each of these dimensions. Time and space emerge as the most fundamental dimensions for scaling. Scaling across hierarchical or integrative levels of organization, which inevitably involves change in time and space, is also important in many studies. As hierarchy theory suggests, response patterns at higher levels of organization tend to be massive and slow, while phenomena at lower levels tend to be fine-grained and fast. Thus, scaling in the three dimensions can be related to one another through space-time correspondence principles along with hierarchy theory. As shown in the case study chapters, observational, experimental, modeling and policy scales can all be distinguished from the intrinsic scale of a phenomenon within each dimension. Each different kind of scale has its own meaning, as determined by a variety of factors, and these do not necessarily correspond to the intrinsic scale of phenomena. In the practice of scaling, or to develop quantitative relationships across scales, the components by which scale is defined (e.g., extent, grain, and coverage) also have to be specified.

Ecological scaling as the study of organism-based allometry has existed for several decades, but the recent burst of interest in spatial scaling coincided with the rapid development of landscape ecology in the past two decades. While the term, scale, has acquired a score of connotations in different sciences, the early definition of scaling used in physics and biological allometry has proven to be too narrow for development of a science of scaling in ecology. Our search of papers on scaling

published in Nature and Science (using the ISI Web of Knowledge[SM]) has shown that nearly all of them deal with power-law scaling one way or another. However, a review of the ecological literature and of the chapters in this book indicates that scaling is more than the search for power laws or systematic size relationships. These sources show that ecological scaling includes, but is more than, organism-centered allometric studies. Scaling is generally defined as the translation of information across spatial, temporal, and organizational scales in this book.

This general definition of scaling neither prescribes its goal as the search for power-law relationships, nor as documentation of narrowly-defined scale-invariant phenomena. This synthesis chapter, therefore, adopts the above broad definition of scaling, with emphasis on the translation of information across space. Why is this broad definition necessary? As the previous chapters have shown, ecological patterns and processes can be related in a number of different ways across scales, and pluralistic theories and methods will be needed to discover them. Power laws are elegant and compelling when they are found to exist, but most scaling issues in practice cannot be equated to a search for such simplistic relationships (see also the next section). Accordingly, a range of scaling methods have been developed in the case studies to deal appropriately with the broad range of scaling problems encountered.

18.3 DEALING WITH SCALE ISSUES

Current literature in ecology, and in this book, requires that three types of scale issues be distinguished: characteristic scales, scale effects, and scaling and associated uncertainty. The chapters of this book have dealt with these issues through a variety of objectives and from different perspectives, as illustrated by the summary in Table 18.1. In this section, we provide a systematic overview of how these three types of scale issues have played out, using material from the chapters as well as from recent literature on scale and scaling.

18.3.1 Characteristic Scales

Characteristic scales are "intrinsic scales" on which phenomena of interest operate, and thus are central to description and understanding of the phenomena (Wu and Li, Chapter 1). Characteristic scales are intrinsic because they are inherent to the system to be observed and do not change at the pleasure of the observer. However, because they are usually determined through observation and analysis, characteristic scales have the possibility of being distorted or misrepresented, which leads to the problem of scale mismatch between the intrinsic and observed scale. In general, fine-grained sampling schemes tend to generate data that blur coarse-scale patterns (i.e., high noise/signal ratio), whereas coarse-grained sampling schemes will surely miss fine-scale patterns. Thus, in any study it is critically important to choose a scale that is commensurate with the characteristic scale of the phenomenon of interest based on relevant empirical knowledge or through an exploratory scale analysis.

Table 18.1. *Major objectives, system properties, scale domains, and scaling issues covered in each chapter of the book.*

Chapter	*Main Objectives*	*System Properties*	*Scale Domain*	*Scaling Issues*
1	• to review the concepts of scale and scaling and propose a three-tiered conceptual framework	• generic	• all scales	• concepts and definitions • characteristic scales • scale effects • scaling
2	• to review perspectives and methods of scaling in different disciplines	• generic	• all scales	• scaling theory and principles • upscaling methods • downscaling methods • problems and prospects
3	• to review major methods in uncertainty analysis as applied in ecology and earth sciences	• generic	• all scales	• concepts of uncertainty • sources of uncertainty • methods for estimating upscaling errors
4	• to develop new multilevel statistical models to link ecological variables at different scales	• generic	• local ecosystem to region	• hierarchical scaling • cross-scale relationships linking multiple organizational levels • characterizing uncertainty at different levels
5	• to develop new statistical methods to link distribution and abundance estimates of tree species	• distribution and abundance of species	• local community to landscape	• downscaling • estimating species abundance based on presence/absence data • uncertainty estimation using variance in specie detectability
6	• to compare and contrast intact and model experimental systems for upscaling biogeochemical processes	• biogeochemical processes in ecosystems	• controlled environment • ecosystem or landscape	• upscaling with experiments in space and time • internal validity vs. external validity • upscaling uncertainties with experiments

Table 18.1 (contd.)

