

# Spatial Extrapolation: The Science of Predicting Ecological Patterns and Processes

JAMES R. MILLER, MONICA G. TURNER, ERICA A. H. SMITHWICK, C. LISA DENT, AND EMILY H. STANLEY

*Ecologists are often asked to contribute to solutions for broadscale problems. The extent of most ecological research is relatively limited, however, necessitating extrapolation to broader scales or to new locations. Spatial extrapolation in ecology tends to follow a general framework in which (a) the objectives are defined and a conceptual model is derived; (b) a statistical or simulation model is developed to generate predictions, possibly entailing scaling functions when extrapolating to broad scales; and (c) the results are evaluated against new data. In this article, we examine the application of this framework in a variety of contexts, using examples from the scientific literature. We conclude by discussing the challenges, limitations, and future prospects for extrapolation.*

*Keywords:* extrapolation, prediction, scale, modeling, evaluation

**S**ome of society's most pressing concerns are ecological, and ecologists are increasingly called upon to explain broadscale problems and contribute to their solutions. But whereas phenomena such as global warming, pollution, biodiversity loss, and land-use change operate over very large areas or over extended periods of time, the field data that characterize ecological research are typically collected over relatively small areas during studies of short duration. Reconciling this mismatch in scales is one of the most formidable challenges confronting environmental scientists (Levin 1992, Peterson and Parker 1998), which may explain why references to scale in the research literature have increased exponentially in recent years (Schneider 2001). Given the logistical, financial, and technical constraints on data collection at broad scales, meeting this challenge depends largely on scientists' ability to make reliable predictions using the data at hand.

When prediction is grounded in current knowledge, it is more precisely termed *extrapolation*. To extrapolate is "to project, extend, or expand (known data or experiences) into an area not known or experienced so as to arrive at...knowledge of the unknown area by inferences based on an assumed continuity, correspondence, or other parallelism between it and what is known" (Gove and Merriam-Webster 1986). This definition encompasses the process of "scaling up," or deriving inferences and rules that can be applied at broad scales on the basis of data collected at smaller scales. It also includes the extension of an ecological relationship from one location to another at approximately the same spatial scale (Turner et al. 1989a). The latter type of extrapolation may be

outside the original extent (i.e., the range over which observations are made) or within that extent, as in interpolation, or "filling in" a series.

Extrapolation in one form or another has always been a part of ecology, but it became a *sine qua non* in the latter half of the 20th century. This reflected a general paradigm shift in the philosophy of science (Popper 1959) and the subsequent efforts of ecologists such as Robert MacArthur to transform their discipline into more of a predictive science (Cody and Diamond 1975). On the heels of this shift, there were expectations within a burgeoning environmental movement that ecologists would provide the scientific knowledge necessary for public policy formation (McIntosh 1985). Technological innovations over the last few decades, especially in the fields of remote sensing and geographic information systems (GIS), greatly enhanced scientists' capacity to meet this challenge by giving them the ability to describe patterns in nature over broader spatial scales and at a greater level of detail than

---

James R. Miller (e-mail: jrmiller@iastate.edu) is an assistant professor in the Department of Natural Resource Ecology and Management and the Department of Landscape Architecture at Iowa State University, Ames, IA 50011. Monica G. Turner is a professor, and Erica A. H. Smithwick is a postdoctoral fellow, in the Department of Zoology at the University of Wisconsin, Madison, WI 53706. C. Lisa Dent was a postdoctoral fellow, and Emily H. Stanley is an assistant professor, at the Center for Limnology, University of Wisconsin. © 2004 American Institute of Biological Sciences.

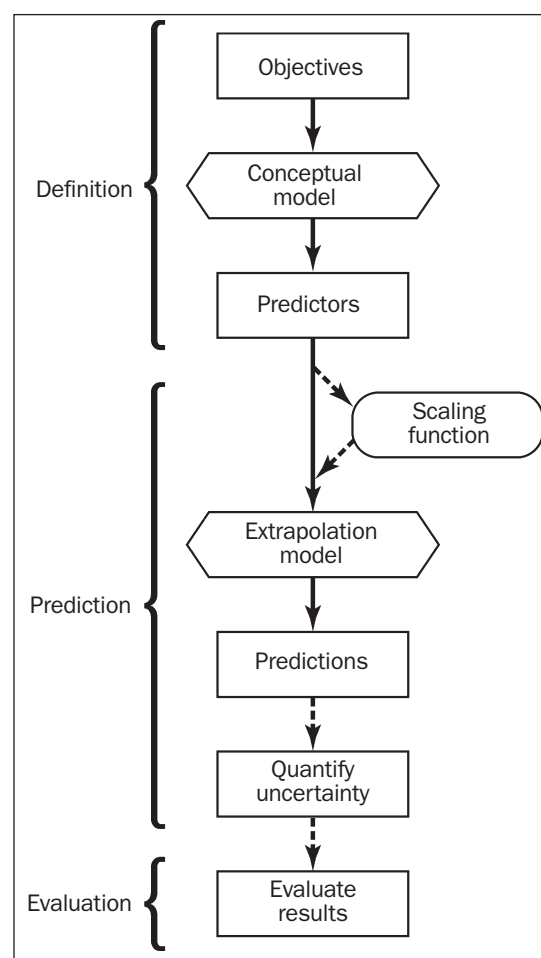
ever before. Concurrent with such developments, however, has been a growing appreciation for the complexity and uncertainty involved in determining which patterns are ecologically meaningful and in predicting their environmental consequences.

Spatial extrapolation in ecology tends to comprise variations on a basic framework (figure 1). We explore that framework in this article, using a variety of process- and organism-based examples, with the goal of drawing on the lessons learned from these examples to inform future research. Throughout, we emphasize work on terrestrial and freshwater systems, not to downplay endeavors in other areas (e.g., marine systems), but rather because these are the fields of study with which we are most familiar. We first discuss how a given extrapolation is initially defined, especially the factors that lead to the inclusion of particular variables in predictive models and the determination of the scales over which these variables are measured. Second, we describe the process of generating predictions. We consider different types of predictive models, as well as the role of scaling functions and the ways in which ecological relationships are affected by changes in scale. Third, we explore techniques for dealing with uncertainty in predictions and discuss the importance of evaluating the accuracy of model predictions, potential sources of error, and procedures for reporting error. We conclude by summarizing both the limitations and the potential value of extrapolation.

### Defining the approach

Typically, the first step in extrapolation is a statement of objectives that, by definition, extend ecological relationships identified in previous studies (the source) to new locations or over broader scales (the target). Prior understanding is summarized in a conceptual model, based either on descriptive data from the literature or on statistical relationships between response variables (the pattern or process that is being predicted) and predictors (Guisan and Zimmermann 2000). At this point, careful attention should be given to the grain (size of the individual units of observation) and extent that characterize the response and predictors in both source and target areas (Turner et al. 1989b).

