Statistical Agent-Based Models for Discrete Spatio-Temporal Systems

Mevin B. HOOTEN and Christopher K. WIKLE

Agent-based models have been used to mimic natural processes in a variety of fields, from biology to social science. By specifying mechanistic models that describe how small-scale processes function and then scaling them up, agent-based approaches can result in very complicated large-scale behavior while often relying on only a small set of initial conditions and intuitive rules. Although many agent-based models are used strictly in a simulation context, statistical implementations are less common. To characterize complex dynamic processes, such as the spread of epidemics, we present a hierarchical Bayesian framework for formal statistical agent-based modeling using spatiotemporal binary data. Our approach is based on an intuitive parameterization of the system dynamics and can explicitly accommodate directionally varying dispersal, long distance dispersal, and spatial heterogeneity.

KEY WORDS: Binary data; Cellular automata; Dynamical system; Hierarchical Bayesian model.

1. INTRODUCTION

Many types of ecological, environmental, and epidemiologic data are collected over discrete spatial and temporal domains (Krebs 1978; Hanski 1999; Waller and Gotway 2004). Moreover, such data are often binary-valued. Consider, for example, a natural process that evolves in some space ($\mathcal{S}$) over a set of time ($\mathcal{T}$) continuously. It is rarely feasible to observe such a process in a continuous fashion; in fact, it is impossible to do so if the process is being observed and (or) recorded digitally. Therefore, data collection schemes are often designed for convenience, and the process is observed in discrete snapshots, say, $y_{i,t}$, where $t \in \mathbf{T} \subset \mathcal{T}$ and $i \in \mathbf{S} \subseteq \mathcal{S}$ for $\dim(\mathbf{S}) < \infty$, where $\mathbf{S}$ and $\mathbf{T}$ are finite sets of spatial locations (or areas) and times (or periods), respectively.

Dynamic models over gridded spatial domains, such as those proposed by Hooten and Wikle (2008), can be used in cases where the data ($y_{i,t}$) are counts on domains with areal spatial support. In contrast, the spatiotemporal matrix models proposed by Hooten et al. (2007) can be used in cases where count data occur in either discrete or continuous space. Both types of models for characterizing spreading phenomena involve population growth and dispersal processes. It is often the case that data reflecting population growth are not available and only the occurrence of a phenomena is observed. In such cases where only binary data exist and without repeated measurements, population growth parameters are not identifiable (without strong prior information). The focus in such situations shifts to the estimation of dispersal-based dynamics.

Conventional methods for modeling spatial processes on partitioned domains often are specified as Markov random field models (also known as conditional autoregressive models and simultaneous autoregressive models) that have now been extended to the spatiotemporal setting (see Cressie 1993; Banerjee, Carlin, and Gelfand 2004 for a complete discussion). The autologistic model has served as the dominant spatial model for binary data (Besag 1974; Cressie 1993; Heikkinen and Hogmander 1994; Hogmander and Moller 1995; Hoeting, Leecaster, and Bowden 2000), where the probability of presence is conditioned on its neighbors. More recently, methods have been developed that extend the autologistic model, as well as other models using a logit transform of presence/absence probability, to spatiotemporal settings (e.g., Zhu, Huang, and Wu 2005; Royle and Dorazio 2006; Royle and Kery 2007).

Many previously developed methods address the dynamic nature of these processes from an autoregressive modeling perspective (e.g., Pace et al. 2000), where a state of the system at the previous time is related to the state at the current time through (possibly time-varying) spatial and temporal autoregressive coefficients. This is a very powerful approach when considering the process as a whole and in settings where a Gaussian autoregressive model can be used. Moreover, Royle and Dorazio (2008) discussed various generalizations to binary data and argued that certain autoregressive generalized linear specifications can be convenient to implement, especially when covariates are involved directly through the link function.

We take an alternative approach in the methods that follow, and introduce a class of models motivated by viewing the movement (i.e., dispersal) of an agent from the perspective of the agent itself, rather than the system as a whole. Such approaches have been termed "agent-based models" (ABMs) in the literature (Grimm and Railsback 2005). For example, an organism, such as an exotic invasive species or a pathogenic species, in a new environment often will move from areas of lower quality to areas of higher quality (where quality is defined in terms of many possible factors, from environmental suitability to overpopulation to availability of hosts). In this way, the system as a whole (i.e., the automaton) can be thought of as a process with numerous automatous components. It is important to note that an agent could be defined at multiple scales. For example, an agent could represent a virus, an individual, a node in a network, or, more generally, a spatial reporting unit (e.g., a census tract, township, county, or state). In Section 2 we present statistical ABMs that allow the spatiotemporal process to propagate through areal reporting units with directionally varying rates

Mevin B. Hooten is Assistant Professor of Statistics, Department of Mathematics and Statistics, Utah State University, Logan, UT 84322-3900 (E-mail: *mevin.hooten@usu.edu*). Christopher K. Wikle is Professor of Statistics, Department of Statistics, University of Missouri, Columbia, MO 65211 (E-mail: *wiklec@missouri.edu*). In addition to invaluable suggestions provided by the associate editor and anonymous reviewers, the authors thank Lance Waller, Devin Johnson, Noel Cressie, Jun Zhu, Andy Royle, Jim Powell, and Ali Arab for providing data, comments, and numerous helpful suggestions.

Journal of the American Statistical Association
March 2010, Vol. 105, No. 489, Theory and Methods
DOI: 10.1198/jasa.2009.tm09036

and spatial heterogeneity that is linked to the gradients in the underlying environment.

1.1 Cellular Automata

The properties and behavior of automata have been studied in nearly every field (Wolfram 1984), but rarely in a statistical context. In essence, an automata can be defined as an entity or agent, whose state (state can be taken to be as abstract as necessary in this context) is governed by a set of simple rules, and whose interaction with its environment can result in extremely complex systemic behavior. In other words, the combined behavior of numerous individual automata results in very complicated large-scale dynamical system behavior incapable of being described in any other way (Wolfram 1983).

Automata are most often defined in a deterministic dynamical system framework in which the state of the "neighborhood" of an agent (in this case, an areal unit of space called a "cell") determines the future state of the agent. Another type of automata can be formulated probabilistically, where the state of an agent is defined by parametric probability distributions and the behavior of the system as a whole given the probability rule is not unique, but can be expressed in terms of likelihood (Lee et al. 1990). A propagating automatous system defined in either manner is capable of exhibiting spatially irregular wavelike behavior commonly found in natural phenomena (Hogeweg 1988). In this case, we consider "spatially irregular" as referring to spatially varying dynamics that are either anisotropic (i.e., varying directionally) or nonstationary (i.e., varying locationally), or both (e.g., Cressie 1993).

There is a fundamental difference in the construction of models through the "top-down" approach versus the "bottom-up" approach (Grimm et al. 2005). Traditionally, bottom-up approaches (i.e., Lagrangian) to studying ecosystem function and ecological processes have been used in simulation settings, whereas top-down approaches (i.e., Eulerian) have been taken using statistical methodology (e.g., Wikle et al. 2001; Wikle 2003). Both approaches have contributed much to scientific theory and application, although they rarely intermingle. In what follows in Section 2, we present a bottom-up hierarchical statistical framework for studying spreading spatiotemporal processes when only binary data are available, using AMBs with dynamical behavior modeled by cellular automata. Then in Section 3, we discuss the relationships between these statistical ABMs and various alternatives, such as partial differential equation models, generalized linear models, and Markov random field models.

1.2 Motivating Example: Rabies Epidemics in Raccoon Populations

Although this article is largely focused on methodological developments, we feel that a preliminary discussion of a motivating example will help elucidate and justify the specifications that follow. Therefore, consider as an example, the ongoing rabies epidemic on the east coast of the United States that began in the mid-1970's (Nettles et al. 1979; Wheeler and Waller 2008). This epidemic is a natural spatiotemporal process where binary data are available on a partitioned spatial domain and exhibit irregular spreading behavior. Specifically, space–time data documenting the spread of rabies through the raccoon population in Connecticut are available in a "time since arrival in township" format (e.g., Smith et al. 2002); that is, after first appearing in 1991 in the western townships of Connecticut (those bordering New York), rabies spread throughout the state over the subsequent 5-year period.

The data set shown in Figure 1 (Section 2.4) illustrates the spatiotemporal behavior of rabies epidemic in Connecticut in 1991–1995, during which several long-distance dispersal events occurred in addition to anisotropic and nonstationary neighborhood-based dispersal patterns. Smith et al. (2002) used an agent-based simulation model in the spirit of an automatous system, but focusing on the temporal domain by characterization of rates of spread from neighboring townships, and found that such models can be useful for studying this particular natural phenomenon. Moreover, they allowed for long-distance dispersal and spatially varying rates of spread between neighboring townships, and found a correlation with environmental covariates such as proximity of a large river (i.e., the Connecticut River) and human population density.

We feel that the spreading rabies epidemic is an excellent example to illustrate the robustness of ABM approaches. Thus the general models developed in Sections 2.2 and 2.3 are demonstrated in Section 2.4 to characterize the rabies epidemic from the same perspective as that of Smith et al. (2002), but in a rigorous probability framework in which inference on the dispersal parameters can be made from a statistical perspective given the available data.

2. AGENT–BASED STATISTICAL MODELS

2.1 Agent-Based Dynamics

The agent-based dynamical system known as a cellular automaton (CA) operates on discrete support and can be used to model naturally propagating phenomena. In such systems, a model defining the dynamics of the process can still be formulated with a state–space representation, the primary difference being that the CA model is not based on a continuous operator such as the case with differential equation operators; rather, the propagating operator h, is now defined by a neighborhood $\mathcal{N}$ and a set of rules $\mathcal{R}$ that map the past state of the neighborhood to the current state. In this way, an automaton (i.e., the entire partitioned domain made up of individual automata or cells) contains numerous interacting components, the states of which are governed by a set of rules based on the previous state. These characteristics of automatous systems make them especially suited to high-performance computing environments, such as parallel processing and cluster computing settings (Wolfram 1988).

