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

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.
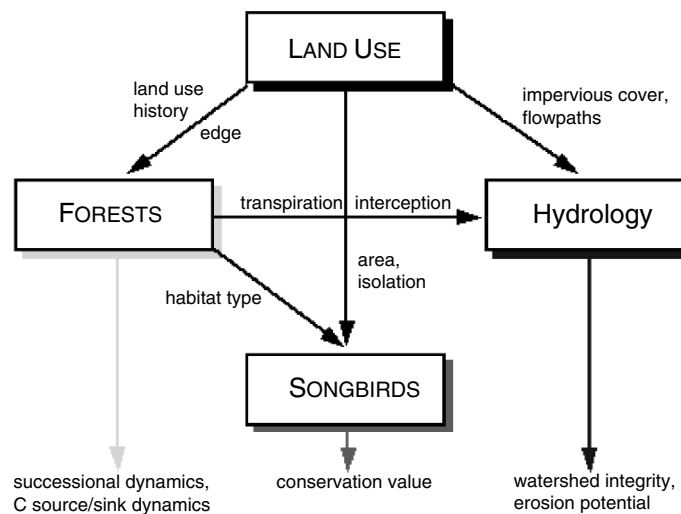

Figure 13.1. Schematic of linkages among research themes in the Triangle Landscape Change Project.

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 $\pm$ 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.
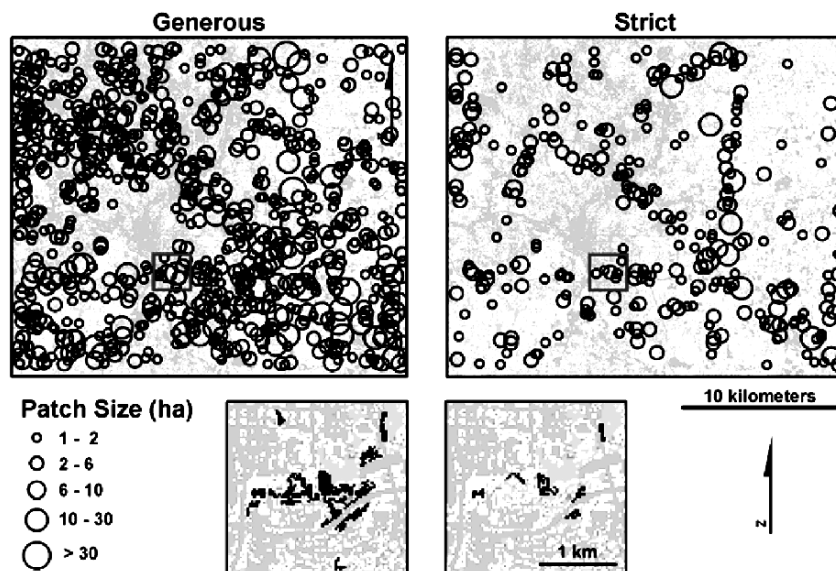

Figure 13.2. *Contrasting maps of potential wood thrush habitat, defined generously (left) and strictly (right) in terms of hardwood basal area. Because the shaded clusters of cells are difficult to resolve, patches are overlaid with circles indicating their relative sizes. Insets are expanded to highlight 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:

$$p_{ij} = exp(-\theta\, d_{ij}) \cdot a_j \qquad (13.1)$$

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 ($N \gg 1000$'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 $\gg 1.0$).