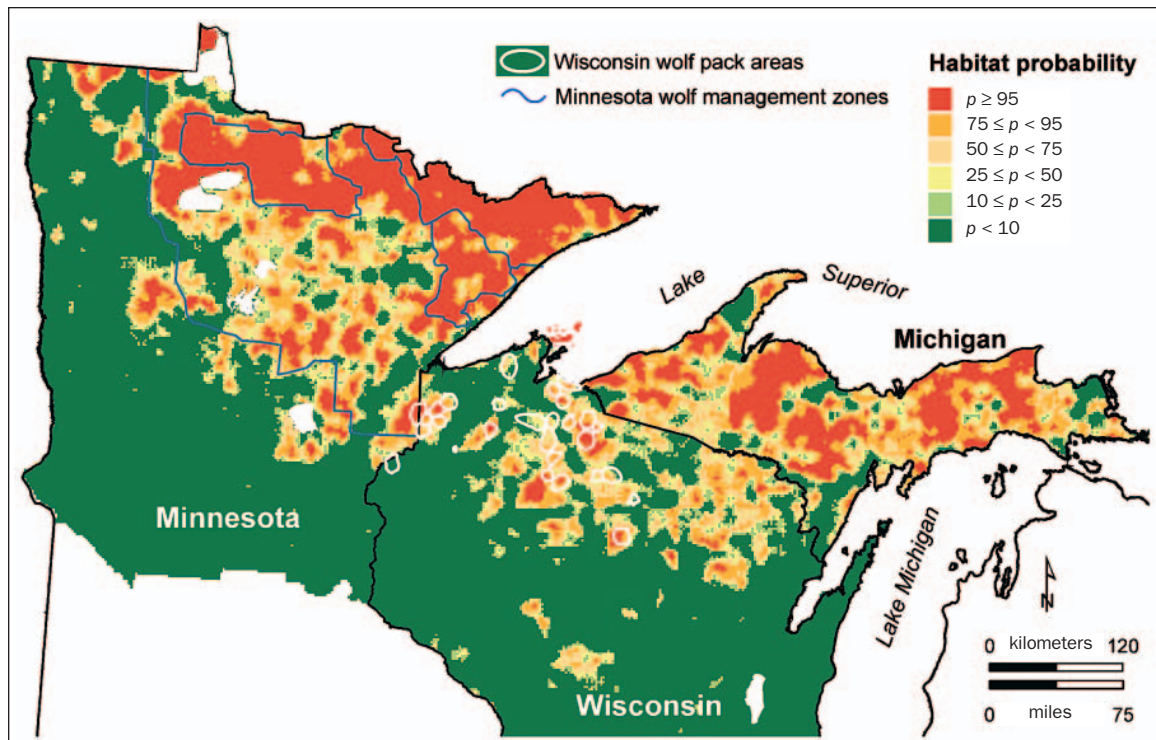
Perhaps the best way to ensure that the grain and extent of both source and target areas are compatible with the goals of a project is to develop new data sets with this purpose in mind. However, the expense of developing data sets that include remotely sensed imagery and field measurements collected over broad areas is often prohibitive within the confines of a single study, prompting many investigators to rely on existing data for the source area, the target area, or both. Mladenoff and colleagues (1995), for example, examined the usefulness of a suite of available data sets in predicting the territory locations of an endangered species, the gray wolf (*Canis lupus*), in northern Wisconsin. These data included information on land cover (from the US Geological Survey [USGS] Land Use and Land Cover data files), deer population density (from the Wisconsin and Michigan Depart-



**Figure 1.** A basic framework describing the various components of extrapolation in ecology. Solid arrows indicate steps that are typically included in this process, whereas dotted arrows indicate steps that are less commonly addressed.

ments of Natural Resources), road density (from the US Census Bureau's TIGER/line [Topologically Integrated Geographic Encoding and Referencing] files), land ownership (from the respective states), and human density (from census blocks). The model with the greatest predictive power (measured using known locations of wolf territories) included a single term for road density, most likely reflecting the probability of human contact (figure 2; Mladenoff et al. 1995).

A limitation of relying on existing data is that the minimum grain size and maximum extent are preset. Preexisting data can be aggregated, however, and this is one way that the relationship between predictors and response variables has been explored across a range of grain sizes. Karl and colleagues (2000) sequentially aggregated data describing vegetation cover and topography in 0.09-hectare (ha) map cells to produce two additional resolutions (4 ha and 10 ha). They examined the effect of this variation on the accuracy of pre-



**Figure 2.** Probability of favorable wolf habitat for Minnesota, northern Wisconsin, and upper Michigan, based on a logistic model using road density as the predictor variable. Modified from Mladenoff and colleagues (1995).

dicted occurrences of breeding bird species in Idaho at two scales of analysis: the site level (homogenous areas of less than 0.5 ha) and the cover-type level (aggregations of many sites of a similar cover type). Models with finer grain size performed better in more heterogeneous areas and at the cover-type (as opposed to site) level. The latter finding, which may be partly a function of the number of individuals of a given species necessary to test habitat-relationship models, suggests that such models may be better suited to coarser scales (Karl et al. 2000).

As an alternative to aggregating data from a single source to explore the effect of grain size, some investigators have substituted different data sets to describe a given variable at different resolutions. Iverson and colleagues (1997) evaluated digital elevation model (DEM) data from four different sources, each with a different resolution, for their effectiveness in estimating an integrated soil moisture index for a managed forest in southern Ohio. These sources were a USGS digital line graph (1:100,000 scale) and DEMs derived from a 7.5-minute USGS digitized contour map (1:24,000 scale), USGS data (1:24,000 scale), and a USGS 3-arc-second DEM (1:250,000 scale). Relative to topographic and moisture indexes, the 1:24,000-scale digitized contour map data and the other 1:24,000-scale data performed reasonably well, but the reliability of the 1:100,000-scale data was ambiguous, and the 1:250,000-scale data were unreliable. Iverson and colleagues attributed these results to the relatively small area of the forest (475 ha) and the amount of topographic relief (less than 100 meters [m] total).

These examples reflect the exploratory approach that has often been taken in identifying suitable predictors and associated grain sizes. Some workers have extended this approach to examine the effects of varying the spatial extent of predictors. For instance, Mitchell and colleagues (2001) compared three models for predicting the presence of forest bird species in South Carolina in an effort to provide managers with a method to assess the effects of forest management over large areas. The models included one based only on microhabitat features measured over 50-m plots, one based only on landscape characteristics (GIS data depicting forest type and age), and one based on both data types. Mitchell and colleagues found that the three model types generally had the same explanatory power, and that landscape models performed particularly well for migrant species that were habitat specialists. If done carefully, inductive approaches such as these may yield important insights as to the appropriate scales for measuring predictor variables and the circumstances in which extrapolation is likely to be effective.

### Generating predictions

Ideally, the choice of a given extrapolation model is based on research objectives and on the nature of the relationship between response and predictors. Model selection, however, often appears to be based more on the traditions in a given discipline than on careful consideration of the alternatives. Among static distribution models (i.e., those that relate geographical distributions to current environmental condi-

tions), regression has traditionally been the tool of choice. Usually logistic regression is applied to binary or categorical data (e.g., the presence or absence of one or more species, the occurrence of a disturbance event, discrete nutrient levels) and linear regression to continuous data (e.g., species abundance, nutrient concentrations). These two methods are occasionally combined in a two-stage approach, first modeling the presence or absence of an organism and then, conditional on the organism's presence, modeling its abundance. Despite their continuing popularity, linear models are associated with assumptions that are difficult to meet with many ecological data sets, especially regarding the statistical distribution of response variables, the form of variance structures, and the independence of observations. Linear models also tend to underestimate the slope of a regression line if there is unmeasured variability associated with independent variables.