Chapter	*Main Objectives*	*System Properties*	• *Scale Domain*	*Scaling Issues*
7	• to discuss an upscaling framework for estimating aboveground net primary production and associated uncertainties	• carbon dynamics • aboveground net primary production	• landscape or region	• upscaling as spatial extrapolation • spatial modeling • reducing scaling uncertainty by spatially decomposing heterogeneous landscapes
8	• to examine the methods and problems of upscaling with remote sensing	• land surface characteristics • remote sensing • biophysical variables	• landscape, region, or beyond	• upscaling with remote sensing • uncertainties due to heterogeneity and nonlinearity • aggregation error • scale effects
9	• to demonstrate a spatially explicit upscaling approach using field observations, remote sensing tools, and a biogeochemistry model	• carbon pools • fluxes • net ecosystem production	• landscape or region	• characteristic scales • upscaling using a spatially nested hierarchy • estimating uncertainty at different spatial scales
10	• to explore challenges and solutions in upscaling nitrogen gas fluxes in landscapes	• nitrogen gas fluxes in ecosystems • landscape-scale variability	• watershed or landscape	• characteristic scales • spatial upscaling with "representative fluxes" and ecosystem modeling • temporal scaling of nitrogen fluxes • sources of uncertainties
11	• to relate stream nitrogen concentrations to regional atmospheric nitrogen deposition, watershed-level land surface characteristics, and within-riparian-zone land cover composition	• total stream nitrogen • biophysical factors at local, watershed, and regional scales	• landscape or region	• characteristic scales • linking biophysical variables at different scales • scaling functions • sources of uncertainty • estimating uncertainty using variance measures

Table 18.1 (conta.)

Chapter	*Main Objectives*	*System Properties*	*• Scale Domain*	*Scaling Issues*
12	• to examine scaling relations of nutrient export to watershed resolution	• nitrogen and phosphorous export from watersheds	• watershed or landscape	• scale effects • scaling relations • quantifying scaling errors
13	• to explore uncertainty issues in integrated, regional-scale assessments using coupled models (meta-models)	• land use change • bird population dynamics	• landscape or region	• upscaling through coupling models at different scales • sources of uncertainties • quantifying uncertainties
14	• to assess scale dependence of relationships between avian nesting success and forest fragmentation	• fragmented forests • bird population dynamics	• landscape or region	• characteristic scales • scale effects • cross-scale relationships • scale-related uncertainties
15	• to quantify scale effects and uncertainties in mapping riparian zones	• agricultural landscapes • riparian systems	• landscape or region	• scale effects • scaling relations • errors in mapping and classification
16	• to explore scaling issues in assessing relationship between lake water quality and lakeshore development	• lake and watershed interactions • water clarity	• watershed or region	• characteristic scales • scale effects • between-scale influences
17	• to examine scaling issues in decision-making processes on water quality through a hierarchy of institutions	• water quality • resource planning and policy • decision-making hierarchies	• watershed or region	• characteristic scales • hierarchical scaling in policy-development and decision-making • relationships between ecological and geopolitical scaling • sources of uncertainty
18	• to provide a synthesis of the book	• generic	• all scales	• characteristic scales • scale effects and scaling in general

Different scales of observation or policy-making may lead to disparate outcomes (see, for example, Wessman and Bateson, Chapter 8, Groffman et al., Chapter 10, Loucks et al., Chapter 17). Indeed, scale mismatching may have been one of the most common problems in ecological studies. Such problems may be a consequence of a flawed study design in which mismatches between different kinds of scale (intrinsic, observational, experimental, analytic, modeling, and policy scale) are encountered. This situation is also an example of scale effects, which will be further discussed. Therefore, to describe and understand a given phenomenon, there may be no single correct scale (Levin 1992, Wu and Levin 1994), but there are certainly scales that are more revealing than others.

The idea of characteristic scale appears to be at odds, however, with the often misinterpreted prevalence of scale-invariant phenomena in nature, inferred frequently from theories such as fractal geometry and self-organized criticality (e.g., Bak 1996). Recent studies based on these theories have claimed that ecological systems are characterized by self-organized criticality and self-similarity, and exhibit scale invariant patterns over several to many orders of magnitude (e.g., Bak 1996, Jørgensen et al. 1998, Sole et al. 1999, Brown et al. 2004). However, others have pointed out that some of these analyses were problematic because of misinterpreting ecological data or overreaching from the results (e.g., Raup 1997, Kirchner and Weil 1998, Dodds et al. 2001, Plotnick and Sepkoski 2001, Cyr and Walker 2004). Studies of both biophysical and socioeconomic systems have shown much evidence that complex systems often exhibit both scale-dependent behaviors and characteristic scales (Clark 1985, Courtois 1985, Urban et al. 1987, Delcourt and Delcourt 1988, Holling 1992). Such findings are consistent with the prediction from hierarchy theory that patterns and processes in complex systems tend to have distinctive characteristic scales, through both internal self-organization and multiple-scale external constraints (O'Neill et al. 1986, Schweitzer 1997, Wu 1999). Several chapters of this book also provide evidence to support the presence of distinctive characteristic scales, illustrated in the context of carbon cycling (Law et al., Chapter 9), nitrogen fluxes (Groffman et al., Chapter 10), avian population dynamics (Lloyd et al., Chapter 14), lake-watershed interactions (Johnston and Shmagin, Chapter 16), and policy-making processes (Loucks et al., Chapter 17).