In a deterministic CA, the state at time t for cells $1, \ldots, m$, is denoted by $\mathbf{u}_t = (u(1,t), u(2,t), \ldots, u(m,t))'$ and is mapped exactly to the previous state ($\mathbf{u}_{t-1}$), so that the state of the automaton $\mathbf{u}_t$ is unique given $\mathbf{u}_{t-1}$. More specifically, each element $u_{i,t} \equiv u(i,t)$ (the automata) results as a function of a forward operator h and the previous set of neighboring automata, $\mathbf{u}_{\mathcal{N}_i,t-1} \equiv \{\text{all } u_{i,t-1} \text{ in the neighborhood, } \mathcal{N}_i, \text{ of cell } i\}$. The function h is often composed of a set of rules, defining the map from all possible states of the set $\mathbf{u}_{\mathcal{N}_i,t-1}$ to the possible states of $u_{i,t}$. An automaton defined with this simple set of rules

can exhibit strikingly complex behavior as a dynamical system. A deterministic automaton on a two-dimensional spatial domain can be defined similarly.

CA systems are capable of exhibiting very complex behavior and, from a modeling perspective, are very powerful tools. The use of deterministic automatons is challenging from a statistical modeling perspective, however, because the rule space quickly becomes cumbersome as the size of the neighborhood increases, the dimensionality of the system increases, and (or) the support of the automata is extended. For example, a mapping function h for a simple first-order neighborhood in one dimension has only $2^3 = 8$ possible neighborhood structures; this leads to $2^{(2^3)} = 256$ possible rules for the binary automaton. In general, for a neighborhood of dimension $d_{\mathcal{N}}$, there will be $2^{(2^{d_{\mathcal{N}}})}$ possible rules. Thus, for a two-dimensional spatial domain, with a binary automaton and Queen's neighborhood (i.e., $d_{\mathcal{N}} = 9$), the number of possible rules is 1.34×10^{154}. If the rule is treated as an unknown set of parameters, then, in terms of estimation, we would need to search its support for rules that capture the observed spatiotemporal behavior of the system under study. Unfortunately, searching the space of rules for CA systems with large neighborhoods is not feasible; thus it is necessary to either simplify the rule space using assumptions or take an alternative approach.

A stochastic or probabilistic automata can be thought of as a random variable conditioned on the neighborhood structure at the previous time. For example, the evolution equation of a probabilistic automata can be written as

$$u_{i,t}|\mathbf{u}_{\mathcal{N}_i,t-1} \sim \left[u_{i,t}|h\left(\mathbf{u}_{\mathcal{N}_i,t-1}\right)\right], \tag{1}$$

whereas the evolution equation of a deterministic automata is written as

$$u_{i,t} = h\left(\mathbf{u}_{\mathcal{N}_i,t-1}\right). \tag{2}$$

In the case of the probabilistic automata, the square bracket notation $[\cdot]$ refers to a probability distribution. For the binary automaton (i.e., $u_{i,t} \in \{0, 1\}$), the probability distribution $[u_{i,t}|h(\mathbf{u}_{\mathcal{N}_i,t-1})]$ could be a Bernoulli distribution with the Bernoulli probability a function of the neighborhood at the previous time and some parameters. Note that the specification in (1) differs from Gaussian state–space models only in that the source of variability in the probabilistic CA (1) is more general than the additive Gaussian error term in a probabilistic partial differential equation (e.g., Hooten and Wikle 2008).

The advantage of the probabilistic specification (1) over the deterministic specification (2) is that from a statistical perspective, it is easier to search the space of rules in terms of probability than it would be to do so exactly. The disadvantage is that the map from the previous state to the current state is no longer unique. Probabilistic CA models are sometimes referred to as interacting particle systems (Turchin 1998) and often are considered in continuous time (i.e., asynchronous updating) to aid in the investigation of their theoretical properties (Durrett and Levin 1994). In what follows, we concentrate on the discrete time situation (i.e., synchronous updating) for computational reasons, although in Section 3.1 we show that scaling up from our agent-based Lagrangian model leads to an Eulerian model based on partial differential equations.

2.2 Anisotropic Stationary Dispersal on Homogeneous Discrete Domains

One approach to modeling a spatiotemporal process such as the rabies process discussed in Section 1.2 is to estimate the dynamics of spread in terms of probability of presence. That is, assume that the phenomenon in question is present ($y_{i,t} = 1$) at the spatial location or cell i and time t with probability $\theta_{i,t}$ and absent ($y_{i,t} = 0$) with probability $1 - \theta_{i,t}$. (To simplify notation, we use only one index to refer to a location in two-dimensional space.) This is a common form for the classes of models known as occurrence and occupancy models (e.g., MacKenzie et al. 2002; MacKenzie et al. 2003; Royle and Dorazio 2008). A more convenient specification for this data model is $y_{i,t}|\theta_{i,t} \sim \text{Bern}(\theta_{i,t})$. Now, rather than model the transformed probability (as either a logit or probit transform) at the next level in the hierarchy, consider the following specification for $\theta_{i,t}$:

$$\theta_{i,t} = y_{i,t-1}\phi + (1 - y_{i,t-1})\left(I_{\mathcal{N}_{i,t-1}}\right)\bar{p}_{i,t} + \left(1 - I_{\mathcal{N}_{i,t-1}}\right)\psi \tag{3}$$

for all i and t in the spatiotemporal support of the data, and where

$$I_{\mathcal{N}_{i,t-1}} = \begin{cases} 1, & \text{if } \sum_{j\in\mathcal{N}_i} y_{j,t-1} > 0 \\ 0, & \text{if } \sum_{j\in\mathcal{N}_i} y_{j,t-1} = 0 \end{cases}$$

and, $\mathcal{N}_i$ is the spatial neighborhood of area i. In a regularly gridded two-dimensional domain with a Queen's neighborhood, $\mathcal{N}_i$ can be depicted as

$$\mathcal{N}_i = \begin{bmatrix} N_{j=i-4} & N_{j=i-1} & N_{j=i+2} \\ N_{j=i-3} & N_{j=i} & N_{j=i+3} \\ N_{j=i-2} & N_{j=i+1} & N_{j=i+4} \end{bmatrix} = \begin{bmatrix} N_1 & N_4 & N_7 \\ N_2 & N_5 & N_8 \\ N_3 & N_6 & N_9 \end{bmatrix}. \tag{4}$$

Note that in (3), because $\mathcal{N}_i$ includes cell i, each term in the sum represents probabilities for following three disjoint situations:

1. $y_{i,t-1}$: Cell i is occupied at time $t - 1$.
2. $(1 - y_{i,t-1})(I_{\mathcal{N}_{i,t-1}})$: At least one neighbor of cell i is occupied at time $t - 1$, but cell i itself is not occupied.
3. $1 - I_{\mathcal{N}_{i,t-1}}$: No cell in $\mathcal{N}_i$ is occupied at time $t - 1$ (including cell i).

In (3), the probability ϕ corresponds to persistence of the phenomena; that is, once the phenomena is present in cell i, the probability that it will be present at the next time step is ϕ. The probability ψ corresponds to out-of-neighborhood dispersal; that is, the phenomena will occur in an area outside of those occupied with probability ψ. This is a somewhat naive specification for what is referred to in the ecological literature as long-distance dispersal (e.g., Clark 2003). In some cases it would be more realistic to let ψ vary with distance from source or other environmental covariates. Without appropriate a priori knowledge of the mechanisms of long-distance dispersal and data to support this notion, a naive specification [such as in (3)] at least allows for its effect.

The probability $\bar{p}_{i,t}$ is the component of the process that handles the neighborhood-based dispersal in this model. When building our agent-based statistical model, many specifications for $\bar{p}_{i,t}$ are possible depending on the spatial domain, level of

realistic detail desired in the model, and estimability of $\bar{p}_{i,t}$. In this case, let

$$\bar{p}_{i,t} = 1 - \exp\big(\big(\mathbf{y}_{\mathcal{N}_i,t-1}\big)' \log(\mathbf{1} - \mathbf{p})\big), \tag{5}$$

where, $\mathbf{y}_{\mathcal{N}_i,t-1} = [y_{N_1,t-1}, \ldots, y_{N_9,t-1}]'$ is a vector of the data (i.e., 1s and 0s) corresponding to the neighborhood of the ith cell. The probabilities in $\mathbf{p}$ are $[\mathbf{p} = (p_{N_9}, p_{N_8}, \ldots, p_{N_1})']$, where p_{N_j} refers to the transition probability from cell i to its jth neighbor and are constant over the spatial domain. This specification allows for the probability of presence at time t, in an area that has been previously unoccupied, to be the union of transition probabilities from occupied neighbors. To show this result [i.e., (5)], denote the event that the phenomenon propagates from neighboring area N_j into area i at time t as event $E_{N_j,t}$; then the probability that area i becomes occupied at time t, given its neighborhood at time $t-1$, is

$$\begin{aligned}
&\mathrm{P}\Big(\bigcup_j E_{N_j,t};\ \text{for all occupied neighbors } j\Big)\\
&\quad = 1 - \mathrm{P}\Big(\bigcap_j E^c_{N_j,t};\ \text{for all occupied neighbors } j\Big)\\
&\quad = 1 - \prod_j \mathrm{P}\big(E^c_{N_j,t}\big)\\
&\quad = 1 - \prod_j \big(1 - \mathrm{P}\big(E_{N_j,t}\big)\big),
\end{aligned}$$

where the assumption is made that the phenomenon moves from neighboring area j into area i independently of the phenomenon moving from neighbor k into area i at time t for all occupied areas j and k in the neighborhood of area i. This assumption of independent propagation, conditional on the parameters $\mathbf{p}$, is strong and implies that phenomena entering a new area from two (or more) neighboring areas will not interact as they do so. This assumption does, however, allow the analytical calculation of $\bar{p}_{i,t}$ in the absence of significant information about the numerous possible intersections of events $(E_{N_j,t})$ that would be required to calculate the probability of the union.

Assuming that the transition probabilities in $\mathbf{p}$ sum to 1 and that the propagating process can be modeled by stationary dynamics, a Dirichlet probability model for $\mathbf{p}$ is the natural choice [i.e., $\mathbf{p}|\mathbf{a} \sim \mathrm{Dir}(\mathbf{a})$]. The Dirichlet parameters $(\mathbf{a})$ can then be modeled as $\log(\mathbf{a}) \sim \mathrm{N}(\boldsymbol{\mu}_a, \boldsymbol{\Sigma}_a)$. The Dirichlet model here is what sets these models apart from generalized linear models. It becomes vital later, in Section 2.3, when we connect the directionally varying dispersal probabilities to underlying environmental gradients.