A number of promising alternatives that do not impose such limitations are gaining in popularity. Classification and regression trees offer a nonparametric alternative to linear regression models and are ideal for exploring and modeling strongly nonlinear data with complex interactions among variables (De'ath and Fabricius 2000). Generalized linear models (GLMs), an extension of linear models, are suitable for data from a variety of probability distributions (normal, binomial, Poisson, negative binomial, or gamma). Generalized additive models, or GAMs, are semiparametric extensions of GLMs; they are applicable when relationships between response and predictors are highly nonlinear and nonmonotonic. Guisan and colleagues (2002) provide a comprehensive overview of these latter two methods, including a number of examples.

The assumption of independence among observations is frequently violated because of spatial dependencies in the data, resulting in inferior models and inaccurate predictions (Carroll and Pearson 2000). A residual plot is sometimes sufficient for detecting spatial patterning, but if the data are indeed dependent, modeling techniques such as autoregressive or geostatistical procedures may be required. Spatial statistics are increasingly being used in the context of extrapolation, and they have great potential to improve the accuracy of predictive models (see Fortin [1999] for an overview of these approaches). Kriging, which may be the most commonly used method of this sort, relies on autocorrelation functions to generate spatially explicit predictions (Webster and Oliver 2001). One application of this method is the creation of DEMs by extrapolating from topographic data at known locations.

Techniques for evaluating competing models of the same form may be useful in this context (Burnham and Anderson 1998). There may also be value in the application of different model types to the same data set, potentially providing insights as to the relative strengths and weaknesses of various modeling techniques in a given context. For example, four different methods were compared in predicting forest composition in North Carolina's Coweeta basin (Bolstad et al. 1998). Kriging was used to extrapolate forest composition from measurements of basal area and stem density collected on a series of small plots (0.08 ha) to the entire basin. The effec-

tiveness of co-kriging, which involves the use of covariates, was also tested to see whether including elevation or terrain shape as covariates with the plot data improved the accuracy of the predictions. In addition, vegetation maps were produced by linear regressions involving elevation and terrain variables combined with cartographic overlay, and also by a mosaic diagram, which is sometimes used to summarize the relationship between elevation, landform, and expected vegetation. When predicted vegetation patterns were compared with known forest composition in a set of independent plots, the mosaic diagram and linear regression models were more accurate than either the kriging or the co-kriging techniques. Bolstad and colleagues (1998) concluded that geostatistical methods were not useful for mapping forest composition in the southern Appalachians, because spatial covariation decreases rapidly with distance and would therefore require a very dense array of sampling plots.

Whereas static distribution models by definition assume equilibrium and a fixed environmental template, nonequilibrium conditions in a dynamically and stochastically changing environment are addressed with simulation models (Guisan and Zimmermann 2000). In simulation models, the potential mechanisms underlying the observed response are represented formally. A dynamic simulation approach is well suited to extrapolating patterns or processes over broad scales, particularly when the pattern of the driving variables may change. Simulation approaches to extrapolation are widely used in ecosystem ecology, because field measurements of process rates across large areas are costly to acquire, and thus relatively few spatially extensive data sets exist. In these approaches, the attributes of individual grid cells serve as inputs, but a simulation (as opposed to statistical) model is used to project the value of the response variable. In addition, interactions among different sites may be represented in a simulation approach. Running and colleagues (1989) were among the first to integrate biophysical information obtained from many sources and combine these data with an ecosystem simulation model to predict spatial patterns of evapotranspiration and net photosynthesis across a large landscape. The results demonstrated the power of these new integrative methods for producing spatially explicit projections of variation in ecosystem processes and offered insights into interactions among the controls on these processes.

As with statistical models, different simulation models have been applied to common data sets. For example, three biogeography models and three biogeochemistry models were compared under existing atmospheric carbon dioxide ( $\text{CO}_2$ ) levels and climatic conditions, and under doubled  $\text{CO}_2$  levels and a range of potential climate scenarios (VEMAP 1995). The biogeography models in the study were used to predict the geographic distribution of major vegetation types, and the biogeochemistry models simulated cycles of carbon, nutrients, and water in terrestrial systems. Numerous models of each type have been developed independently in recent years and exercised over large areas, or over the entire globe, using a variety of climate-change scenarios. Because understand-



ing the controls of ecosystem structure and function is not sufficient to identify the best models or to judge their predictions as correct, a comparative approach involving a common data set is a reasonable way to gauge model differences, with the ultimate goal of providing more realistic simulations (VEMAP 1995).

**Scaling functions.** Because processes, patterns, and organism responses are scale dependent, a procedure for dealing with issues of scale is key when extrapolating from small plots to larger areas. The most commonly used procedure is direct scaling (King 1991), which assumes that the relationship of a variable to changes in scale is linear or additive. The quantity of interest is thus multiplied by the proportionate increase in unit area. When variability in this quantity is associated with, say, different forms of land cover, an overall estimate is obtained by repeating this process for each cover type that is present and either summing or averaging over the extent of the study (see King [1991] for variations on this approach).

Direct scaling, though simple, may be an appropriate choice in some instances. Extrapolation of ecosystem process rates often relies on area-weighted averaging, with the assumption that landscape elements do not interact horizontally. In the case of carbon flux, for example, vertical exchanges from the atmosphere to the biosphere through photosynthesis, or from the biosphere to the atmosphere through respiration, are the only ones considered. (See Houghton and colleagues [2000] for an example in which annual carbon fluxes stemming from deforestation and agricultural abandonment in the Brazilian Amazon were measured using direct scaling.)

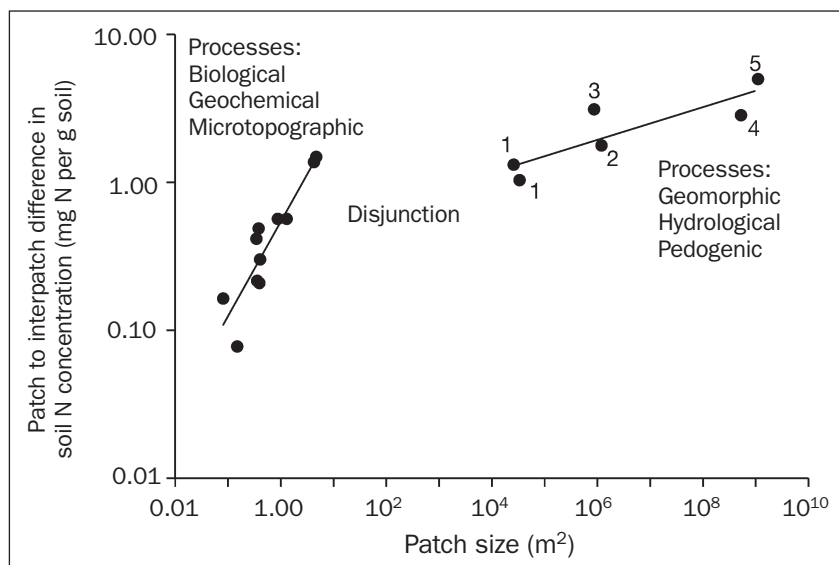
Even when horizontal interactions are an important consideration, direct scaling may still be quite effective, at least

over very large areas. Caraco and Cole (1999) examined nitrate export in 35 large river systems with a worldwide distribution and found that a simple model based on human point-source and nonpoint-source nitrogen loads explained much of the variation ( $r^2 > 0.8$ ) among watersheds. For each river, point-source inputs were derived from per capita sewage production and urban population estimates, and nonpoint-source inputs were calculated as the product of nitrogen fertilizer per unit of agricultural land and the total amount of agricultural land in the watershed. Conversely, Poiani and colleagues (1996) reported that such an approach was inadequate for describing nitrogen export to wetlands in nine relatively small watersheds in New York State. They found that the spatial characteristics of these watersheds and the amount of cropland were strong determinants of nitrogen delivery to groundwater-dominated wetlands at this scale. Nitrogen loads were attenuated as a function of slope, soil porosity, and flow path length.