In reality, neither all patterns and processes always have a clearly identifiable hierarchical structure, nor do they all exhibit scale-invariant behavior. These two perspectives should be viewed as complementary, rather than opposing to each other. For example, as discussed by Wu and Li (Chapter 2), extrapolation along a scaling ladder (or the hierarchical patch dynamics scaling approach) integrates both perspectives. Dealing with scale issues requires as much appreciation of scale-dependent phenomena as seeking scale-invariant instances. For either one, the kinds of phenomena and the ranges of scale (or scale domains) in which scale-dependence or scale-invariance occurs must be specified if the higher-level, comprehensive integration is to be achieved. The existence of characteristic scales suggests that scale analysis should be a necessary first step in dealing with complex phenomena (Levin 1992, Wu and Loucks 1995). Numerous landscape metrics and spatial statistical methods can be used for this purpose (Turner et al. 2001). Jones et al. (Chapter 11) provides an example of using classification and regression trees

(CART) to analyze relationships between total stream nitrogen and its controlling variables at local riparian, watershed, and regional scale. While multiple-scale dynamic models are commonly used for this purpose (Wu and Li, Chapter 2), statistical methods such as multilevel statistical models (Berk and Leeuw, Chapter 4) may also be effective in identifying and linking characteristic scales.

18.3.2 Scale Effects

Scale effects occur whenever changes in the scale of observation, analysis, modeling, or experimentation lead to changes in the results of a study. The idea of characteristic scales suggests that scale effects are bound to occur whenever the scale of observation involves a mismatch with the intrinsic scale of a phenomenon. Such effects, although generally expected, may not be specifically predictable. In contrast, theories of scale-invariance and self-similarity tend to imply that scale effects either do not occur or can be readily predicted mathematically. Empirical studies have shown that scale effects may result in inaccurate classifications or distorted maps (see Wessman and Bateson, Chapter 8 and Hollenhorst et al., Chapter 15), and altered or erroneous statistical and modeling results (see Jones et al., Chapter 11, Wickham et al., Chapter 12, Lloyd et al., Chapter 14, Johnston and Shmagin, Chapter 16). Bradford and Reynolds (Chapter 6) show that, in experimental studies, scale effects may be more common than ecologists tend to admit when microcosms or artificial systems are used to mimic natural systems. In this case, a crucial issue at stake is the tradeoff between the internal and external validity of experiments (Naeem 2001). Thus, scale effects do seem often to impede our ability to accurately interpret the results of a study, be it observational or experimental, and add to the uncertainty of scaling operations. In general, an increase in grain size may lead to lower variability in system variables due to averaging or smoothing effects, while an increase in extent may lead to higher variability due to the inclusion of more diverse conditions.

The studies in this book and elsewhere show that scale effects are pervasive in natural and social systems, and commonly found in basic research studies as well as in policy-making and political processes (Loucks et al., Chapter 17). It is interesting to note the problem of gerrymandering, dating from more than a hundred years ago, as an example of scale effects as well as the interaction between observational science and social and political processes. Elbridge Gerry (1744-1814), the governor of Massachusetts from 1810 to 1812, signed a bill into law that redistricted the state allegedly to benefit his Republican Party in elections. As a result of the redistricting, one of the congressional districts was shaped like a salamander, and the term gerrymander was derived from the two words: Gerry and salamander (http://webster.com/). The purposeful manipulation of the local boundaries of electoral districts (i.e., changing grain size and configuration) altered the outcome of an election process at a larger scale. Scale effects have long been studied in human geography as part of what has been known generally as "the modifiable area unit problem" or MAUP (Openshaw 1984). Although MAUP studies clearly are relevant to understanding scale effects and space-scale interactions in general, the subject has generally been ignored in ecological literature until recently (Jelinski and Wu 1996).

In parallel, plant community ecologists have long studied the effects of changes in sample size and position on vegetation pattern results from field surveys. The diversity of studies on scale effects in the previous chapters demonstrates that, today, most ecologists are aware of these effects. However, no discipline outside landscape ecology, which focuses on the relationships among pattern, process, and scale (Turner et al. 2001, Wu and Hobbs 2002), has placed more emphasis on understanding scale effects.

Of course, scale effects may also be artifacts if the scales of study are entirely arbitrary, in which case the actual patterns and processes become distorted. When the scales of study are determined based on understanding of the phenomena of concern, however, observed scale effects can be used to improve understanding of scaling relationships and the accuracy of scaled outcomes (Jelinski and Wu 1996, Wu 2004). Hence, future studies of scale effects will have to move beyond merely reporting their occurrence to focus work on the development of more sophisticated scaling relations and scale-dependent understanding (Wu 2004).

18.3.3 Approaches to Scaling

As the previous chapters have shown, scaling has become an increasingly important element of ecological research. While ecologists are among those who are most aware of scale issues, most scaling theories and methods have originated in physics, meteorology, and hydrology, and some of these methods have remained underutilized in ecology. Chapters 1 and 2 have reviewed the full range of scaling methods, breaking them into two complementary general approaches according to their conceptual foundations: the similarity-based scaling approach, widely used in geophysical and biological sciences, is rooted in the idea of similitude or self-similarity, whereas the dynamic model-based approach includes scaling methods that emphasize processes and mechanisms. Similarity-based scaling methods may start with first principles and proceed deductively with mathematical analysis (the analytical approach), or seek scaling relations inductively with statistical regressions (the empirical approach).