It is important to point out two consequences concerning the effect of the probabilities, $\mathbf{p}$, on the propagating phenomenon: First, this specification assumes that the neighborhood-based dispersal of the process is dynamically anisotropic but stationary; that is, though the process can spread with different probabilities in different directions, these probabilities remain homogeneous over the entire spatial domain. The addition of persistence and nonneighborhood dispersal terms allows for more realistic behavior. A simpler form of dispersal would have a phenomena propagating to all neighboring areas with an equal probability. Second, in the case with a regularly gridded spatial domain and Queen's neighborhood, allowing the dimension of $\mathbf{p}$ (i.e., $d_{\mathcal{N}}$) to be eight assumes that the propagating phenomena will indeed propagate to a neighboring area. Letting $d_{\mathcal{N}} = 9$ in this setting allows the phenomena to remain in the current area with some probability; however, because the probability of persistence is being modeled in θ by ϕ, the addition of p_i seems redundant. It is important to note that in this setting, p_i will not affect the persistence, due to the $(y_{i,t-1}\phi)$ term in (3), but it will affect the overall rate of dispersal. A larger p_i will slow the propagating phenomena without affecting the anisotropy of the process, thus providing the model with additional flexibility.

In a Bayesian implementation of the model, priors for ϕ and ψ can be specified in terms of Beta distributions and result in conjugate full-conditional distributions. Thus the priors $[\phi] = \mathrm{Beta}(\alpha_\phi, \beta_\phi)$ and $[\psi] = \mathrm{Beta}(\alpha_\psi, \beta_\psi)$, as well as the hyperpriors $\alpha_\phi, \beta_\phi, \alpha_\psi, \beta_\psi$, $\boldsymbol{\mu}_a$, and $\boldsymbol{\Sigma}_a = \sigma_a^2\mathbf{I}$, result in the joint posterior for this model,

$$\begin{aligned}
&[\mathbf{p}, \mathbf{a}, \phi, \psi | \{y_{i,t} \text{ for } t = 1, \ldots, T \text{ and } i = 1, \ldots, m\}]\\
&\qquad \propto \prod_{t=1}^{T}\prod_{i=1}^{m} [y_{i,t}|\theta_{i,t}(\mathbf{p}, \phi, \psi)][\mathbf{p}|\mathbf{a}][\mathbf{a}][\phi][\psi]. \qquad (6)
\end{aligned}$$

It can be shown (see the Appendix for details) that the full-conditional distributions for the model parameters, required for implementation in a Markov chain Monte Carlo (MCMC) setting, are

$$\phi|\cdot \sim \mathrm{Beta}\Bigg(\sum_{t=2}^{T}\sum_{i\in\mathcal{A}_{\phi_{t-1}}} y_{i,t} + \alpha_\phi, \sum_{t=2}^{T}\sum_{i\in\mathcal{A}_{\phi_{t-1}}} (1 - y_{i,t}) + \beta_\phi\Bigg),$$

$$\psi|\cdot \sim \mathrm{Beta}\Bigg(\sum_{t=2}^{T}\sum_{i\in\mathcal{A}_{\psi_{t-1}}} y_{i,t} + \alpha_\phi, \sum_{t=2}^{T}\sum_{i\in\mathcal{A}_{\psi_{t-1}}} (1 - y_{i,t}) + \beta_\psi\Bigg),$$

$$\mathbf{p}|\cdot \sim [\mathbf{p}|\cdot] \propto \prod_{t=2}^{T}\prod_{i\in\mathcal{A}_{\bar{p}_i}} \mathrm{Bern}(y_{i,t}|\bar{p}_{i,t}) \times \mathrm{Dir}(\mathbf{p}|\mathbf{a}),$$

$$\mathbf{a}|\cdot \sim [\mathbf{a}|\cdot] \propto \mathrm{Dir}(\mathbf{p}|\mathbf{a}) \times \mathrm{N}(\log(\mathbf{a})|\boldsymbol{\mu}_a, \sigma_a^2\mathbf{I}),$$

where the $\mathcal{A}$ are sets of indexes defined by

$$\begin{aligned}
\mathcal{A}_{\phi_{t-1}} &= \{j|\forall j \text{ such that } y_{j,t-1} = 1\},\\
\mathcal{A}_{\psi_{t-1}} &= \big\{j|\forall j \text{ such that } \big(1 - I_{\mathcal{N}_{i,t-1}}\big) = 1\big\},\\
\mathcal{A}_{\bar{p}_i} &= \big\{j|\forall j \text{ such that } (1 - y_{j,t-1})\big(I_{\mathcal{N}_{i,t-1}}\big) = 1\big\}.
\end{aligned}$$

In cases where $y_{i,t}$ is not observed for all i in the spatial domain, the posterior predictive distribution can be found for the missing data $(\tilde{\mathbf{y}}_t)$ using composition sampling,

$$\begin{aligned}
[\tilde{\mathbf{y}}_t|\mathbf{y}_t, \forall t] = \int\cdots\int &[\tilde{\mathbf{y}}_t|\mathbf{p}, \mathbf{a}, \phi, \psi]\\
&\times [\mathbf{p}, \mathbf{a}, \phi, \psi|\mathbf{y}_t, \forall t]\, d\mathbf{p}\, d\mathbf{a}\, d\phi\, d\psi.
\end{aligned}$$

2.3 Anisotropic Nonstationary Dispersal on Heterogeneous Discrete Domains

The model presented in the previous section is very robust in that it allows for anisotropic dynamic behavior in the process that is capable of dispersing at varying rates. Like the matrix model used by Hooten et al. (2007), it relies heavily on the

dynamics of the process to model the data. From a scientific perspective, it is often of interest to make inference about the propagating nature of the process not only in the dynamics, but also in the structure regarding the likely heterogeneous environment. In the situation where a phenomenon is propagating in space over time, it is reasonable to think that certain areas of the spatial domain are more suitable than others. The ecological literature refers to this notion as "habitat suitability" or "habitat preference," depending on the residence of the phenomenon in question relative to the surrounding environment (e.g., Hirzel et al. 2002). From a dynamic modeling perspective, it seems natural to think that a phenomenon will most likely move from areas of undesirable habitat to areas of desirable habitat. Note that the term "habitat" is used very generally here and in regard to the successful propagation of the phenomenon under study.

Distinguishing areas of suitable or desirable habitat from unsuitable areas is a long-studied subject in ecology (Manly et al. 2002). One of the simplest and most common approaches to mapping habitat suitability is to find the associations between organism abundance (or presence) and environmental covariates (e.g., Hooten, Larsen, and Wikle 2003; Gelfand et al. 2006). Due to limitations imposed through data collection schemes, such analyses are often focused in a static temporal domain over space. The study of phenomena that are actively propagating into new areas or reoccupying previous areas must involve some dynamic component; the challenge is in accounting for covariate effects on the dynamics.

In what follows, we begin by specifying a model for suitability (herein "suitability" and "preference" are used interchangeably) and then connect it to the transition probabilities in our neighborhood-based dispersal ABM. First, let the suitability of a spatial domain, partitioned into m areas, be denoted by $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \ldots, \alpha_m]'$. Scientifically, a large portion of the variability in $\boldsymbol{\alpha}$ is often thought to be explained by a set of environmental covariates, $\mathbf{X}$, that may be linked to $\boldsymbol{\alpha}$ by some function f, which could be parametric or semiparametric and possibly could contain correlated and (or) nonstationary error. Assuming a linear function f and Gaussian error $\boldsymbol{\varepsilon}$, the standard linear model is

$$\boldsymbol{\alpha} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}. \tag{7}$$

Assuming that the error contains residual spatial dependence, and letting $\boldsymbol{\varepsilon} \sim \mathrm{N}(\mathbf{0}, \boldsymbol{\Sigma}_\alpha)$, where $\boldsymbol{\Sigma}_\alpha = \sigma_\alpha^2 \exp(-\frac{\mathbf{D}}{\theta_\alpha})$, and $\mathbf{D}$ is an $m \times m$ matrix of Euclidean distances between area centroids, it is then possible to estimate the residual spatial structure as well as the covariate effects ($\boldsymbol{\beta}$) on habitat suitability. Note that if warranted, other suitable spatial covariance models could be used to define $\boldsymbol{\Sigma}_\alpha$, or, alternatively, a Markov random field model (rather than a geostatistical model) could be used to accommodate second-order spatial structure in $\boldsymbol{\alpha}$. In this specification (7), each coefficient β_j positively associates the jth covariate with suitability and thus increases the probability of movement into more suitable cells from less suitable cells.

Because the propagating phenomena are attracted to areas of higher suitability, an attraction model is useful for explaining the probability of movement. One such model is based on partial differential equations and can be used in this context. Consider the spatial field, $\boldsymbol{\alpha}$, exhibiting the preferred habitat of an organism. Given that the phenomena is present at location i, the probability that it will propagate to location j depends, at least in part, on the quality of habitat relative to other surrounding habitats. The continuous version of this type of attraction can be represented as a partial differential equation,

$$\tilde{\mathbf{A}} = \nabla_{\mathcal{S}}(\boldsymbol{\alpha}) = \frac{\partial \boldsymbol{\alpha}}{\partial \mathcal{S}}, \tag{8}$$

where the partial derivative of the environmental preference field, $\boldsymbol{\alpha}$, is taken with regard to space ($\mathcal{S}$) and thus the gradient operator ($\nabla_{\mathcal{S}}$) will be infinite-dimensional in the sense that there will be a different tangent for every directional derivative. In a discrete context however, the gradient in (8) can be approximated by a finite number of directional difference equations. In the specific case of the regular two-dimensional gridded spatial domain and Queen's neighborhood, $\tilde{\mathbf{A}}$ is a matrix, and the directional derivative at location i can be approximated by a vector of differences $\tilde{\mathbf{a}}_i$, where $\tilde{\mathbf{A}} = [\tilde{\mathbf{a}}_1, \ldots, \tilde{\mathbf{a}}_m]'$. Each element of $\tilde{\mathbf{a}}_i$ is then a function of the spatial field of habitat suitability ($\boldsymbol{\alpha}$),

$$\tilde{\mathbf{a}}_i = \begin{cases} \tilde{a}_{i,1} = \dfrac{\alpha_{N_1,i} - \alpha_i}{d_{N_1,i}}, \\ \tilde{a}_{i,2} = \dfrac{\alpha_{N_2,i} - \alpha_i}{d_{N_2,i}}, \\ \vdots \\ \tilde{a}_{i,9} = \dfrac{\alpha_{N_9,i} - \alpha_i}{d_{N_9,i}}, \end{cases}$$

where N_j, i refers to the jth neighbor of cell i in the Queen's neighborhood [$\mathcal{N}_i$ in (4)] and $d_{N_j,i} = d(N_j, i)$ is the Euclidean distance between the centroids of cell i and its jth neighboring cell.