How do researchers decide when direct scaling is adequate or when a different method is necessary? Ludwig and colleagues (2000) described a general approach to dealing with issues of scale that is rooted in scaling functions. These functions provide the conceptual framework for defining the collective scaling dynamics of a system and the basis for proposing scaling rules that relate changes in scale to consequences for particular phenomena in a particular place. From these rules, one can derive the scaling equations necessary for generating predictions (Ludwig et al. 2000). To illustrate this approach, Ludwig and colleagues (2000) proposed that in savanna ecosystems, the amount of a resource per unit area (in this case, soil nitrogen) in vegetation patches increases with the size of the patch. This scaling rule was based on functions related to

surface water flow, to the redistribution of nutrients and organic matter, and to the ways in which patches capture these materials. Data from savanna landscapes in northern Australia were used to test the scaling rule and then to develop a scaling equation for predicting the conservation of soil nutrients under different landscape disturbances. Although the scaling rule was supported over a wide range of patch sizes, there was an apparent disjunction in scaling relations between small patches and large landscape patches (figure 3), a result that Ludwig and colleagues (2000) attributed to different processes operating at the two scales. In other words, the rule applied generally to a wide range of landscapes, but the scaling equation necessary for extrapolation had a much narrower range of applicability.

The importance of this last point cannot be overstated. When scaling relationships are nonlinear but still monotonic, extrapolation may be possible through the derivation of power equations. In fact, relationships of this



**Figure 3.** Apparent threshold response of interpatch differences in soil nitrogen to variation in patch size. This disjunction is probably related to different landscape processes acting at different scales. Modified from Ludwig and colleagues (2000). Abbreviations: g, gram; m<sup>2</sup>, square meters; mg, milligram; N, nitrogen.

sort are common in nature, with examples ranging from body-size allometry to species–area curves (Schneider 2001). Extrapolation is ill-advised, however, across domains of scale that are delineated by critical thresholds where abrupt or nonlinear changes occur (O'Neill et al. 1989, Wiens 1989). Such thresholds, exemplified by the disjunction apparent in figure 3, present a particularly vexing challenge in extrapolation because they are often difficult or impossible to anticipate in the absence of adequate empirical data. Ludwig and colleagues (2000) noted that their scaling rule still held for the larger, landscape patches (although the scaling equation did not), but this is not always the case. Andrén (1994) found that, for a variety of animal species, the relationship between habitat suitability and fragmentation in the surrounding landscape exhibits a threshold when the loss of habitat exceeds 70%. Below that threshold, the overall amount of habitat is the primary determinant of population size, whereas once that threshold of loss is reached, the arrangement of habitat remnants becomes crucial.

Extrapolation based on scaling rules or equations that are inappropriate when crossing critical scaling thresholds results in aggregation error (O'Neill 1979), so called because it arises from the variation among aggregated components (see O'Neill and King [1998] for examples that describe this sort of error). Extrapolation procedures that minimize aggregation error have been proposed (e.g., Rastetter et al. 1992), but these have generally not received much attention from ecologists (O'Neill and King 1998).

**Measuring uncertainty in predictions.** There is always a measure of uncertainty associated with extrapolation. As Stewart (2000) has pointed out, a prediction based on current knowledge represents just one of a number of possibilities. He goes on to distinguish between aleatory uncertainty, stemming from random processes (e.g., the roll of a fair die), and epistemic uncertainty, a function of incomplete knowledge of the factors that determine events. Total uncertainty is the sum of these two forms. The important question is not how to eliminate these sources of doubt (an impossible task), but rather how to quantify uncertainty and then incorporate this information into model predictions (Flather et al. 1997).

Quantification of uncertainty in predicted process rates or species distributions has not received much empirical attention, but a growing number of examples provide some guidance. Here we emphasize spatially explicit depictions of uncertainty, because they are particularly valuable if model predictions are to be used in formulating policy or management decisions. For instance, Mladenoff and colleagues (1995) devised an effective, spatially explicit representation of uncertainty in which the probabilities that are the products of logistic regression are treated categorically for display purposes (figure 2). Another example is provided by Pidgeon and colleagues (2003), who used area-weighted averaging to extrapolate avian nest success and abundance from data collected over 42 plots (600 m<sup>2</sup> each) to an entire landscape in central New Mexico. The resulting map (figure 4) depicted not only

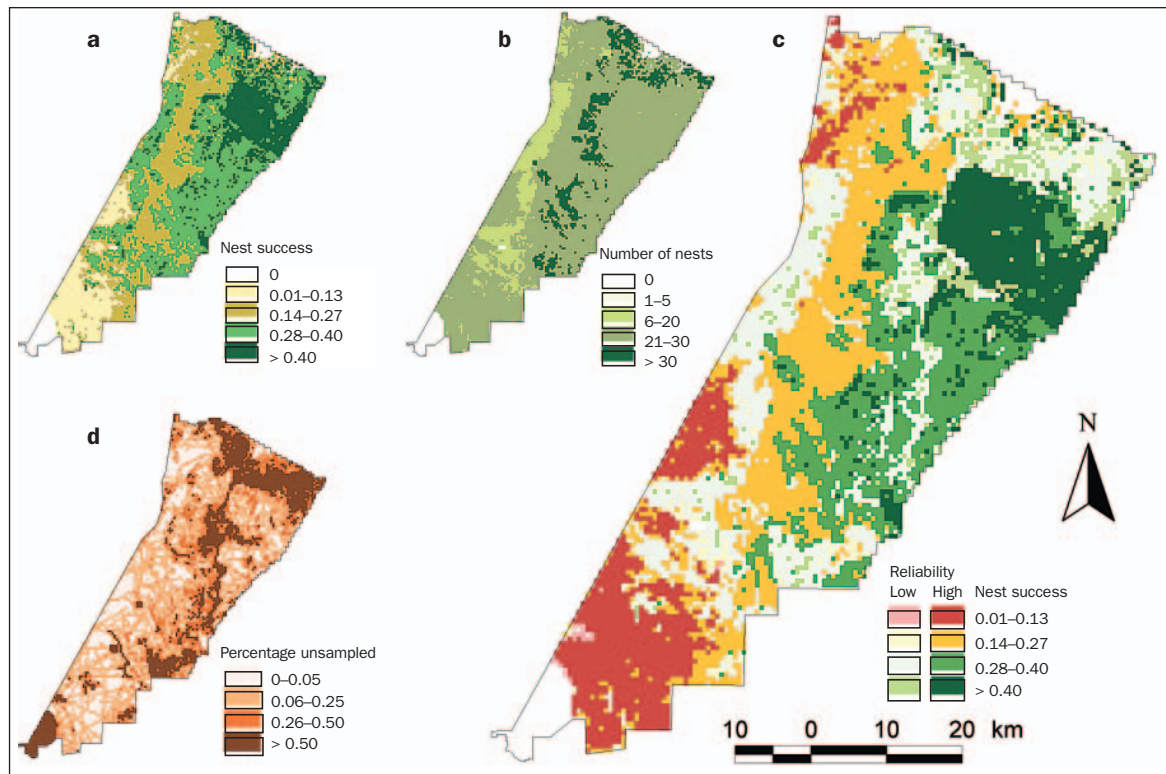
estimates of breeding productivity but also quantification of uncertainty, expressed as a binomial level of confidence (high or low) in these estimates (based on the number of nests used in the calculations and the habitat types present in each 600-m<sup>2</sup> cell).