For similarity-based scaling, methods available from current literature and the earlier chapters include dimensional analysis, similarity analysis, biological allometry, and spatial allometry, all of which draw on the principles of similarity (geometric, physical, and functional) and self-similarity (fractal scaling). Dimensional and similarity analysis are fundamental to modeling and scaling in general, but we have not yet seen how effective these methods are for complex ecological and socioeconomic processes that are not explained well by physical laws alone. Biological allometry, where the techniques of dimensional and similarity analysis are invoked often, has dominated the literature in "ecological scaling" for many years. However, organism-based allometric scaling may have little relevance for spatial scaling problems, unless space can be incorporated into the scaling relation through, for example, population density or home range. In contrast, spatial allometry relates ecological variables directly to spatial scale, facilitating cross-scale predictions when the domains of applicability can be determined. While this book is

not focused on the similarity-based methods, a review of them, as used in biology and geophysical sciences, has been provided by Wu and Li in Chapter 2.

Dynamic model-based scaling includes explicit upscaling and downscaling methods. Wu and Li (Chapter 2) review several upscaling methods used in ecology and geophysical sciences, including extrapolation by lumping, extrapolation by effective parameters, direct extrapolation, extrapolation by expected value, explicit integration, spatially interactive modeling, and hierarchical scaling. A major difference among these methods lies in how spatial heterogeneity is treated in the model-based scaling procedure. Extrapolation by lumping essentially ignores spatial heterogeneity, and is more likely to produce results with high uncertainty. Extrapolation by effective parameters treats spatial heterogeneity in an aggregated way, and has had success in hydrology and meteorology. It may be equally useful in scaling up population and ecosystem processes in situations where the procedures for deriving effective parameters are applicable. Both direct extrapolation and extrapolation by expected value treat spatial heterogeneity in explicit ways, and are widely used methods in ecology and earth sciences. Explicit integration, although probably the most elegant and accurate, is not generally practical. When horizontal flows, time delays, and feedbacks become significant and when spatial heterogeneity can no longer be decomposed discretely or characterized statistically, spatially interactive modeling may be the only sensible alternative.

The discussion by Peters et al. (Chapter 7) treats several upscaling methods in terms of their degrees of spatial explicitness (also see Peters et al. 2004). In their view upscaling methods are of three kinds: non-spatial, spatially implicit, and spatially explicit. The non-spatial method refers mainly to extrapolation by lumping; the spatially implicit method to direct extrapolation, and the spatially explicit method to spatially interactive modeling. These authors also illustrate how and when the different scaling methods should be used through an example of the extrapolation of net primary production in a desert landscape. We should note that the definition of "spatially implicit" in Chapter 7 (and also Peters et al. 2004) is different from that commonly used in ecology and earth sciences. Spatial explicit models usually refer to those that consider spatial interactions of processes of interest explicitly, or represent the spatial locations of model variables or parameters explicitly. Thus, extrapolation by lumping and extrapolation by effective parameters both are spatially implicit methods. Extrapolation by expected value incorporates spatial heterogeneity in terms of probability density functions and thus is quasi-spatial; direct extrapolation, explicit integration, and spatially interactive modeling are all spatially explicit methods. With rapidly increasing computational capabilities and available remote sensing data, direct extrapolation and similar methods are becoming the most widely used approach in landscape and regional case studies.

Although most of the chapters in this book deal with upscaling, Wu and Li (Chapter 2) reviewed the two major downscaling approaches: empirically-based statistical downscaling and downward nested modeling. In contrast with upscaling, downscaling seeks to derive detailed patterns within a spatial domain by disaggregating coarse-grained information. A number of sophisticated statistical and modeling techniques have been developed for downscaling the outputs of General Circulation Models (GCMs) from regions to local landscapes or ecosystems and for

estimating fine-scale patterns of hydrological and soil properties from coarse-grained information (see Chapter 2 and references therein). While many downscaling studies have been carried out in the context of global climate, hydrological, and soil sciences, He and Reed (Chapter 5) present a new downscaling method for a time-honored ecological problem–linking species distribution to abundance. Their statistical models, based on the combinatorial theory of occupancy, allow for estimation of the number of organisms (abundance) from species presence-absence maps (distribution). Not surprisingly, the accuracy of these models was found to decrease with decreasing map resolution, a manifestation of scale effects and a source of uncertainty. This method is similar to the statistical downscaling methods reviewed by Wu and Li (Chapter 2) in that they all assume some statistical distribution of the variable to be downscaled and then seek model parameters that satisfy the assumption. However, He and Reed's method (Chapter 5) is suited for discrete variables, whereas most other downscaling methods deal with continuous variables associated with hydrological, soil, and climatic processes.

