# Invasions, Epidemics, and Binary Data in a Cellular World

Mevin B. Hooten
Department of Mathematics and Statistics
Utah State University

Christopher K. Wikle
Department of Statistics
University of Missouri

**Abstract**

Hierarchical models for estimating and predicting spatio-temporal ecological processes have been useful in situations where large-scale monitoring projects provide data in the form of counts. It is often the case in ecological studies, however, that it is only feasible to collect binary (i.e., presence / absence) data on large spatial and temporal domains. In such settings, the propagation of underlying ecological phenomena may still behave in a well-defined theoretical fashion (e.g., according to IDE- or PDE-based dynamics), though limited information prohibits the implementation of conventional spatio-temporal process models. When there is simply not enough information to inform such scientifically based parameterizations, we propose the use of a scientifically naive, rule-based dynamic process model within a hierachical framework. Utilizing only simple model specifications can still accomodate multiple sources of uncertainty (e.g., observational and model misspecification error) yet exhibit complicated large-scale dynamical behavior for characterizing the spread of invasive species in the presence of binary data.

KEY WORDS: Cellular Automata, Spatio-Temporal Models, Dynamical Systems, Hierarchical Models

## 1. Introduction

Many types of ecological, environmental, and epidemiological data are collected over discrete spatial and temporal domains (Krebs 1978; Hanski 1999; Waller and Gotway 2004). Moreover, such data are often binary valued. For example, consider a natural process $\mathcal{Z}$ 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 $\mathcal{Z}$ 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.

Models such as those proposed by Hooten and Wikle (2007) can be employed in cases where the data ($y_{i,t}$) are counts on domains with areal spatial support; whereas those proposed by Hooten et al. (2007) can be employed in cases where count data occur in continuous space. Both types of models previously discussed 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 priors). The focus in such situations shifts to the estimation of dispersal-based dynamics.

Conventional methods for modeling spatial processes on partitioned domains include conditional and simultaneously specified spatial Gaussian models (also known as conditional autoregressive models and simultaneous autoregressive models) which have now been extended to the spatio-temporal setting (see Cressie 1993 and Banerjee et al. 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 et al. 2000), where the probability of presence is conditioned on its neighbors. More recently, methods have been developed that extend the auto-logistic model, as well as other models using a logit transform of presence/absence probability, to spatio-temporal settings (e.g., Zhu et al. 2005; Royle and Kery 2006; Royle 2006).

Many previously developed methods address the problem 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 employed. Despite its ubiquitous appearance in the literature, the logit transform is not the most intuitive nor robust way to characterize changes in probability. In fact, the main reason for its use is the ability to use Gaussian model error (e.g., logistic regression). Addressing the problem more directly and modeling the untransformed probability components of presence and absence allows for the characterization of the dynamics of a propagating phenomenon over partitioned spatial and temporal domains.

The methods that follow introduce a class of models that are motivated by viewing the movement (i.e., dis-

persal) of a phenomenon from the perspective of the phenomenon itself, rather than the system as a whole. For example, a species of organism, such as an exotic invasive species or a pathogenic species, in a new environment will often move from areas of lower quality to areas of higher quality (quality 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. Such "automata" can act in their own manner given the state of their surrounding environment.

The properties and behavior of automata have been studied in nearly every field and are so powerful that they have been hypothesized to form the underpinnings of life itself (Wolfram 1984). In essence, an automata is defined as an entity, whose movement (movement can 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 "life-like" systematic behavior. In other words, the combined behavior of numerous individual automata results in very complicated dynamical system behavior incapable of being described in any other way (Wolfram 1983).

Automata are most often defined in a deterministic dynamical system framework where the state of the "neighborhood" of an entity (if an areal unit of space, called a "cell") determines the future state of the entity. Another type of automata can be formulated probabilistically, where movement of an entity into its neighborhood is defined by parametric probability distributions and the behavior of the system as a whole (i.e., the "automaton") 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 wave-like behavior commonly found in natural phenomena (Hogeweg 1988). Using more traditional terminology found in the spatial statistics literature, "spatially irregular" refers to spatial structure that is either anisotropic (i.e., varying directionally) or non-stationary (i.e., varying locationally) or both (e.g., Cressie 1993).