In a similar vein, Hansen and colleagues (2000) used multiple regression to extrapolate aboveground net primary productivity (ANPP) from a series of small plots to a large portion of the Greater Yellowstone ecosystem, using a suite of abiotic and biotic predictor variables. Two maps were produced: one depicting the predicted mean ANPP for each 30-m<sup>2</sup> cell, grouped into four classes (figure 5a), and the other showing the coefficients of variation for these predictions (figure 5b). A spatially explicit display of the uncertainty and variation associated with predictions is useful in pinpointing locations that require greater sampling intensity or in identifying the need for additional predictors in the extrapolation model. Such displays may also be valuable to decisionmakers by identifying locations in the landscape where confidence in the model results is high. Failure to clearly articulate uncertainty may result in poor decisions and undermine future contributions of scientific research to policy formation (Pielke et al. 2000).

## Evaluating results

Once predictions have been generated, the logical next step is an assessment of their accuracy. This step, referred to as model evaluation (Oreskes et al. 1994, Guisan and Zimmermann 2000), assesses the correspondence between what is predicted and what is subsequently observed. In some cases, evaluation with field data may not be meaningful or even possible. For example, some of the difficulties with testing simulation model results against empirical data over broad scales were noted by Kucharik and colleagues (2000). They evaluated a dynamic global ecosystem model (DGEM) by comparing biome-specific predictions with global-scale observations of water balance, carbon balance, and vegetation structure. Simulated patterns were in reasonable agreement with field estimates, but the authors advised that comparisons of DGEM output with empirical data should be interpreted cautiously, for two reasons. First, model results were estimates of pools or processes of large areas (1° latitude × 1° longitude) that were assumed to be homogenous, whereas empirical data were collected in plots as small as 10 m<sup>2</sup>. Second, the empirical evidence available for model evaluation was surprisingly scarce and was poorly distributed over the spatial and temporal scales relevant to continental- or global-scale change.

Extrapolations of this sort are often made because the measurement of broadscale patterns or processes is intractable. Rastetter (1996) asserted that long-term climate change models are not amenable to testing and that this is unlikely to change in the foreseeable future. Nonetheless, he concluded that such models remain an essential part of efforts to determine the global consequences of human activities; untested predictions, based on the best science available, are



**Figure 4.** Map layers used in predicting nest success for the black-throated sparrow in central New Mexico: (a) nest success estimates in seven habitat types, (b) number of nests used to estimate nest success in each habitat type, (c) estimates of mean nest success and confidence levels of estimates, and (d) location of sampled and unsampled habitat types. Modified from Pidgeon and colleagues (2003).

still better than proceeding blindly. Comparisons of model output with field data, though not as informative as rigorous testing, are useful in assessing the relative contribution of various processes to climate change and in testing the consistency of interpretations of empirical findings (Rastetter 1996). Moreover, broadscale extrapolation models may also have value in identifying data needs and knowledge gaps and in describing the potential consequences of alternative management actions (He and Mladenoff 1999).

An admission that evaluation is not always possible, however, should not be construed to mean that it is unnecessary. In many cases, the products of extrapolation are amenable to testing, and there is much to be gained by doing so. The omission of this final step in the extrapolation process has resulted in a proliferation of models of questionable value, heuristic considerations notwithstanding.

Evaluation is only meaningful when it is based on data that were not used in formulating the extrapolation model, and there are two ways to accomplish this. The first involves the use of two independent data sets, one for calibrating the model and one for testing it. The second, sometimes referred to as the training-testing method, is a variation of the first: A single data set is split, with half the data used in model development and the second half withheld for evaluation. Either method allows for an independent assessment of the extrapolation, usually followed by deliberation over any

sources of error that may have reduced the accuracy of model predictions.

**Sources of error.** Error, or a lack of correspondence between predictions and new observations, can arise from many sources, some of which have been described in the preceding sections. Thapa and Bossler (1992) describe a variety of errors associated with data collection, which may have a multiplicative effect as the number of data sources increases or as information is aggregated at larger scales. It may be possible to correct for systematic errors with a functional relationship, but this is not possible when errors are random (Thapa and Bossler 1992).

Errors in model predictions may also derive from limitations on the types of data that can be collected over broad scales. This point is illustrated by Orrock and colleagues (2000) in their study of the southern red-backed vole (*Clethrionomys gapperi*), a species requiring habitat features that are not easily identified with current remote-sensing technology. Low-resolution imagery was adequate for identifying forest types where suitable habitat might be found, but higher-resolution data gathered from field surveys were necessary to predict vole presence and abundance accurately. By contrast, Mladenoff and colleagues (1999) experienced a high success rate using data on road density (which are highly accurate and widely available in digital form) to predict locations for wolf



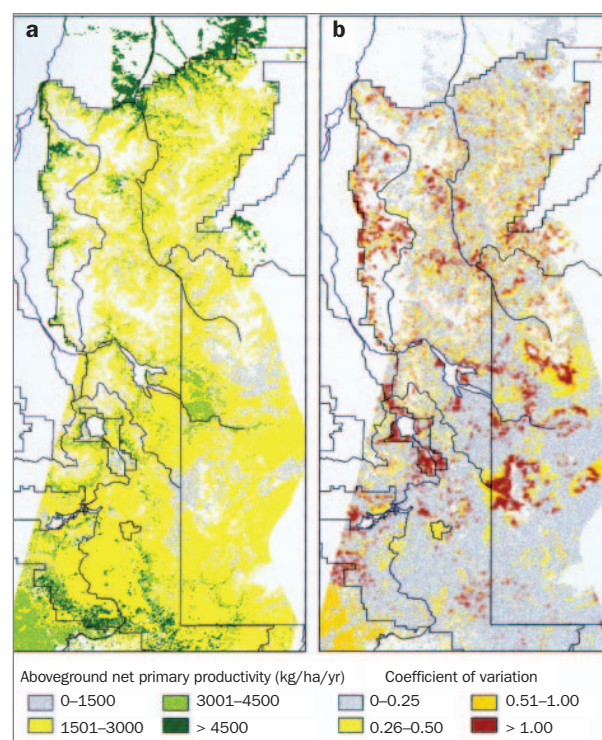
packs that were colonizing new areas in the upper Midwest; 18 of 23 packs were established in areas that were classified as favorable. The gray wolf, a habitat generalist, has a long history of persecution by humans, and road density serves as a reliable index of the probability of human contact.

Error may also emanate from difficulties in surveying certain species or from model parameterization based on limited data. Edwards and colleagues (1996) found that the error rates for the predicted occurrence of amphibians and reptiles in Utah's national parks were higher than the rates for birds and mammals. They attributed these results partly to problems in inventorying the herpetofauna and partly to a greater historical emphasis on avian and mammalian natural history.