Which of these scaling methods should be chosen for specific research problems in ecology? In practice, it is frequently the case that several different scaling methods are used together in a single study. This has been true of complex scaling projects that either cover a wide range of scales or consider a diversity of processes (Reynolds and Wu 1999, Wu 1999, Law et al., Chapter 9). Also, models that are spatially more realistic tend to have higher explanatory potential, but not necessarily higher predictive accuracy. For a particular scaling problem, therefore, one cannot expect a single best method or approach; some methods may be more effective and accurate for certain goals than others. Therefore, the choice of scaling methods should be resolved in relation to the purpose of the study, the acceptable level of uncertainty, and data availability. Multiple methods are usually preferred for purposes of comparison and confirmation.

18.3.4 Approaches to Uncertainty

No matter what scaling methods are used, uncertainty in scaling is inevitable due to spatial heterogeneity, nonlinear relationships, lack of reliable data, and problems in scaling techniques. All of these are illustrated by Peters et al. (Chapter 7) and Urban et al. (Chapter 13). However, uncertainty analysis, or accuracy assessment, has not consistently been a part of ecological scaling. In Chapter 3, Li and Wu have provided an overview of uncertainty analysis, focusing on the sources of uncertainty, evaluation of scaling algorithms, error propagation, and presentation of prediction accuracy. Uncertainty analysis should be regarded as an essential part of the scaling process because it provides critical information about confidence in the results and the adequacy of the models and algorithms used. The main purposes of uncertainty analysis, therefore, are to quantify the various sources, assess the effects of uncertainty on scaling results, and identify critical factors in models (see He and Reed, Chapter 5, Peters et al., Chapter 7, Law et al., Chapter 9, Urban et al., Chapter 13). The methods used for uncertainty analysis include probability theory, Taylor series expansion, Monte Carlo simulation, generalized likelihood uncertainty estimation, Bayesian statistics, and sequential partitioning.

Several of the earlier chapters provide examples of how to deal with scaling uncertainty. In particular, Peters et al. (Chapter 7) advocate a general approach to reducing scaling uncertainty by dividing a complex landscape into a number of regions for which different scaling methods are selected. The idea behind this approach is a spatial extension of the decomposability principle of hierarchy theory, and consistent with the hierarchical patch dynamics paradigm (Wu and Loucks 1995, Wu 1999). While Law et al. (Chapter 9), Groffman et al. (Chapter 10), Jones et al. (Chapter 11), Wickham et al. (Chapter 12), Lloyd et al. (Chapter 14), and Hollenhorst et al. (Chapter 15) estimate uncertainty using conventional measures such as standard deviations and variances, Urban et al. (Chapter 13) has explored new methods for estimating error propagation and communicating scaling uncertainty to scientists as well as landscape managers and planners. Given the increasing role of large-scale modeling and scaling in ecological research and environmental decision-making (e.g., Johnston and Shmagin, Chapter 16, Loucks et al., Chapter 17), the obligation to understand, report, and reduce uncertainties in scaling are becoming increasingly important.

18.4 TOWARDS A PLURALISTIC SCALING PARADIGM

Inevitably, one must ask now whether some overarching pattern is evident in the material just summarized. The reviews in Chapters 1 and 2 by Wu and Li show that scaling has often been associated with power laws, fractals, and self-organized criticality. Such scaling laws would be elegant and powerful for ecosystems and landscapes if they could be validated through empirical evidence. The recent resurgence of interest in biological allometry is epitomized by the development of a "metabolic theory of ecology" (Brown et al. 2004), which attempts to use organismal allometry with a temperature correction to predict "ecological processes at all levels of organization from individuals to the biosphere." Such grand theory based on first principles in physics, chemistry, and biology, would be eminently useful in ecological scaling, but skepticism and sharp criticisms are rooted in the dearth of empirical support, mathematical limitations, diminishing rigor at organizational levels beyond whole organisms, and an inability to deal with heterogeneous structures and transient dynamics (Dodds et al. 2001, Bokma 2004, Cyr and Walker 2004, Kozlowski and Konarzewski 2004). Can a pluralistic approach be an alternative?

Not all ecosystems and landscapes, or their properties, exhibit fractal characteristics, and "self-organized criticality is not likely to be a universal feature" (Levin 1999). In ecological systems, scale invariance may be common, but scale-dependence is ubiquitous. The previous chapters illustrate the complexity of scaling problems in ecology, and explain why holistic approaches have had limited scientific success, despite often appearing to be of high ecological relevance. At the same time, extreme reductionist approaches, although mechanistically appealing, often fail to resolve ecological problems that hinge on emergent properties, self-organization, and other nonlinear interactions. Ecologists have long called for an integration between the two kinds of approaches, and such an integration could be accomplished in the context of a pluralistic approach (McIntosh 1987, Wu and

Loucks 1995). A pluralistic scaling paradigm would be able to deal with the diverse problems of transferring information across the various kinds of scale. Such a paradigm is implied in the previous chapters as well as in other recent publications.

Pluralism does not mean an anarchic development of views and approaches free of an underlying common framework. As a scaling paradigm, pluralism accepts the organized diversity of scaling problems seen in the previous paragraphs, and discourages exaggeration of a monistic theory or methodology. It allows promoting of alternative but complementary perspectives arising out of interdisciplinary sources. A pluralistic scaling paradigm should start with the clearly defined concepts of scale and scaling that we have sought to provide here. It also reconciles their different connotations within and among disciplines.