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.
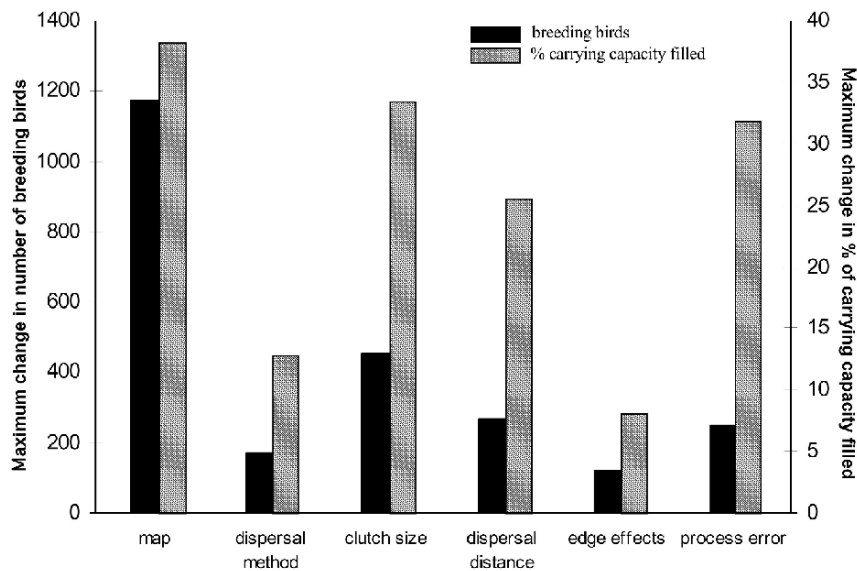

Figure 13.3. *Relative importance of each simulated factor as a source of uncertainty in the metapopulation model, based on ANOVA. Analyses were conducted separately for total population size (left-side axis) and percent of carrying capacity (right-side axis) because the two habitat maps had very different 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).
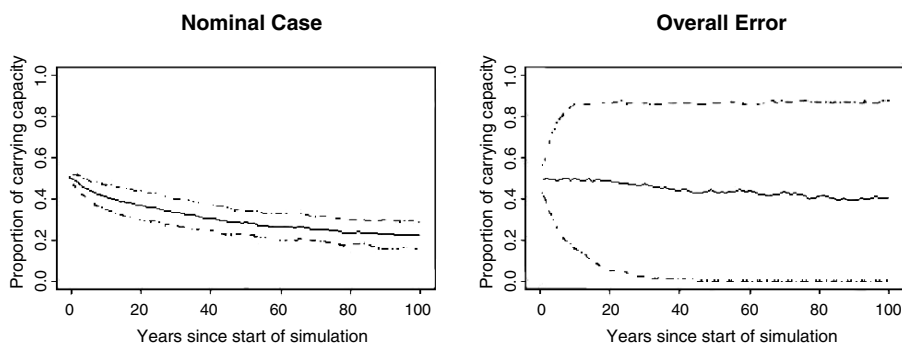

Figure 13.4. *Illustration of total uncertainty as propagated through the model, compared to the confidence limits generated by stochastic "process error" alone.*

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.
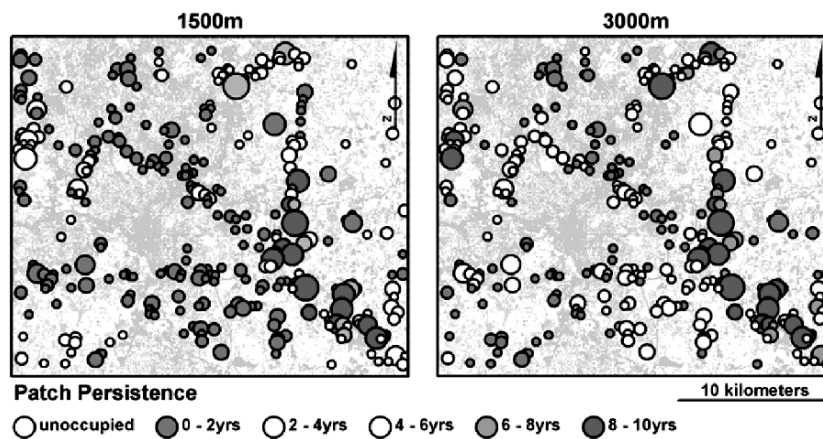

Figure 13.5. *Local effects of model uncertainty, as generated by variation in dispersal distance as modeled using least-cost paths as dispersal routes. Circles are overlaid on habitat clusters to indicate their relative size (see Figure 13.2). Effects of connectivity on population persistence are expressed 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.