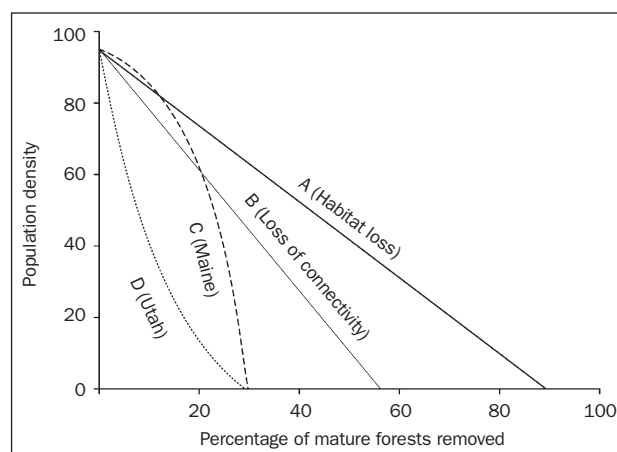
In many cases, errors may stem from extrapolation models that are based on correlative habitat relationships whose underlying mechanisms are poorly understood. As a case in point, the American marten (*Martes americana*) is generally characterized as having a strong affinity for mature, closed-canopy coniferous forests (Bissonette et al. 1997). Yet a predictive distribution model based on this apparent affinity would be overly restrictive in some regions: In Maine, for example, martens also use deciduous forests and regenerating stands. The most likely reason is that the attributes of forest structure required by this species, which are found only in mature coniferous stands throughout much of its range, are found in a wide variety of forest types in Maine (Bissonette et al. 1997).

Mechanistic explanations such as this may account for habitat selection over larger areas, from microhabitat features to selection at the stand scale. Nonetheless, the work of Bissonette and colleagues (1997) underscores the possibility that such explanations may not hold as the extent of an investigation is increased. They show that American martens are apparently sensitive to broadscale landscape patterns, even though the mechanisms affecting habitat selection operate at or below the scale of the home range. Population declines in this species, which avoids large unforested areas, deviated from predictions of a linear decrease based on loss of habitat or connectivity; instead, they exhibited a nonlinear response in both Utah and Maine (figure 6; Bissonette et al. 1997). Furthermore, the response curve for martens in Utah initially declines more sharply than the curve for Maine, suggesting that the Utah populations may be more sensitive to lower levels of fragmentation. When mature forest represents less than 70% of the landscape, the curves for both states converge to indicate a lack of habitat suitability (figure 6), adding further support for the fragmentation threshold identified by Andr  n (1994). The description of habitat relationships for the American marten by Bissonette and colleagues (1997) highlights the potential for error when extrapolating from one system to another (i.e., from Utah to Maine) and also for aggregation error (O'Neill 1979) when extrapolating relationships from stands to landscapes (or between landscapes with different levels of forest fragmentation).

In many cases, extrapolation errors arise from the failure to consider effects associated with the nature of the landscape



**Figure 5.** (a) Predicted distribution of aboveground net primary productivity (ANPP; kilograms per hectare per year) for a portion of the Greater Yellowstone ecosystem, based on cover type and elevation, and (b) estimates of the coefficient of variation in predicted ANPP. Modified from Hansen and colleagues (2000).



**Figure 6.** Responses of pine marten to habitat fragmentation. Curve A represents the expected response to increasing fragmentation if martens are influenced only by habitat loss. Curve B represents the expected relationship if martens are also influenced by loss of connectivity. Curves C and D represent actual marten responses to habitat fragmentation levels in Maine and Utah. Modified from Bissonette and colleagues (1997).



mosaic. Reiners and Driese (2001) point out that in most predictive models, the characteristics of individual grid cells, or clusters of like cells (i.e., habitat patches), tend to be considered independently. There is a tacit assumption in these models that the presence or abundance of a species (or the occurrence or rate of a process) in a given location is invariant with respect to landscape position. Yet vegetation composition and structure in a given patch may be strongly dependent on the surrounding landscape in terms of seed sources and the propagation of disturbances such as fire or windthrow. It has also long been recognized that many animal species require a variety of habitat types for daily activities, such as resting and feeding, and may depend on different habitats in different seasons. Moreover, numerous studies

have demonstrated that for some vertebrate and invertebrate species, the suitability of a habitat patch or mosaic of patches is affected by the surrounding landscape in ways that may not be manifest in local habitat structure, especially in human-altered areas (Mazerolle and Villard 1999). The failure to incorporate considerations of landscape position in extrapolation models, even though the importance of these considerations is widely appreciated, most likely results from the complexity their inclusion would introduce and from a poor understanding of underlying mechanisms.

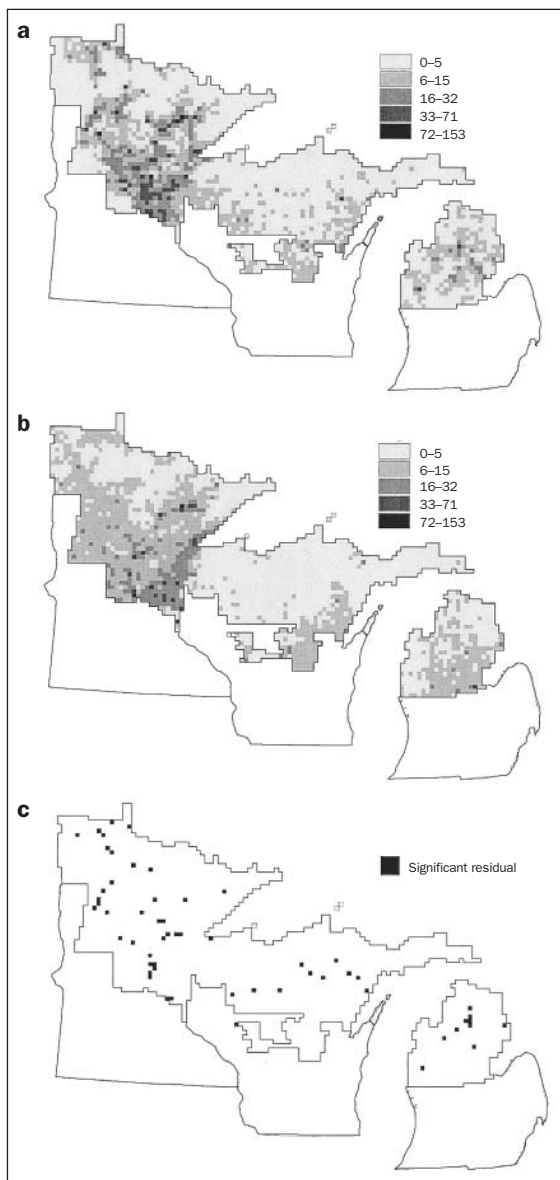
**Reporting error.** When extrapolation is evaluated, accuracy is often reported as a correlation coefficient for a continuous response and as the percentage of correct or incorrect predictions for a categorical response. For a binary response, such as presence or absence, incorrect predictions may be further classified as errors of commission (false positive) and omission (false negative). Errors of commission are especially difficult to interpret for mobile organisms, because one cannot be sure whether the species was detected as the result of model inaccuracies or sampling error (Haila et al. 1993).