The definitional hierarchy outlined in Chapter 1 can serve as a point of departure. Because human influences have become pervasive in all ecological systems, scaling over large areas requires considering explicitly how biophysical and socioeconomic processes interact at different but hierarchically inked scales. Thus, the pluralistic scaling paradigm is inherently interdisciplinary, integrating natural and social sciences. In it the two general scaling methodologies, the similarity-based and dynamic model-based approaches, can be brought together through a complementary, rather than an adversarial, conceptual framework. Hierarchy theory may provide such a scaling framework for both the natural and the social sciences (Wagenet 1998, Marceau 1999, Wu 1999, Haila 2002).

Because all environments have a hierarchical structure (MacArthur 1972), and because "space is inherently hierarchical" (Meentemeyer 1989), a hierarchical frame work for pluralistic scaling is not only intuitive but also captures the essential scale-dependent complexity of biophysical and socioeconomic systems. As a general strategy, the "scaling ladder approach" (sensu Wu 1999) provides general guidelines for decomposing heterogeneous landscapes or regions into nested spatial hierarchies, along which information can be transferred. The scaling ladder approach is based on the hierarchical patch dynamics paradigm (Wu and Loucks 1995) that integrates hierarchy theory with the patch dynamics perspective. The approach has proven useful in scaling landscape patterns and processes (Hay et al. 2001, 2002, Poole 2002, Wu and David 2002, Burnett and Blaschke 2003, Hall et al. 2004, Poole et al. 2004).

While scale-invariance may exist over broad geographic regions in some circumstances, most ecological patterns and processes show scaling thresholds at which abrupt changes in scaling relationships occur, corresponding to shifts in underlying mechanisms. In the hierarchical context of the scaling ladder approach, both similarity-based and dynamic model-based scaling methods are useful for transferring information between adjacent hierarchical levels (or scaling thresholds). To transfer information across a broad range of scales along the scaling ladder (e.g., from single leaves, canopies, ecosystems, landscapes, to regions or the entire biosphere), there may be more scientific justification and technical feasibility through use of a hierarchy of scale-specific models rather than single monolithic models with several hierarchical levels built in (Wu 1999). Such multiple-step procedures require novel model-linking techniques, including nested modeling and meta-modeling (e.g., Reynolds et al. 1993, Wu and David 2002, Urban et al., Chapter 13),

and combine both bottom-up and top-down conceptualizations. Because both ecological and socioeconomic systems are complex adaptive systems (Levin 1999), their structure and function can change in response to changing environments. Such responses need to be accommodated through pluralistic scaling. Accordingly, the structure of scaling ladders–patch hierarchies used for scaling particular ecological patterns and processes–may also change when the time horizon involved is much longer than the characteristic spatial and temporal scales of the phenomenon of interest.

18.5 CONCLUSIONS

Throughout this volume we have tried to view scaling consistently as the process of translating information across space, time, and organizational levels. Scaling is ubiquitous and of paramount importance in ecology. Although ecologists are acutely aware of such issues as characteristic scale and scale effects, the commonly used scaling methods have tended to be inadequate for dealing quantitatively with the spatial heterogeneity and nonlinearity embedded in ecological systems. While the availability of accurate multiple-scale data sets are, and will always be, crucial to successful scaling, we argue that a key impediment to be overcome now derives from the limited scaling methodologies currently in wide use in ecology. The field can benefit significantly from, and contribute to, the development of a coherent science of scaling by embracing a number of theories and methods from the physical and geophysical sciences, and moving forward with an ecologically comprehensive, pluralistic scaling paradigm.

We would fall short of a reader's expectations for a book on scaling if no guidelines for further development of scaling were offered. However, it is still difficult to provide a general "recipe" for scaling considering the idiosyncrasies of many specific scaling problems and the diversity of scaling methods available. Still, the following general guidelines, although by no means inclusive, should be useful for the practice of spatial scaling.

18.5.1 Some general principles for scaling

- The most effective scaling strategies are those that integrate bottom-up and top-down approaches through combining field observations, experimentation, with mathematical modeling. In developing models for scaling, bottom-up approaches supply mechanistic details, whereas top-down approaches provide constraints and boundary conditions.
- The relationships between pattern and process, be they physical, biological, or social, are multifaceted and scale-dependent. Only when pattern and process operate at similar time scales within the same geographic region, can they possibly have interactive relationships. If spatial patterns change much more slowly than the processes that influence them, then the relationship between pattern and process can be reduced to the one-directional effect of pattern on process. This general principle can be used as

a guide for simplifying ecological complexity during study design, and for coupling biophysical and socioeconomic patterns, processes and outcomes from scaling.

- The feasibility and accuracy of translating information across scales depend greatly on properly identifying scaling thresholds. Thus, scale analysis using landscape metrics and spatial statistics should be a first step in scaling. Key processes or variables with similar scales of variation should then be grouped, and examined for potential interactions within each group and for hierarchical linkages between different groups.
- Ecological systems can be considered spatially nested hierarchies *a priori*, or based on cross-scale analyses, which provide the context necessary for scale invariance to be properly interpreted. Nested hierarchies also facilitate mechanistically transferring information across multiple domains of scales.