The utility of $\tilde{\mathbf{A}}$ is as a likelihood of movement from one area to another for a propagating phenomena. A transformation of $\tilde{\mathbf{A}}$ allows for its inclusion as parameters in a model similar to that of Section 2.2. Because the transition probabilities, $\mathbf{p}$, are the focal point of the neighborhood-based dynamics in this class of models, allowing them to depend hierarchically on the approximate gradients $\tilde{\mathbf{A}}$ is critical to the utilization of covariate information and nonstationary anisotropic dynamics in the process. Therefore, let $\mathbf{A}$ be a function of the directional gradient fields in $\tilde{\mathbf{A}}$ [e.g., $\mathbf{A} = g(\tilde{\mathbf{A}})$] such that they meet the requirements for hyperparameters of a Dirichlet distribution (i.e., positive, finite support). The choice of function g is subjective and may be somewhat arbitrary. In this case, we use the probit or standard normal cdf function that maps a set of real numbers to the set of positive real numbers, $\mathbf{a}_i = g(\tilde{\mathbf{a}}_i) = \Phi(\tilde{\mathbf{a}}_i)$. A more flexible form of such a transformation is one in which the parameters $\mathbf{A}$ are proportional only to the transforming function of $\tilde{\mathbf{A}}$,

$$\mathbf{a}_i = g(\tilde{\mathbf{a}}_i) = c\Phi(\tilde{\mathbf{a}}_i), \tag{9}$$

where c is a multiplicative scalar modeled at a lower level (or user-defined, serving as a tuning parameter that could be estimated in a similar fashion as in Box–Cox transformations or cross-validation). Letting the vector of transition probabilities vary for each spatial location i and depend on the parameters $\mathbf{a}_i$ through the Dirichlet distribution gives the probability model that characterizes the neighborhood-based dynamics of the propagating phenomena,

$$\mathbf{p}_i | \mathbf{a}_i \sim \mathrm{Dir}(\mathbf{a}_i) \quad \text{for } i = 1, \ldots, m. \tag{10}$$

Recall, from (3), that the actual probability of presence in cell i when cell i was previously unoccupied and at least one neighbor at the previous time was occupied is dictated by the parameter $\bar{p}_{i,t}$. The calculation of $\bar{p}_{i,t}$ in the stationary model was fairly straightforward, because $\mathbf{p}$ did not vary with space. Using the same data model from the previous section [i.e., $y_{i,t}|\theta_{i,t} \sim \text{Bern}(\theta_{i,t})$] and specification for the probability of presence ($\theta_{i,t}$) as in (3), a different function mapping $\mathbf{p}_{\mathcal{N}_i}$ to $\bar{p}_{i,t}$ must be used. For this nonstationary model, let $\bar{p}_{i,t}$ be defined as

$$\bar{p}_{i,t} = 1 - \exp\left(\left(\mathbf{y}_{\mathcal{N}_i,t-1}\right)' \log\left(\mathbf{1} - \mathbf{p}_{\mathcal{N}_i}\right)\right), \tag{11}$$

where $\mathbf{y}_{\mathcal{N}_i,t-1}$ is defined as in the previous section, $\mathbf{p}_{\mathcal{N}_i} = [p_{N_9,i}, p_{N_8,i}, \ldots, p_{N_1,i}]'$, and $p_{N_j,i}$ refers to the transition probability from the jth neighboring cell into cell i. In this way, the probability of presence in cell i at the current time is the probability of the unions of transitions into cell i from occupied neighboring cells at the previous time.

To complete the hierarchical model, let $\boldsymbol{\beta} \sim \text{N}(\boldsymbol{\mu}_\beta, \boldsymbol{\Sigma}_\beta)$, where $\boldsymbol{\mu}_\beta$ is specified as a hyperprior and $\boldsymbol{\Sigma}_\beta$ is modeled at a lower level [i.e., $\boldsymbol{\Sigma}_\beta^{-1} \sim \text{Wish}((\nu \mathbf{S}_\beta)^{-1}, \nu)$] to handle any multicollinearity between covariates. Let the persistence and nonneighborhood-based probability of presence be defined as before, $\phi \sim \text{Beta}(\alpha_\phi, \beta_\phi)$ and $\psi \sim \text{Beta}(\alpha_\psi, \beta_\psi)$. Persistence and nonneighborhood dispersal also could be generalized so that they vary in space or time; the simplest approach would be to linearly link the probit or logit of ϕ and ψ to a set of spatial or temporal covariates.

The spatial covariance parameters on the habitat preference spatial field, σ_α^2 and θ_α, can be modeled in the conventional manner where $\sigma_\alpha^2 \sim \text{InvGamma}(q, r)$ and $\theta_\alpha \sim \text{Gamma}(\alpha_\theta, \beta_\theta)$. Identifying the spatial range parameter, θ_α, may be difficult if there is minimal residual spatial structure in the $\boldsymbol{\alpha}$ process. In addition, due to the formulation of the model, a large θ_α (implying greater spatial structure) may be counterproductive considering that the model is trying to find sharp differences between neighboring cells (i.e., due to the directional gradient fields in $\tilde{\mathbf{A}}$).

The posterior distribution of interest can now be written as

$$\begin{aligned}[\mathbf{p}_1, \ldots, \mathbf{p}_m, \mathbf{a}_1, \ldots, \mathbf{a}_m, \phi, \psi, \boldsymbol{\beta}, \boldsymbol{\Sigma}_\beta, \sigma_\alpha^2, \theta_\alpha | y_{i,t}, \forall i, t] \\ \propto \prod_{t=1}^{T} \prod_{i=1}^{m} [y_{i,t}|\theta(\mathbf{p}, \phi, \psi)_{i,t}] \prod_{i=1}^{m} [\mathbf{p}_i|\mathbf{a}(\boldsymbol{\alpha})_i] \\ \times [\boldsymbol{\alpha}|\boldsymbol{\beta}, \sigma_\alpha^2, \theta_\alpha][\boldsymbol{\beta}|\boldsymbol{\Sigma}_\beta][\boldsymbol{\Sigma}_\beta][\phi][\psi][\sigma_\alpha^2][\theta_\alpha].\end{aligned}$$

The full-conditional distributions for the model parameters necessary for implementation of this nonstationary anisotropic model can be found in a similar fashion as those in the previous section. Details of the derivation and implementation are provided in the Appendix.

2.4 Application: Rabies Epidemic in Raccoon Populations

This section presents results and relevant discussion pertaining to the statistical ABM model fits using the Connecticut rabies data set as an example (Figure 1).

The rabies data set introduced in Section 1.2 was converted to a gridded areal unit format from a township areal unit format (as in Smith et al. 2002) for use with the proposed ABMs. The models proposed were presented with sufficient generality to model the data in their original format, and although such generalizations are the focus of ongoing research, one could foresee letting the neighborhood-based dispersal probability vectors ($\mathbf{p}$) vary in dimension over space. This would allow cells to have fewer or greater neighbors depending on their spatial arrangement. In addition, one could account for the effect of size and shape of neighbors on dispersal by weighting those dispersal probabilities appropriately in a fashion similar to that in which proximity matrixes are specified in spatial autoregressive models (e.g., Schabenberger and Gotway 2005, chap. 6).

Figure 1. Presence or absence of rabies in raccoon populations in Connecticut over 48 regularly spaced time periods beginning in 1991 (top left) and ending in 1995 (bottom right). Black cells denote the presence of rabies, and white cells denote the absence of rabies. For reference, the upper left edge of the images is New York state, and the lower right portion of the images (with diagonal shading) is the Long Island Sound.

With regard to the persistence of the process, we note that epidemic behavior is often characterized by a travelling epidemic wave in which, after infection of the local population, local extinction occurs and the propagating disease, having no more hosts in that area, subsides. To effectively capture the behavior in such cases, the persistence term (ϕ) becomes critical. In the case of the rabies data, without knowledge of subsidence, the data were transformed such that each infected area was assumed to persist. In this case, with rabies in a wild animal population in which the disease spread readily and was largely unmanaged at the time of data collection, the assumption of persistence is not unreasonable; however, such an assumption would not be appropriate in all epidemiologic situations and, in this case, over longer periods could be directly linked to remediation measures such as vaccination baits (Slate et al. 2005).

In either event, the persistence term (ϕ) should accommodate differing levels of dispersal.

The models from Sections 2.2 and 2.3 were implemented using MCMC via Gibbs and Metropolis–Hastings updates. The algorithms took on the order of hours to complete (on a Linux workstation with 3-GHz processors) where convergence in the MCMC samples was reached relatively quickly (i.e., <1000 iterations), and the remaining 19,000 realizations were thinned and then used to calculate posterior statistics. In what follows, we present only those results pertaining to the more general of the two models (i.e., the nonstationary anisotropic model) for the sake of space; however, we do provide details on the model comparison between the two fits to assess whether the more complicated model is warranted in this case.

2.4.1 Nonstationary Model Fit. The posterior mean propagation of probability of presence ($\theta_{i,t}$) for the nonstationary model appears similar to that of the stationary model (Figure 2). However, the posterior standard deviation (Figure 3) for probability of presence ($\theta_{i,t}$) indicates that the uncertainty in areas without data remains, but the uncertainty corresponding to the wave front of the epidemic is less than that resulting from the stationary model. This suggests that even though the pattern of uncertainty is similar, the nonstationary model more precisely estimates the neighborhood-based component of the probability of presence ($\theta_{i,t}$).