Identifying accurate predictions using logistic regression requires an essentially arbitrary choice of a threshold probability that separates correct from incorrect observations. This threshold probability is often set at 0.5 (e.g., Mladenoff et al. 1995, 1999) but could be set higher or lower, depending on the relative importance of false negatives and false positives (Stewart 2000). An alternative technique measures discrimination capacity as the area under a relative operating curve that tracks the proportion of correct and incorrect predictions over a wide range of thresholds (Pearce and Ferrier 2000). If the area under the curve is 0.95, for example, this indicates that the model under examination can discriminate between occupied and unoccupied sites 95% of the time.

Spatially explicit depictions of model output, like spatially explicit depictions of uncertainty, may have advantages over tabular results in suggesting ways that future efforts might be allocated and in pinpointing locations where confidence in results is high. In one example, Cardille and colleagues (2001) developed GLMs to determine which abiotic, biotic, and human variables best explained fire activity between 1985 and 1995 in the upper Midwest. They evaluated their model using the training-testing method and depicted the results using maps of predicted and observed fire counts (figure 7).

## Conclusions

Extrapolation has become a major research focus in applied ecology (e.g., Scott et al. 2002), and despite the wide assortment of methods being applied in a variety of contexts, several common patterns have begun to emerge. In the most reliable extrapolations, response variables tend to be closely associated with environmental features that can be accurately described using remote sensing technology. Given a strong conceptual model, the choice of response and predictors is still constrained by the availability of data for the target system. This limitation may be alleviated to some extent



**Figure 7.** (a) Observed fire counts, (b) predicted fire counts, and (c) residuals of fire-count model for forest regions in the upper Midwest. Modified from Cardille and colleagues (2001).

by technological advances, such as the development of new airborne scanners (Lefsky et al. 2002), that permit the detection of a wider array of environmental features at ever-finer resolutions.

A second limitation on the ability to generate accurate predictions is a poor understanding of the mechanisms and feedbacks that underlie many ecological patterns. Correlative relationships may be adequate for extrapolation over a narrow range of spatial and temporal scales, but generally the most accurate extrapolations are based on relatively simple relationships grounded in mechanisms that are well understood. Controlled experiments are often quite useful in identifying such mechanisms, but they are typically conducted only over limited extents (Kemp et al. 2001) and may thus fail to identify spatial contingencies or multiple causes. This situation is likely to persist, given the difficulties of acquiring adequate sample sizes and achieving sufficient replication at broad scales (Hargrove and Pickering 1992).

Scaling functions may provide the link between fine-scale experiments and broadscale applications to some degree, but the existing data in most cases are inadequate to develop these functions (Ludwig et al. 2000). When it is possible to derive a scaling rule, the domains of scale that define its range of applicability are difficult to identify a priori. These limitations suggest that multiple approaches, including experiments to unravel mechanisms as well as inductive methods, are necessary to achieve a better understanding of scaling issues (Wiens 1995). Indeed, induction may be the only way to identify the critical thresholds, the scales at which different organisms and processes respond to their environments, and the ways that these responses vary geographically.

In effect, a conceptual model represents a testable hypothesis, and extrapolation is a means to assess the robustness of underlying relationships. There are numerous opportunities for learning throughout this process. These include the application of a single model to data describing patterns at different spatial scales and the comparison of several models using a common data set. Advances in statistical techniques enhance the ability of researchers to tease apart complex relationships, while increasingly sophisticated remote-sensing and graphical tools permit more accurate descriptions of spatial patterns and suggest directions for future research. Extrapolation is best viewed not as an end point, but rather as part of a cycle involving the application and subsequent revision of what is known. By examining the conditions under which extrapolation fails or succeeds, ecologists are likely to gain a better understanding of ecological patterns and underlying processes.

## Acknowledgments

This paper is dedicated to the memory of Lisa Dent, a valued colleague and friend who is sorely missed. We thank Matthias Bürgi, Jeff Cardille, Mark Dixon, Hojeong Kang, Dan Kashian, and Tania Schoennagel for spirited discussions of the ideas presented here. Comments on earlier drafts from Matt Greenstone, Tania Schoennagel, Jennifer Fraterrigo, and three

anonymous reviewers greatly improved the manuscript. This research was funded by the Environmental Protection Agency STAR (Science to Achieve Results) Program (grant no. R826600).

## References cited

- Andrén A. 1994. Effects of habitat fragmentation on birds and mammals in landscapes with different proportions of suitable habitat: A review. *Oikos* 71: 355–366.
- Bissonette JA, Harrison DJ, Hargis CD, Chapin TG. 1997. The influence of spatial scale and scale-sensitive properties on habitat selection by American marten. Pages 368–385 in Bissonette JA, ed. *Wildlife and Landscape Ecology: Effects of Pattern and Scale*. New York: Springer-Verlag.
- Bolstad PV, Swank W, Vose J. 1998. Predicting southern Appalachian over-story vegetation with digital terrain data. *Landscape Ecology* 13: 271–283.
- Burnham KP, Anderson DR. 1998. *Model Selection and Inference: A Practical Information-Theoretic Approach*. New York: Springer-Verlag.
- Caraco NF, Cole JJ. 1999. Human impact on nitrate export: An analysis using major world rivers. *Ambio* 28: 167–170.
- Cardille JA, Ventura SJ, Turner MG. 2001. Environmental and social factors influencing wildfires in the upper Midwest, United States. *Ecological Applications* 11: 111–127.
- Carroll SS, Pearson DL. 2000. Detecting and modeling spatial and temporal dependence in conservation biology. *Conservation Biology* 14: 1893–1897.
- Cody ML, Diamond JM. 1975. *Ecology and Evolution of Communities*. Cambridge (MA): Harvard University Press, Belknap Press.
- De'ath G, Fabricius KE. 2000. Classification and regression trees: A powerful yet simple technique for ecological data analysis. *Ecology* 81: 3178–3192.
- Edwards TC Jr, Deshler ET, Foster D, Moisen GG. 1996. Adequacy of wildlife habitat relation models for estimating spatial distributions of terrestrial vertebrates. *Conservation Biology* 10: 263–270.
- Flather CH, Wilson KR, Dean DJ, McComb WC. 1997. Identifying gaps in conservation networks: Of indicators and uncertainty in geographic-based analyses. *Ecological Applications* 7: 531–542.
- Fortin MJ. 1999. Spatial statistics in landscape ecology. Pages 253–279 in Klopatek JM, Gardner RH, eds. *Landscape Ecological Analysis: Issues and Applications*. New York: Springer-Verlag.
- Gove PB, Merriam-Webster, eds. 1986. *Webster's Third New International Dictionary of the English Language, Unabridged*. Springfield (MA): Merriam-Webster.
- Guisan A, Zimmermann NE. 2000. Predictive habitat distribution models in ecology. *Ecological Modelling* 135: 147–186.
- Guisan A, Edwards TC Jr, Hastie T. 2002. Generalized linear and generalized additive models in studies of species distributions: Setting the scene. *Ecological Modelling* 157: 89–100.
- Haila Y, Hanski I, Raivio S. 1993. Turnover of breeding birds in small forest fragments: The sampling colonization hypothesis corroborated. *Ecology* 74: 714–725.
- Hansen AJ, Rotella JJ, Kraska MPV, Brown D. 2000. Spatial patterns of primary productivity in the Greater Yellowstone Ecosystem. *Landscape Ecology* 15: 505–522.
- Hargrove WW, Pickering J. 1992. Pseudoreplication: A sine qua non for regional ecology. *Landscape Ecology* 6: 251–258.
- He HS, Mladenoff DJ. 1999. Spatially explicit and stochastic simulation of forest-landscape fire disturbance and succession. *Ecology* 80: 81–99.
- Houghton RA, Skole DL, Nobre CA, Hackler JL, Lawrence KT, Chomentowski WH. 2000. Annual fluxes of carbon from deforestation and regrowth in the Brazilian Amazon. *Nature* 403: 301–304.
- Iverson LR, Dale ME, Scott CT, Prasad A. 1997. A GIS-derived integrated moisture index to predict forest composition and productivity of Ohio forests (U.S.A.). *Landscape Ecology* 12: 331–348.
- Karl JW, Heglund PJ, Garton EO, Scott JM, Wright NM, Hutto RL. 2000. Sensitivity of species habitat-relationship model performance to factors of scale. *Ecological Applications* 10: 1690–1705.