18.5.2 Selecting appropriate methods for scaling

- Spatial heterogeneity is the most pervasive and critical factor to influence the process of scaling. Accordingly, quantifying spatial heterogeneity at multiple scales, whenever feasible, should be a priority in the early stage of a scaling study. This analysis may provide critical information for selecting appropriate scaling methods and reducing overall scaling uncertainty.
- Scaling methods have to be selected with sensitivity to particular study goals because they differ in efficiency and accuracy. Each is constrained by a different set of assumptions, data requirements, capabilities, and acceptable levels of uncertainty.
- Similarity-based methods, often relying on relatively simple statistics such as regression and correlation, can be quite useful for prediction and for suggesting possible underlying explanations for observed patterns. However, only dynamic modeling methods, based on processes and mechanisms, have the potential to achieve reliable predictions for evolving systems in changing environments.
- Most existing scaling methods operate only by not crossing scaling thresholds or organizational levels. Scaling across multiple levels of organization often requires a hierarchical approach. In particular, most models are scale-specific and should be used only within the domain of scales for which they are designed. Applying models outside their intended scale domains is expected to result in high uncertainties.

18.5.3 Scaling with known uncertainty

- Errors are bound to occur in scaling, and uncertainty analysis must be considered as an integral part of scaling because it provides critical information on the adequacy of models or algorithms used in the scaling process. Thus, it is not adequate simply to ask how to scale; rather, one needs to ask how to scale with known uncertainty.

- Scaling uncertainty comes from the model structure, parameters, driving variables, and scaling algorithms. Errors from these different sources may propagate to produce nonlinear effects on the accuracy of scaling results. Some uncertainties can be quantified and reduced (e.g., measurement and sampling errors); others can be quantified but are hard to reduce (e.g., natural variability in data); and still others may not even be quantifiable (e.g., model uncertainty). Wherever possible, one should identify and reduce the critical sources of errors.
- Scaling results should be presented along with uncertainty measures such as probability distributions, variance, coefficient of variation (CV), confidence levels, and root mean square error (RMSE). Predictions without accuracy information are of little value, and may even be misleading no matter how impressive the numbers appear to be.

The importance of scaling can hardly be overemphasized. Every time an average of some property is derived across space, time, or organizational levels, scaling is at work. The accumulation of our experience and knowledge is essentially a product of scaling. As ecologists, we must be more conscious about scaling and associated uncertainties. To better understand and manage the diversity and complexity of ecological systems, we need to make more efforts to develop a coherent science of ecological scaling. We certainly hope that this book will be able to help, in some ways, to achieve this goal.

ACKNOWLEDGEMENTS

JW's research on scaling has been supported in part by grants from US EPA's Science to Achieve Results (STAR) program (R827676-01-0), US NSF (DEB 9714833, CAPLTER), and National Natural Science Foundation of China (30028002).

REFERENCES

Bak, P. 1996. How Nature Works: The Science of Self-Organized Criticality. Copernicus, New York.

Barenblatt, G. I. 1996. Scaling, Self-Similarity, and Intermediate Asymptotics. Cambridge University Press, Cambridge.

Bokma, F. 2004. Evidence against universal metabolic allometry. Functional Ecology 18:184-187.

Brown, J. H., J. F. Gillooly, A. P. Allen, V. M. Savage, and G. B. West. 2004. Toward a metabolic theory of ecology. Ecology 85:1771-1789.

Burnett, C., and T. Blaschke. 2003. A multi-scale segmentation/object relationship modeling methodology for landscape analysis. Ecological Modelling 168:233-249.

Clark, W. C. 1985. Scales of climate impacts. Climatic Change 7:5-27.

Courtois, P. J. 1985. On time and space decomposition of complex structures. Communications of the ACM 28:590-603.

Cyr, H., and S. C. Walker. 2004. An illusion of mechanistic understanding. Ecology 85:1802-1804.

Delcourt, H. R., and P. A. Delcourt. 1988. Quaternary landscape ecology: relevant scales in space and time. Landscape Ecology 2:23-44.

Dodds, P. S., D. H. Rothman, and J. S. Weitz. 2001. Re-examination of the "3/4-law" of metabolism. Journal of Theoretical Biology 209:9-27.

Haila, Y. 2002. Scaling environmental issues: problems and paradoxes. Landscape and Urban Planning 61:59-69.

Hall, O., G. J. Hay, A. Bouchard, and D. J. Marceau. 2004. Detecting dominant landscape objects through multiple scales: an integration of object-specific methods and watershed segmentation. Landscape Ecology 19:59-76.

Hay, G. J., P. Dube, A. Bouchard, and D. J. Marceau. 2002. A scale-space primer for exploring and quantifying complex landscapes. Ecological Modelling 153:27-49.

Hay, G., D. J. Marceau, P. Dubé, and A. Bouchard. 2001. A multiscale framework for landscape analysis: object-specific analysis and upscaling. Landscape Ecology 16:471-490.

Holling, C. S. 1992. Cross-scale morphology, geometry, and dynamics of ecosystems. Ecological Monographs 62:447-502.