Figure 4 shows the covariates ($\mathbf{X}$) used in fitting the anisotropic nonstationary model from Section 2.3 to the Connecticut rabies data set given in Figure 1. All of the covariates used in this analysis were binary, with the black regions in Figure 4 corresponding to the occurrence of the environmental characteristic. Specifically, the X_1 and X_2 covariates correspond to the west and east sides of the Connecticut River, respectively. The X_3 covariate corresponds to land area in Connecticut (and some in New York on the western edge of the maps), which is not adjacent to the Connecticut River. The X_4 covariate corresponds to the Connecticut coastline where the state meets the Long Island Sound (diagonally shaded) in the bottom-right portion of the

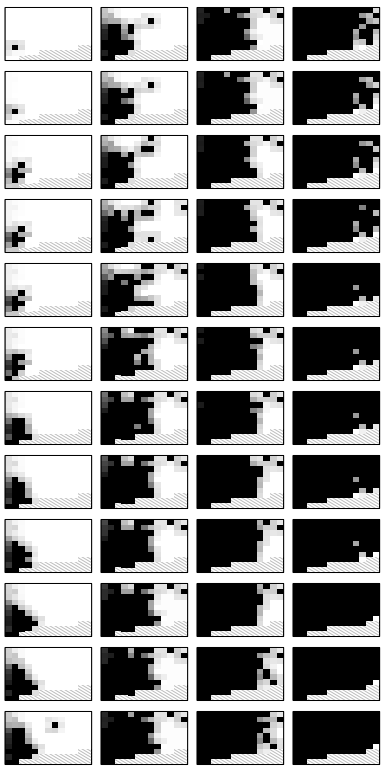

Figure 2. Posterior mean for probability of presence ($[\theta_{i,t}|\mathbf{y}]$) from nonstationary model in the same sequence as the data in Figure 1. Values of intensity range between 0 (white) and 1 (black).

Figure 3. Posterior standard deviation for probability of presence ($[\theta_{i,t}|\mathbf{y}]$) from the nonstationary model in the same sequence as the data in Figure 1. Values of intensity range between 0 (white) and 1 (black).

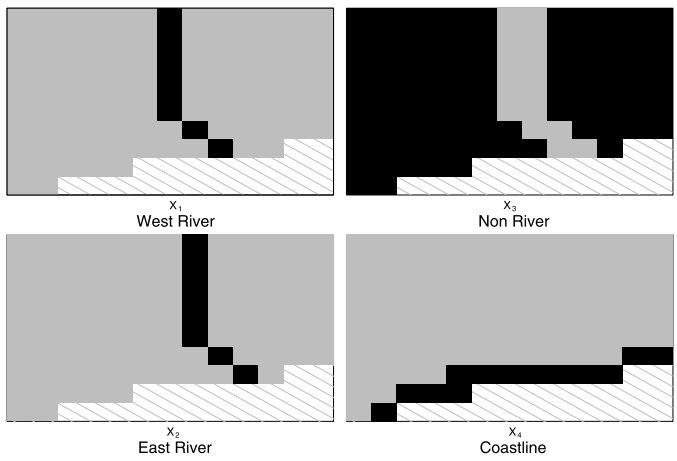

Figure 4. The indicator covariates used in the nonstationary model. $\mathbf{x}_1$ and $\mathbf{x}_2$ represent the west and east sides of the Connecticut River, respectively, while $\mathbf{x}_3$ represents all land area not adjacent to the Connecticut River and $\mathbf{x}_4$ represents the Connecticut coastline.

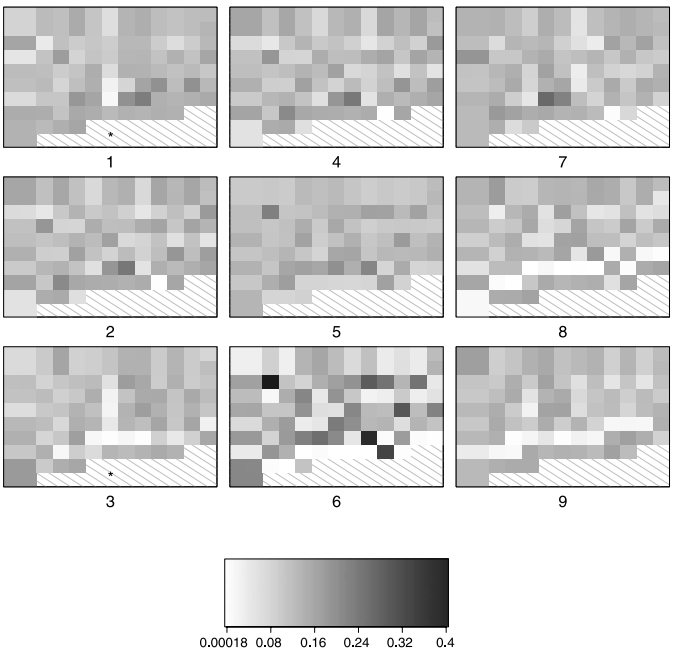

Figure 5. Posterior mean of [$\mathbf{P}|\mathbf{y}$] from the nonstationary model with respect to the Queen's neighborhood arrangement. The jth map refers to the posterior mean spatial field of the neighborhood-based transition probabilities in the jth direction. The small stars correspond to the column referenced in Section 2.4.1.

figures. A covariate corresponding to human population density may prove useful for future analyses, although it is important to note that major population centers in Connecticut coincide with the Connecticut River (X_1 and X_2) and coastline (X_4) covariates and thus is implicitly accounted for in $\mathbf{X}$.

The posterior distributions of the neighborhood-based dispersal for the anisotropic nonstationary model could technically be displayed separately; however, they would have to be shown for each spatial location i, because $\mathbf{p}_i$ now varies in space. Instead, consider the posterior mean of each spatial field of the neighborhood-based dispersal probabilities in Figure 5. The jth map in Figure 5 corresponds to the intensity in probability of the phenomenon (rabies in this case) dispersing in that direction, where the center plot (i.e., $j = 5$) is the probability of nondispersal. For example, the map in Figure 5 labeled "3" corresponds to the probability of neighborhood-based dispersal in the southwest direction given that the phenomenon is currently in the grid cell of interest.

The map of habitat preference ($\boldsymbol{\alpha}$) shown in Figure 6(a) corresponds to areas (shown as darker grid cells) in which the rabies epidemic spreads the fastest in the direction of more suitable habitat. The posterior standard deviation corresponding to habitat preference, shown in Figure 6(b), shows the location of areas of high uncertainty regarding habitat preference over the spatial domain. The posterior distributions for the covariate effect parameters ($\boldsymbol{\beta}$) are provided in Figure 7 and, when premultiplied by the covariates (X), correspond to the mean of habitat preference ($\boldsymbol{\alpha}$).

The maps corresponding to the spatial fields of the directional neighborhood-based dispersal parameters ($\mathbf{p}_j$) in Figure 5 indicate that there are indeed differences in directional dispersal depending on location. For example, in map 1 (northwest direction) and map 3 (southwest direction) of Figure 5, there is a distinct column (indicated by a small star in Figures 5 and 6) of low dispersal corresponding to the west side of the Connecticut River covariate (X_1 in Figure 4). From Figure 6(a), it is evident that the Connecticut river is an area of high habitat preference; this has been discussed in previous studies (e.g., Smith et al. 2002) and also is substantiated by the posterior distribution of $\boldsymbol{\beta}$ in Figure 7, where the river coefficients exceed 0. Thus, as the rabies epidemic spreads from the western region of Connecticut, it exhibits a different pattern of neighborhood-based dispersal than that seen in the other areas. Specifically, it shows a very low probability of dispersal in the northwest and southwest directions as it approaches the western edge of the Connecticut River, and a higher probability of dispersal in the north, south, northeast, and southeast directions. These results are generally in line with previous analyses of these data (e.g., Smith et al. 2002; Waller and Gotway 2004, chap. 9).

The posterior distributions for persistence (ϕ) and nonneighborhood dispersal (ψ) for the anisotropic nonstationary model were very similar to those of the stationary model. Specifically, the posterior probability of persistence resulting from fitting the nonstationary model is close to 1, suggesting that the rabies epidemic in Connecticut was very persistent during the period when these data were collected. The ramifications of this inference were discussed earlier. The posterior mean probability of nonneighborhood dispersal is estimated to be near 0.01,

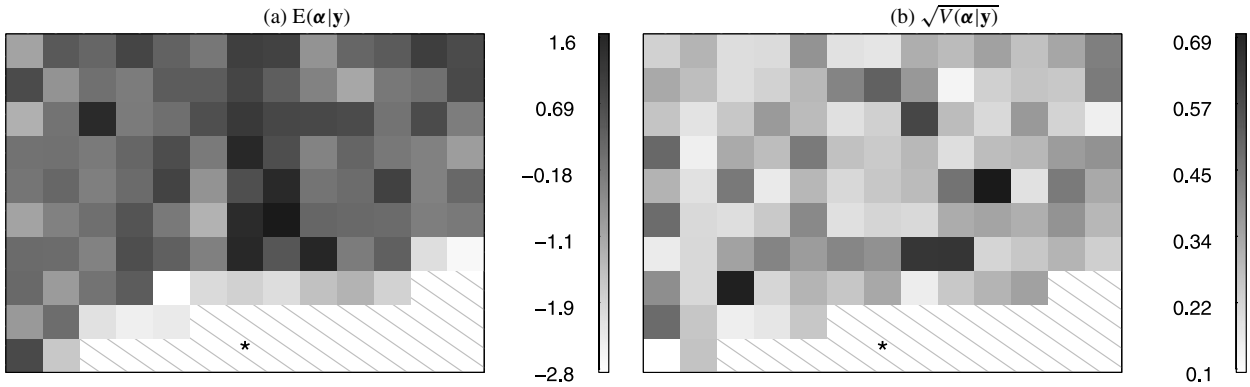

Figure 6. Posterior mean and standard deviation of habitat preference ($\boldsymbol{\alpha}$). Large values in the E($\boldsymbol{\alpha}|\mathbf{y}$) correspond to areas of higher suitability. The small star corresponds to the column referenced in Section 2.4.1.

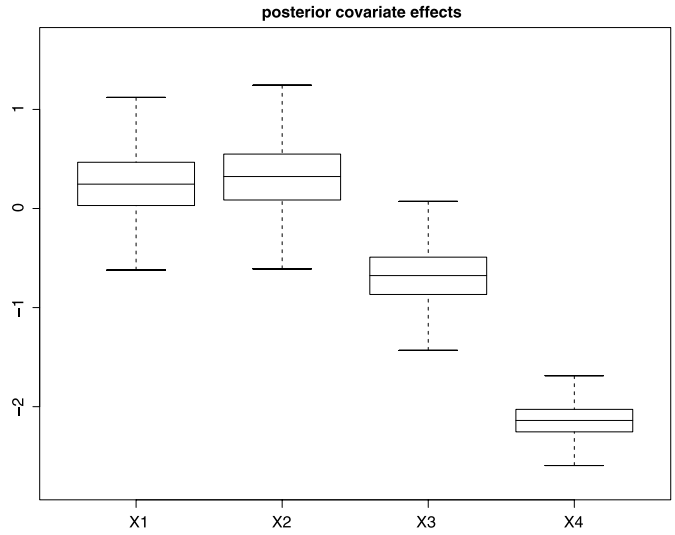

Figure 7. Posterior distributions of the covariate effects ($\boldsymbol{\beta}$), corresponding to the covariates in Figure 4. Positive β_j corresponds to a larger positive influence of the jth covariate on suitability $\boldsymbol{\alpha}$. This in turn will increase the probability of movement into a more suitable cell from a less suitable cell. $\mathbf{x}_1$ and $\mathbf{x}_2$ represent the west and east sides of the Connecticut River, respectively, while $\mathbf{x}_3$ represents all land area not adjacent to the Connecticut River and $\mathbf{x}_4$ represents the Connecticut coastline.

suggesting a 1% chance of long-distance rabies dispersal for nonneighboring areas.