- Kemp WM, Petersen JE, Gardner RH. 2001. Scale-dependence and the problem of extrapolation. Pages 3–57 in Gardner RH, Kemp WM, Kennedy VS, Petersen JE, eds. *Scaling Relations in Experimental Ecology*. New York: Columbia University Press.
- King AW. 1991. Translating models across scales in the landscape. Pages 479–517 in Turner MG, Gardner RH, eds. *Quantitative Methods in Landscape Ecology*. New York: Springer-Verlag.
- Kucharik CJ, Foley JA, Delire C, Fisher VA, Coe MT, Lenters JD, Young-Molling C, Ramankutty N. 2000. Testing the performance of a dynamic global ecosystem model: Water balance, carbon balance, and vegetation structure. *Global Biogeochemical Cycles* 14: 795–825.
- Lefsky MA, Cohen WB, Parker GG, Harding DJ. 2002. Lidar remote sensing for ecosystem studies. *BioScience* 52: 19–30.
- Levin SA. 1992. The problem of pattern and scale in ecology. *Ecology* 73: 1943–1967.
- Ludwig JA, Wiens JA, Tongway DJ. 2000. A scaling rule for landscape patches and how it applies to conserving soil resources in savannas. *Ecosystems* 3: 84–97.
- Mazerolle MJ, Villard M-A. 1999. Patch characteristics and landscape context as predictors of species presence and abundance: A review. *Écoscience* 6: 117–124.
- McIntosh RP. 1985. *The Background of Ecology*. Cambridge (United Kingdom): Cambridge University Press.
- Mitchell MS, Lancia RA, Gerwin JA. 2001. Using landscape-level data to predict the distribution of birds on a managed forest: Effects of scale. *Ecological Applications* 11: 1692–1708.
- Mladenoff DJ, Sickley TA, Haight RG, Wydeven AP. 1995. A regional landscape analysis and prediction of favorable gray wolf habitat in the northern Great Lakes region. *Conservation Biology* 9: 279–294.
- Mladenoff DJ, Sickley TA, Wydeven AP. 1999. Predicting gray wolf landscape recolonization: Logistic regression models vs. new field data. *Ecological Applications* 9: 37–44.
- O'Neill RV. 1979. Transmutation across hierarchical levels. Pages 59–78 in Innis GS, O'Neill RV, eds. *Systems Analysis of Ecosystems*. Fairland (MD): International Co-operative.
- O'Neill RV, King AW. 1998. Homage to St. Michael; or, why are there so many books on scale? Pages 4–15 in Peterson DL, Parker VT, eds. *Ecological Scale: Theory and Applications*. New York: Columbia University Press.
- O'Neill RV, Johnson AR, King AW. 1989. A hierarchical framework for the analysis of scale. *Landscape Ecology* 3: 193–205.
- Oreskes N, Shrader-Frechette K, Belitz K. 1994. Verification, validation, and confirmation of numerical models in the earth sciences. *Science* 263: 641–646.
- Orrock JL, Pagels JF, McShea WJ, Harper EK. 2000. Predicting presence and abundance of a small mammal species: The effect of scale and resolution. *Ecological Applications* 10: 1356–1366.
- Pearce J, Ferrier S. 2000. Evaluating the predictive performance of habitat models developed using logistic regression. *Ecological Modelling* 133: 225–245.
- Peterson DL, Parker VT. 1998. *Ecological Scale: Theory and Applications*. New York: Columbia University Press.
- Pidgeon AM, Radeloff VC, Mathews NE. 2003. Landscape-scale patterns of black-throated sparrow (*Amphispiza bilineata*) abundance and nest success. *Ecological Applications* 13: 530–542.
- Pielke RA Jr, Sarewitz D, Byerly R Jr. 2000. Decision making and the future of nature: Understanding and using predictions. Pages 361–387 in Sarewitz D, Pielke RA Jr, Byerly R Jr. *Prediction: Science, Decision Making, and the Future of Nature*. Washington (DC): Island Press.
- Poiani KA, Bedford BL, Merrill MD. 1996. A GIS-based index for relating landscape characteristics to potential nitrogen leaching to wetlands. *Landscape Ecology* 11: 237–255.
- Popper KR. 1959. *The Logic of Scientific Discovery*. New York: Basic Books.
- Rastetter EB. 1996. Validating models of ecosystem response to global change. *BioScience* 46: 190–198.
- Rastetter EB, King AW, Cosby GM, Hornberger GM, O'Neill RV, Hobbie JE. 1992. Aggregating fine-scale ecological knowledge to model coarser-scale attributes of ecosystems. *Ecological Applications* 2: 55–70.
- Reiners WA, Driese KL. 2001. The propagation of ecological influences through heterogeneous environmental space. *BioScience* 51: 939–950.
- Running SW, Nemani RR, Peterson DL, Band LE, Potts DF, Pierce LL. 1989. Mapping regional forest evapotranspiration and photosynthesis by coupling satellite data with ecosystem simulation. *Ecology* 70: 1090–1101.
- Schneider DC. 2001. The rise of the concept of scale in ecology. *BioScience* 51: 545–553.
- Scott JM, Heglund PJ, Morrison ML, eds. 2002. *Predicting Species Occurrences: Issues of Accuracy and Scale*. Washington (DC): Island Press.
- Stewart TR. 2000. Uncertainty, judgment, and error in prediction. Pages 41–57 in Sarewitz D, Pielke RA Jr, Byerly R Jr, eds. *Prediction: Science, Decision Making, and the Future of Nature*. Washington (DC): Island Press.
- Thapa K, Bossler J. 1992. Accuracy of spatial data used in geographic information systems. *Photogrammetric Engineering and Remote Sensing* 58: 835–841.
- Turner MG, Dale VH, Gardner RH. 1989a. Predicting across scales: Theory development and testing. *Landscape Ecology* 3: 245–252.
- Turner MG, O'Neill RV, Gardner RH, Milne BT. 1989b. Effects of changing spatial scale on the analysis of landscape pattern. *Landscape Ecology* 3: 153–162.
- [VEMAP] Vegetation/Ecosystem Modeling and Analysis Project. 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<sub>2</sub> doubling*. *Global Biogeochemical Cycles* 9: 407–437.
- Webster R, Oliver M. 2001. *Geostatistics for Environmental Scientists*. Chichester (United Kingdom): John Wiley and Sons.
- Wiens JA. 1989. Spatial scaling in ecology. *Functional Ecology* 3: 385–397.
- . 1995. Landscape mosaics and ecological theory. Pages 1–26 in Hansson L, Fahrig L, Merriam G, eds. *Mosaic Landscapes and Ecological Processes*. London: Chapman and Hall.