Jelinski, D. E., and J. Wu. 1996. The modifiable areal unit problem and implications for landscape ecology. Landscape Ecology 11:129-140.

Jørgensen, S. E., H. Mejer, and S. N. Nielsne. 1998. Ecosystem as self-organizing critical systems. Ecological Modelling 111:261-268.

Kirchner, J. W., and A. Weil. 1998. No fractals in fossil extinction statistics. Nature 395:337-338.

Kozlowski, J., and M. Konarzewski. 2004. Is West, Brown and Enquist's model of allometric scaling mathematically correct and biologically relevant? Functional Ecology 18:283-289.

Levin, S. A. 1992. The problem of pattern and scale in ecology. Ecology 73:1943-1967.

Levin, S. A. 1999. Fragile Dominions: Complexity and the Commons. Perseus Books, Reading.

MacArthur, R. H. 1972. Geographical Ecology: Patterns in the Distribution of Species. Princeton University Press, Princeton.

Marceau, D. J. 1999. The scale issue in social and natural sciences. Canadian Journal of Remote Sensing 25:347-356.

McIntosh, R. P. 1987. Pluralism in Ecology. Annual Review of Ecology and Systematics 18:321-341.

Meentemeyer, V. 1989. Geographical perspectives of space, time, and scale. Landscape Ecology 3: 163-173.

Naeem, S. 2001. Experimental validity and ecological scale as criteria for evaluating research programs. Pages 223-250 *in* R. H. Gardner, W. M. Kemp, V. S. Kennedy, and J. E. Petersen, editors. Scaling Relations in Experimental Ecology. Columbia University Press, New York.

O'Neill, R. V., D. L. DeAngelis, J. B. Waide, and T. F. H. Allen. 1986. A Hierarchical Concept of Ecosystems. Princeton University Press, Princeton.

Openshaw, S. 1984. The Modifiable Areal Unit Problem. Geo Books, Norwich.

Peters, D. P. C., J. E. Herrick, D. L. Urban, R. H. Gardner, and D. D. Breshears. 2004. Strategies for ecological extrapolation. Oikos 106:627-636.

Plotnick, R. E., and J. J. Sepkoski. 2001. A multiplicative multifractal model for originations and extinctions. Paleobiology 27:126-139.

Poole, G. C. 2002. Fluvial landscape ecology: addressing uniqueness within the river discontinuum. Freshwater Biology 47:641-660.

Poole, G. C., J. A. Stanford, S. W. Running, C. A. Frissell, W. W. Woessner, and B. K. Ellis. 2004. A patch hierarchy approach to modeling surface and subsurface hydrology in complex flood-plain environments. Earth Surface Processes and Landforms 29:1259-1274.

Raup, D. M. 1997. A breakthrough book? Complexity 2:30-32.

Reynolds, J. F., and J. Wu. 1999. Do landscape structural and functional units exist? Pages 273-296 *in* J. D. Tenhunen and P. Kabat, editors. Integrating Hydrology, Ecosystem Dynamics, and Biogeochemistry in Complex Landscapes. Wiley, Chichester.

Reynolds, J. F., D. W. Hilbert, and P. R. Kemp. 1993. Scaling ecophysiology from the plant to the ecosystem: a conceptual framework. Pages 127-140 *in* J. R. Ehleringer and C. B. Field, editors. Scaling Physiological Processes: Leaf to Globe. Academic Press, San Diego.

Schweitzer, F., editor. 1997. Self-Organization of Complex Structures: From Individual to Collective Dynamics. Gordon and Breach Science Publishers, Amsterdam.

Sole, R. V., S. C. Manrubia, and P. Bak. 1999. Criticality and scaling in evolutionary ecology. Trends in Ecology and Evolution 14:156.

Turner, M. G., R. H. Gardner, and R. V. O'Neill. 2001. Landscape Ecology in Theory and Practice: Pattern and Process. Springer-Verlag, New York.

Urban, D. L., R. V. O'Neill, and H. H. Shugart. 1987. Landscape ecology. Bioscience 37:119-127.

Wagenet, R. J. 1998. Scale issues in agroecological research chains. Nutrient Cycling in Agroecosystems 50:23-34.
Wu, J. 1999. Hierarchy and scaling: extrapolating information along a scaling ladder. Canadian Journal of Remote Sensing 25:367-380.
Wu, J. 2004. Effects of changing scale on landscape pattern analysis: scaling relations. Landscape Ecology 19:125-138.
Wu, J., and J. L. David. 2002. A spatially explicit hierarchical approach to modeling complex ecological systems: theory and applications. Ecological Modelling 153:7-26.
Wu, J., and R. Hobbs. 2002. Key issues and research priorities in landscape ecology: an idiosyncratic synthesis. Landscape Ecology 17:355-365.
Wu, J., and S. A. Levin. 1994. A spatial patch dynamic modeling approach to pattern and process in an annual grassland. Ecological Monographs 64(4):447-464.
Wu, J., and O. L. Loucks. 1995. From balance-of-nature to hierarchical patch dynamics: a paradigm shift in ecology. Quarterly Review of Biology 70:439-466.

INDEX

F

G

H

I

K

L

M

N

O

P

R

S

T

U

V

W