The posterior mean of the scale parameter (σ_α^2) corresponding to the spatial covariance of $\boldsymbol{\alpha}$ was found to be 0.489 with a posterior standard deviation of 0.068, compared with the prior mean of 1 and standard deviation of 4. The posterior distribution of the spatial range parameter (θ_α) in the covariance of $\boldsymbol{\alpha}$ was found to be very close to 0 (suggesting very little, if any, residual spatial structure in the habitat preference covariance) and thus was removed from the analysis. Note that the full-conditional distributions for all other parameters retain the same form, only with θ_α set to 0.

Although we have only discussed the results from fitting the nonstationary model here, we used deviance information criterion (DIC) as a form of model comparison for the stationary and nonstationary models. In this case the DIC provides a measure of how well the top-level parameters in the hierarchy fit the data (Gelman et al. 2004). Here the top-level parameters (i.e., $\mathbf{p}$, ϕ, and ψ) are the most important because they provide insight into the dynamics of dispersal in spreading phenomena.

For the more complex anisotropic nonstationary model, the DIC was calculated to be 702.6, and the effective number of parameters was estimated as 1.92. In comparison, the effective number of parameters for the simpler stationary model was estimated to be 36.5, whereas the DIC value was 723.9. These results indicate that the nonstationary model, with a lower DIC, is the better-fitting model. On the other hand, parsimony is an important factor in model comparison, so there certainly is value in the stationary model, because it is more computationally efficient and retains more degrees of freedom. The nonstationary model, like the stationary model, has a dynamical component, but also has an explicit connection to scientifically meaningful environmental covariates and thus is capable of fitting the data better while performing inference on the connection between different types of dispersal and the heterogeneous environment over which the epidemic is dispersing.

3. COMPARISON WITH OTHER METHODS

3.1 Partial Differential Equation Models

The methods we present here for modeling spatiotemporal dynamic processes through an agent-based framework rely, at least in part, on short distance movement as characterized by a stochastic cellular automaton. That is, we are able to account for complicated spatiotemporal behavior in the dynamic process as a whole by starting with an agent-based construction founded on first principles. In this section we show that this underlying first-principle specification leads to a classical advection-diffusion partial differential equation when scaling up from a Lagrangian model to an Eulerian model.

For simplicity, consider the simple two-dimensional situation in which an organism or epidemic resides in a given cell and can move to any of the first-order neighboring cells or stay where it is with some probability (i.e., Rook's neighborhood). Now we seek a realistic model that can explain both nonstationary and anisotropic dynamical behavior. We let these movement probabilities $[\mathbf{p}(s_1, s_2) = \{p(s_1 - \Delta s, s_2), p(s_1, s_2 + \Delta s), p(s_1, s_2), p(s_1, s_2 - \Delta s), p(s_1 + \Delta s, s_2)\}]$ vary in both direction and location; however, they will remain independent from the dynamic process itself. This latter assumption is not necessary, but will simplify the derivation of our resulting Eulerian differential equation.

Following Turchin (1998), we first create a Lagrangian recurrence equation to describe our two-dimensional agent-based system in terms of probability of presence or occupancy, $\bar{p}(s_1, s_2, t)$, at cell location (s_1, s_2) and time t,

$$\begin{aligned}\bar{p}(s_1, s_2, t) = {} & p(s_1, s_2)\bar{p}(s_1, s_2, t - \Delta t) \\ & + p(s_1 + \Delta s, s_2)\bar{p}(s_1 - \Delta s, s_2, t - \Delta t) \\ & + p(s_1, s_2 - \Delta s)\bar{p}(s_1, s_2 + \Delta s, t - \Delta t) \\ & + p(s_1, s_2 + \Delta s)\bar{p}(s_1, s_2 - \Delta s, t - \Delta t) \\ & + p(s_1 - \Delta s, s_2)\bar{p}(s_1 + \Delta s, s_2, t - \Delta t), \quad (12)\end{aligned}$$

where Δs and Δt represent the distance between first-order neighbors and time steps, respectively. The recurrence equation (12) allows the probability of presence in a cell at the current time to be a function of the probability of movement from potentially occupied neighboring cells at the previous time.

Because (12) involves terms depending on $s_1 \pm \Delta s$, $s_2 \pm \Delta s$, and $t - \Delta t$, to arrive at a differential model, we need to extract them. This can be done by approximating $\bar{p}$ with a truncated Taylor series, then substituting the expansion approximations into the recurrence equation (12) and combining like terms to yield the following partial differential model:

$$\begin{aligned}0 = {} & -\Delta t \frac{\partial \bar{p}}{\partial t} + \Delta s \left(\frac{\partial \bar{p}}{\partial s_1}(p(s_1 - \Delta s, s_2) - p(s_1 + \Delta s, s_2)) \right. \\ & \left. + \frac{\partial \bar{p}}{\partial s_2}(p(s_1, s_2 - \Delta s) - p(s_1, s_2 + \Delta s)) \right)\end{aligned}$$

$$+\frac{\Delta s^2}{2}\left(\frac{\partial^2 \bar{p}}{\partial s_1\,\partial s_1}(p(s_1-\Delta s,s_2)+p(s_1+\Delta s,s_2))\right.$$
$$\left.+\frac{\partial^2 \bar{p}}{\partial s_2\,\partial s_2}(p(s_1,s_2-\Delta s)+p(s_1,s_2+\Delta s))\right), \qquad (13)$$

where all $\bar{p} \equiv \bar{p}(s_1, s_2, t)$. Now (13) simplifies to a partial differential equation (PDE) of the form

$$\frac{\partial \bar{p}}{\partial t} = -\frac{\partial \bar{p}\delta_1}{\partial s_1} - \frac{\partial \bar{p}\delta_2}{\partial s_2} + \frac{\partial^2 \bar{p} D_1}{\partial s_1\,\partial s_1} + \frac{\partial^2 \bar{p} D_2}{\partial s_2\,\partial s_2}, \qquad (14)$$

and is valid as long as the changes in time and space are sent to 0 in such a way that $\Delta s/\Delta t$ and $\Delta s^2/\Delta t$ are constant. This can be considered either a grid with cells becoming increasingly smaller and monitored at increasingly smaller increments or as cell sizes and monitoring stay constant but the spatial and temporal extents of the grid become increasingly large. Although admittedly, this limiting situation is somewhat contrived, it still provides a theoretical connection between the agent-based Lagrangian models discussed herein with the statistical implementation of Eulerian PDE models in other previous studies (e.g., Wikle 2003; Hooten and Wikle 2008). The particular Eulerian model in (14) contains both drift (δ) and dispersion (D) terms that are functions of the movement probabilities $\mathbf{p}$ and can vary by location.

An important note about the foregoing derivation is that if one retains additional higher-order terms in the Taylor series approximation, then more complex differential models for $\bar{p}$ will result (Turchin 1998). That is, even though agent-based models can be used to motivate the use of partial differential equations to describe natural dynamical processes, the Lagrangian framework is inherently less constrained. We see this as an advantage of using agent-based models, especially in situations with binary data on discrete spatial support and potentially nonstationary and anisotropic dynamics.

3.2 Generalized Linear and Markov Random Field Models

Another common class of models that can be used for binary data on discrete spatiotemporal support are generalized linear models (GLMs). Recall the likelihood specification that we introduced in Section 2.2: $y_{i,t}|\theta_{i,t} \sim \text{Bern}(\theta_{i,t})$. Now, using a simpler formulation that considers dynamics in time only, specify a model for $\theta_{i,t}$ on the logit scale rather than on the probability scale as in (3). In the setting of occupancy modeling, Royle and Dorazio (2008) showed that an alternative specification for the Bernoulli success probability is $\pi_{i,t} = \text{P}(y_{i,t} = 1|y_{i,t-1})$, where $\text{logit}(\pi_{i,t}) = a_t + b_t y_{i,t-1}$. The structure of this latter equation yields a temporal autologistic model in which a_t and b_t are reparameterizations of the parameters in the probability scale model. In this case, the logit transform serves as the link function in a GLM specification.

To accommodate static spatial structure in such a model, one also may consider the following spatial autologistic formulation: $\text{logit}(\pi_{i,t}) = a_t + b_t y_{i,t-1} + \beta c_i \sum_{j\in\mathcal{N}_i} y_{j,t}$, where, $1/c_i$ is the cardinality of $\mathcal{N}_i$ (Royle and Dorazio 2008). Now we have an auto-logistic model for occupancy that contains both time and space dependencies but is not spatiotemporally dynamic; that is, occupancy at location i during the current time depends only on neighboring occupancy at the current time and occupancy in the same location at the previous time. Zhu, Huang, and Wu (2005) also implemented a version of this model and discussed some of the difficulties, including well-posedness issues that result from potential nonstationarity in the spatial structure. A generalization would be to allow neighboring locations at the previous time to affect the occupancy at the current time at location i. In this vein, Royle and Dorazio (2008) suggested formulations for dynamic survival and colonization models sans implementation.

In both cases (i.e., Zhu, Huang, and Wu 2005 and Royle and Dorazio 2008), the covariates are accommodated in a static fashion; that is, they are linked directly to $y_{i,t}$. As described in Section 2.3, we are interested in characterizing directionally varying and spatially heterogeneous dynamics; thus we linked the directional transition probabilities in the neighborhood-based dispersal model to the directional derivatives of the suitability surface, which in turn depends on the covariates and, potentially, on additional spatial structure via $\boldsymbol{\varepsilon}$ in (7). It could be argued that the additional spatial structure induced through $\boldsymbol{\varepsilon}$ makes the ABM a more robust model, although in the case of the rabies example, we found no evidence of additional structure beyond that provided by the covariates. Our ABM specification also allows for potential long-distance dispersal and persistence that may be unrelated to the covariates. We believe that duplicating these features of our models would be challenging using autologistic GLM specifications, although new developments in autoregressive modeling, such as those described by Rue, Martino, and Chopin (2009), may such models more convenient to specify and fit in the future.

4. CONCLUSION

In this work, we have developed statistical ABMs for modeling a spreading phenomena over a partitioned landscape, given only binary data, and presented various generalizations along with connections to alternative methods. To illustrate the utility of these methods, we applied the models to the rabies epidemic in raccoon populations in Connecticut in 1991–1995. Both models take a probabilistic cellular automata approach. The first model is specified so that the focus is on estimating the dynamics governing the process in terms of stationary neighborhood-based dispersal. The second model uses environmental covariates to help determine how patterns of dispersal might differ given heterogeneous environmental conditions.

Although the simpler model (Section 2.2) assumes stationary dynamics, it can accommodate anisotropic dispersal behavior based on transition probabilities into neighboring areas. The generalized ABM, presented in Section 2.3, allows for anisotropic and nonstationary dynamics while accounting for associations between dispersal and environmental factors. Both models allow for the possibility of long-distance dispersal and varying levels of persistence, although in the rabies example the persistence parameter is not informative.

The general statistical agent-based modeling framework presented here could be extended to address other issues as well. For example, the assumption of perfect detection of the process under study is made in the specification of the model likelihood. Specifically, for the application presented here, there may

be some probability associated with the actual detection of rabies in a given area, suggesting that perhaps rabies is present in an area earlier than it was recorded. One approach to addressing such a situation is to allow for a detection probability parameter in the data model and then use a latent process model to model the "true" presence or absence of rabies in an area. Royle and Dorazio (2008) and Royle and Kery (2007) took this approach in a logit-transformed framework and were able to make use of repeated measurements to estimate such parameters. In the situation considered here, repeated measurements are not available, and with no other means to estimate detectability, such methods are futile. It could be argued, however, that by modeling the probability of presence in the manner adopted here, the posterior distribution of $\theta_{i,t}$ (e.g., Figure 2) absorbs the detection probability in $\theta_{i,t}$ at locations where rabies has yet to be observed. The framework presented here also could be extended to accommodate data with count support by extending the data model from a Bernoulli to a binomial, negative binomial, or Poisson.

Overall, the methods presented here place agent-based simulation models, such as those presented by Smith et al. (2002), into a rigorous statistical framework in which the parameters governing dispersal can be estimated and inference made about the effects of environmental covariates on dispersal. Furthermore, these methods also provide an additional way to study resource selection in terms of characterizing important factors governing the spread of epidemics (e.g., river corridors, shorelines, population centers) and could ultimately be used in policy making and management decisions. In fact, Hooten, Johnson, and Lowry (2010) discussed how similar approaches can be taken to connect individual-based models for animal movement to heterogeneous landscape information. Such individual-based models still fall under the ABM umbrella where, instead of expressing the agent in terms of an areal reporting unit, the animal itself is considered an agent. These models can be useful in linking individual behavior to environmental features as well as other individuals for the purposes of studying disease transmission. From a broader perspective, the statistical ABM allows for the specification of these formal linkages and the estimation of key components controlling the process under study.

APPENDIX A: FULL–CONDITIONALS

In the anisotropic stationary model proposed in Section 2.2, the full-conditional distributions were reported without proof. Consider now the derivation of $[\phi|\cdot]$:

$$
\begin{aligned}
[\phi|\cdot] &\propto \prod_{t=2}^{T} \prod_{i\in\mathcal{A}_{\phi_{t-1}}} \text{Bern}(y_{i,t}|\theta(\phi)_{i,t}) \times \text{Beta}(\alpha_\phi, \beta_\phi) \\
&\propto \left[\prod_{t=2}^{T} \prod_{i\in\mathcal{A}_{\phi_{t-1}}} \theta(\phi)_{i,t}^{y_{i,t}} (1-\theta(\phi)_{i,t})^{1-y_{i,t}}\right] \phi^{\alpha_\phi-1}(1-\phi)^{\beta_\phi-1} \\
&\propto \left[\prod_{t=2}^{T} \prod_{i\in\mathcal{A}_{\phi_{t-1}}} \phi^{y_{i,t}} (1-\phi)^{1-y_{i,t}}\right] \phi^{\alpha_\phi-1}(1-\phi)^{\beta_\phi-1} \\
&\propto \phi^{\sum_{t=2}^{T}\sum_{i\in\mathcal{A}_{\phi_{t-1}}} y_{i,t}} (1-\phi)^{\sum_{t=2}^{T}\sum_{i\in\mathcal{A}_{\phi_{t-1}}} (1-y_{i,t})} \\
&\quad \times \phi^{\alpha_\phi-1}(1-\phi)^{\beta_\phi-1} \\
&\propto \phi^{\sum_{t=2}^{T}\sum_{i\in\mathcal{A}_{\phi_{t-1}}} y_{i,t}+\alpha_\phi-1} \\
&\quad \times (1-\phi)^{\sum_{t=2}^{T}\sum_{i\in\mathcal{A}_{\phi_{t-1}}} (1-y_{i,t})+\beta_\phi-1} \\
\Rightarrow \quad \phi|\cdot &\sim \text{Beta}\left(\sum_{t=2}^{T}\sum_{i\in\mathcal{A}_{\phi_{t-1}}} y_{i,t}+\alpha_\phi, \right. \\
&\qquad \left. \sum_{t=2}^{T}\sum_{i\in\mathcal{A}_{\phi_{t-1}}} (1-y_{i,t})+\beta_\phi\right)
\end{aligned}
$$

where the index set $\mathcal{A}_{\phi_{t-1}}$ is defined as in Section 2.2. The full-conditional for ψ can be derived in the same manner where ϕ in the foregoing equations is replaced with ψ, and the index set $\mathcal{A}_{\psi_{t-1}}$ is used instead of $\mathcal{A}_{\phi_{t-1}}$.

The parameters $\mathbf{p}$ and $\mathbf{a}$ are sampled from the full-conditionals via Metropolis–Hastings,

$$
\begin{aligned}
\mathbf{p}|\cdot \sim [\mathbf{p}|\cdot] &\propto \prod_{t=2}^{T} \prod_{i\in\mathcal{A}_{\bar{p}_t}} \text{Bern}(y_{i,t}|\bar{p}_{i,t}) \times \text{Dir}(\mathbf{p}|\mathbf{a}), \\
\mathbf{a}|\cdot \sim [\mathbf{a}|\cdot] &\propto \text{Dir}(\mathbf{p}|\mathbf{a}) \times \text{N}(\log(\mathbf{a})|\boldsymbol{\mu}_a, \sigma_a^2\mathbf{I}),
\end{aligned}
$$

where the proposal $\mathbf{p}^*$ is sampled from a Dirichlet distribution given the current Gibbs sample $\mathbf{a}^{(k)}$. The proposed parameter vector is then accepted with probability

$$
\frac{\prod_{t=2}^{T}\prod_{i\in\mathcal{A}_{\bar{p}_i}} \text{Bern}(y_{i,t}|\bar{p}^*_{i,t})}{\prod_{t=2}^{T}\prod_{i\in\mathcal{A}_{\bar{p}_i}} \text{Bern}(y_{i,t}|\bar{p}^{(k-1)}_{i,t})},
$$

where the index set $\mathcal{A}_{\bar{p}_t}$ is defined as in Section 2.2. The proposal $\mathbf{a}^*$ is sampled via a Gaussian random walk [i.e., $\mathbf{a}^* \sim \text{N}(\mathbf{a}^{(k-1)}, \sigma^2_{\text{tune}}\mathbf{I})$]. Because the proposal distribution is symmetric with respect to $\mathbf{a}^{(k-1)}$ and $\mathbf{a}^*$, the proposed vector can be accepted with probability

$$
\frac{\text{Dir}(\mathbf{p}^{(k)}|\mathbf{a}^*) \times \text{N}(\log(\mathbf{a}^*)|\boldsymbol{\mu}_a, \sigma_a^2\mathbf{I})}{\text{Dir}(\mathbf{p}^{(k)}|\mathbf{a}^{(k-1)}) \times \text{N}(\log(\mathbf{a}^{(k-1)})|\boldsymbol{\mu}_a, \sigma_a^2\mathbf{I})}.
$$

In the Gibbs sampler for the anisotropic nonstationary model proposed in Section 2.3, the full-conditional distributions for ϕ and ψ can be derived in the same manner as before. Then a proposal for all $\mathbf{p}_i$ simultaneously is obtained by sampling $\mathbf{p}^*_i \sim \text{Dir}(\mathbf{a}^{(k)}_i)$ for all $i = 1, \ldots, m$, thus yielding $\mathbf{P}^* = [\mathbf{p}^*_1, \ldots, \mathbf{p}^*_m]$, and accepted with probability

$$
\frac{\prod_{t=2}^{T}\prod_{i\in\mathcal{A}_{\bar{p}_i}} \text{Bern}(y_{i,t}|\bar{p}^*_{i,t})}{\prod_{t=2}^{T}\prod_{i\in\mathcal{A}_{\bar{p}_i}} \text{Bern}(y_{i,t}|\bar{p}^{(k-1)}_{i,t})}.
$$

Proposals for $\boldsymbol{\alpha}$ can be taken from a random walk, $\boldsymbol{\alpha}^* \sim \text{N}(\boldsymbol{\alpha}^{(k-1)}, \sigma^2_{\text{tune}}\mathbf{I})$, and accepted with probability

$$
\frac{\prod_{i=1}^{m} \text{Dir}(\mathbf{p}^{(k)}_i|\mathbf{a}(\boldsymbol{\alpha})^*_i) \times \text{N}(\boldsymbol{\alpha}^*|\boldsymbol{\beta}^{(k)}, (\sigma^2_\alpha)^{(k)}, \theta^{(k)}_\alpha)}{\prod_{i=1}^{m} \text{Dir}(\mathbf{p}^{(k)}_i|\mathbf{a}(\boldsymbol{\alpha})^{(k-1)}_i) \times \text{N}(\boldsymbol{\alpha}^{(k-1)}|\boldsymbol{\beta}^{(k)}, (\sigma^2_\alpha)^{(k)}, \theta^{(k)}_\alpha)}.
$$

The parameter θ_α also can be sampled via Metropolis–Hastings with the random-walk proposal, $\theta^*_\alpha \sim \text{N}(\theta^{(k-1)}_\alpha, \sigma^2_{\text{tune}})$ and accepted with probability

$$
\frac{\text{N}(\boldsymbol{\alpha}^{(k)}|\boldsymbol{\beta}^{(k)}, (\sigma^2_\alpha)^{(k)}, \theta^*_\alpha) \times \text{Gamma}(\theta^*_\alpha|\alpha_\theta, \beta_\theta)}{\text{N}(\boldsymbol{\alpha}^{(k)}|\boldsymbol{\beta}^{(k)}, (\sigma^2_\alpha)^{(k)}, \theta^{(k-1)}_\alpha) \times \text{Gamma}(\theta^{(k-1)}_\alpha|\alpha_\theta, \beta_\theta)}.
$$

Full-conditional distributions for the remaining parameters (i.e., $\boldsymbol{\beta}$, $\boldsymbol{\Sigma}_\beta$, and σ^2_α) are conjugate and can be trivially derived in the fashion described by Gelman et al. (2004).

APPENDIX B: HYPERPRIOR SPECIFICATIONS

The tables in this appendix provide the hyperprior specifications used in fitting these models.

Table B.1. Hyperparameters used in stationary model from Section 2.2

Hyperparameter	Value
$\mu_{a,i}$	1
$\mathbf{\Sigma}_a$	$1 \times \mathbf{I}$
α_ϕ	10
β_ϕ	1
α_ψ	1
β_ψ	10

Table B.2. Hyperparameters used in nonstationary model from Section 2.3

Hyperparameter	Value
$\mu_{\beta,i}$	0
$\mathbf{S}_\beta$	$0.5 \times \mathbf{I}$
ν	50
r	0.8
q	2.25
μ_θ	1
σ^2_θ	1
α_ϕ	10
β_ϕ	1
α_ψ	1
β_ψ	10

[Received January 2009. Revised May 2009.]

REFERENCES

Banerjee, S., Carlin, B. P., and Gelfand, A. E. (2004), *Hierarchical Modeling and Analysis for Spatial Data*, New York: Chapman & Hall/CRC. [236]

Besag, J. (1974), "Spatial Interaction and the Statistical Analysis of Lattice Systems," *Journal of the Royal Statistics Society, Ser. B*, 36, 192–236. [236]

Clark, J. S. (2003), "Uncertainty and Variability in Demography and Population Growth: A Hierarchical Approach," *Ecology*, 84, 1370–1381. [238]

Cressie, N. A. C. (1993), *Statistics for Spatial Data: Revised Edition*, New York: Wiley. [236,237]

Durrett, R., and Levin, S. (1994), "Stochastic Spatial Models: A User's Guide to Ecological Applications," *Philosophical Transactions of the Royal Society, Ser. B*, 343, 329–350. [238]

Gelfand, A. E., Silander, J. A., Wu, S., Latimer, A., Lewis, P. O., Rebelo, A. G., and Holder, M. (2006), "Explaining Species Distribution Patterns Though Hierarchical Modeling," *Bayesian Analysis*, 1, 41–91. [240]

Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (2004), *Bayesian Data Analysis* (2nd ed.), Boca Raton: Chapman & Hall/CRC. [244,246]

Grimm, V., and Railsback, S. F. (2005), *Individual-Based Modeling and Ecology*, Princeton, NJ: Princeton University Press. [236]

Grimm, V., Revilla, E., Berger, U., Jeltsch, F., Mooij, W. M., Railsback, S. F., Thulke, H.-H., Weiner, J., Wiegand, T., and DeAngelis, D. L. (2005), "Pattern-Oriented Modeling of Agent-Based Complex Systems: Lessons From Ecology," *Science*, 310, 987–991. [237]

Hanski, I. (1999), *Metapopulation Ecology*, New York: Oxford University Press. [236]

Heikkinen, J., and Hogmander, H. (1994), "Fully Bayesian Approach to Image Restoration With an Application in Biogeography," *Applied Statistics*, 43, 569–582. [236]

Hirzel, A. H., Hausser, J., Chessel, D., and Perrin, N. (2002), "Ecological-Niche Factor Analysis: How to Compute Habitat-Suitability Maps Without Absence Data?" *Ecology*, 83, 2027–2036. [240]

Hoeting, J. A., Leecaster, M., and Bowden, D. (2000), "An Improved Model for Spatially Correlated Binary Responses," *Journal of Agricultural, Biological, and Environmental Statistics*, 5, 102–114. [236]

Hogeweg, P. (1988), "Cellular Automata as a Paradigm for Ecological Modeling," *Applied Mathematics and Computation*, 27, 81–100. [237]

Hogmander, H., and Moller, J. (1995), "Estimating Distribution Maps From Atlas Data Using Methods of Statistical Image Analysis," *Biometrics*, 51, 393–404. [236]

Hooten, M. B., and Wikle, C. K. (2008), "A Hierarchical Bayesian Non-Linear Spatio-Temporal Model for the Spread of Invasive Species With Application to the Eurasian Collared-Dove," *Environmental and Ecological Statistics*, 15, 59–70. [236,238,245]

Hooten, M. B., Johnson, D. S., and Lowry, J. H. (2010), "Agent-Based Inference for Animal Movement and Selection," to appear. [246]

Hooten, M. B., Larsen, D. R., and Wikle, C. K. (2003), "Predicting the Spatial Distribution of Ground Flora on Large Domains Using a Hierarchical Bayesian Model," *Landscape Ecology*, 18, 487–502. [240]

Hooten, M. B., Wikle, C. K., Dorazio, R. M., and Royle, J. A. (2007), "Hierarchical Spatio-Temporal Matrix Models for Characterizing Invasions," *Biometrics*, 63, 558–567. [236,239]

Krebs, C. J. (1978), *Ecology: The Experimental Analysis of Distribution and Abundance*, New York: Harper & Row Publishers. [236]

Lee, Y. C., Qian, S., Jones, R. D., Barnes, C. W., Flake, G. W., O'Rourke, M. K., Lee, K., Chen, H. H., Sun, G. Z., Zhang, Y. Q., and Chen, D. (1990), "Adaptive Stochastic Cellular Automata: Theory," *Physica D*, 45, 159–180. [237]

MacKenzie, D. I., Nichols, J. D., Hines, J. E., Knutson, M. G., and Franklin, A. B. (2003), "Estimating Site Occupancy, Colonization, and Local Extinction When a Species Is Detected Imperfectly," *Ecology*, 84, 2200–2207. [238]

MacKenzie, D. I., Nichols, J. D., Lachman, G. B., Droege, S., Royle, J. A., and Langtimm, C. A. (2002), "Estimating Site Occupancy Rates When Detection Probabilities Are Less Than One," *Ecology*, 83, 2248–2255. [238]

Manly, B., McDonald, L., Thomas, D., McDonald, T., and Erikson, W. (2002), *Resource Selection by Animals* (2nd ed.), Boston: Kluwer Academic. [240]

Nettles, V. F., Shaddock, J. H., Sikes, R. K., and Reyes, C. R. (1979), "Rabies in Translocated Raccoons," *American Journal of Public Health*, 69, 601–602. [237]

Pace, R. K., Barry, R., Gilley, O. W., and Sirmans, C. F. (2000), "A Method for Spatial–Temporal Forecasting With an Application to Real Estate Prices," *International Journal of Forecasting*, 16, 229–246. [236]

Royle, J. A., and Dorazio, R. M. (2008), *Hierarchical Modeling and Inference in Ecology: The Analysis of Data From Populations, Metapopulations and Communities*, Amsterdam, The Netherlands: Academic Press. [236,238,245,246]

——— (2006), "Hierarchical Models of Animal Abundance and Occurrence," *Journal of Agricultural, Biological, and Environmental Statistics*, 11, 249–263. [236]

Royle, J. A., and Kery, M. (2007), "A Bayesian State–Space Formulation of Dynamic Occupancy Models," *Ecology*, 88, 1813–1823. [236,246]

Rue, H., Martino, S., and Chopin, N. (2009), "Approximate Bayesian Inference for Latent Gaussian Models by Using Integrated Nested Laplace Approximations," *Journal of the Royal Statistical Society, Ser. B*, 71, 319–392. [245]

Schabenberger, O., and Gotway, C. A. (2005), *Statistical Methods for Spatial Data Analysis*, New York: Chapman & Hall/CRC. [241]

Slate, D., Rupprecht, C. E., Rooney, J. A., Donovan, D., Lein, D. H., and Chipman, R. B. (2005), "Status of Oral Rabies Vaccination in Wild Carnivores in the United States," *Virus Research*, 111, 68–76. [241]

Smith, D. L., Lucey, B., Waller, L. A., Childs, J. E., and Real, L. A. (2002), "Predicting the Spatial Dynamics of Rabies Epidemics on Heterogeneous Landscapes," *Proceedings of the National Academy of Sciences of the United States of America*, 99, 3668–3672. [237,241,243,246]

Turchin, P. (1998), *Quantitative Analysis of Movement*, Sunderland, MA: Sinauer Associates. [238,244,245]

Waller, L. A., and Gotway, C. A. (2004), *Applied Spatial Statistics for Public Health Data*, New Jersey: Wiley. [236,243]

Wheeler, D. C., and Waller, L. A. (2008), "Mountains, Valleys, and Rivers: The Transmission of Raccoon Rabies Over a Heterogeneous Landscape," *Journal of Agricultural, Biological, and Environmental Statistics*, 13, 388–406. [237]

Wikle, C. K. (2003), "Hierarchical Bayesian Models for Predicting the Spread of Ecological Processes," *Ecology*, 84, 1382–1394. [237,245]

Wikle, C. K., Milliff, R. F., Nychka, D., and Berliner, L. M. (2001), "Spatiotemporal Hierarchical Bayesian Modeling: Tropical Ocean Surface Winds," *Journal of the American Statistical Association*, 96, 382–397. [237]

Wolfram, S. (1983), "Statistical Mechanics of Cellular Automata," *Reviews of Modern Physics*, 55, 601–644. [237]

——— (1984), "Cellular Automata as Models for Complexity," *Nature*, 311, 419. [237]

——— (1988), "Cellular Automaton Supercomputing," in *High Speed Computing: Scientific Applications and Algorithm Design*, Champaign, IL: University of Illinois Press. [237]

Zhu, J., Huang, H.-C., and Wu, C.-T. (2005), "Modeling Spatial-Temporal Binary Data Using Markov Random Fields," *Journal of Agricultural, Biological, and Environmental Statistics*, 10, 212–225. [236,245]