This fundamental difference in the construction of models is known as the "top-down" versus "bottom-up" approach (Grimm et al. 2005). Traditionally, bottom-up approaches to studying ecosystem function and ecological processes have been used in simulation settings whereas top-down approaches have been taken using statistical methodology. Both have contributed much to ecology in theory and application, though they rarely intermingle.

### 1.1 Application: Rabies Epidemics in Raccoon Populations

An example of a natural spatio-temporal process where binary data is available on a partitioned spatial domain and exhibiting irregular spreading behavior is the ongoing rabies epidemic on the east coast of the United States of America that began in the mid-1970's (Nettles et al. 1979). Specifically, space-time data documenting the spread of rabies through the raccoon population in Connecticut is available as "time since arrival in township" format (e.g., Smith et al. 2002). That is, after appearing first in 1991 in the western townships of Connecticut (those bordering New York), rabies spread throughout the state over the following five year period. Although epidemics are not typically thought of as invasive species, they propagate through populations of organisms in space and time in a similar manner as non-epidemiological invasions propagate in space and time.

The dataset shown in Figure 1 illustrates the spatio-temporal behavior of rabies epidemic in Connecticut from 1991–1995, where several long-distance dispersal events occurred in addition to anisotropic and non-stationary neighborhood-based dispersal patterns. Smith et al. (2002) employed a bottom-up simulation model in the spirit of an automatous system, though focusing on the temporal domain by characterization of rates of spread from neighboring townships, and found that such models can be useful in studying this particular natural phenomenon. Moreover, they allowed for long-distance dispersal, 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.

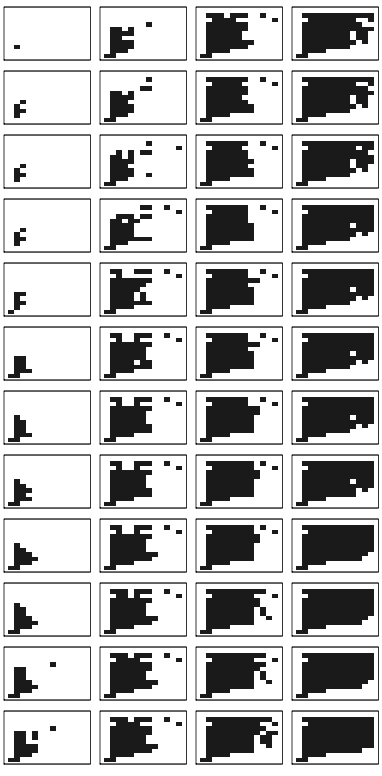


Figure 1: Presence / Absence of rabies in raccoon populations in Connecticut over 48 regularly spaced time periods beginning in 1991 (top left) and ending in 1995 (bottom right). Black cells denote presence while white cells denote absence of rabies.

The models developed in the next section approach the problem from the same perspective as Smith et al. (2002), but in a rigorous probability framework where inference on the dispersal parameters can be made from a statistical perspective given the available data.

## 2. Methods

### 2.1 CA-Based Dynamics

A dynamical system known as a cellular automaton (CA) can also 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 state of which is 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 (e.g., parallel processing and cluster computing settings; Wolfram 1988).

In a deterministic CA, the current state ($\mathbf{u}_t$) is exactly mapped to the previous state ($\mathbf{u}_{t-1}$) by a function $h$, so that the state of the automaton $\mathbf{u}_t$ is unique given $\mathbf{u}_{t-1}$ and $h$. More specifically, each element $u_{i,t}$ (the automata) is exactly mapped to the previous set of neighboring automata $u_{\mathcal{N}_i,t-1}$ by the function $h$. The function $h$ is often made up of a set of rules, defining the map from all possible states of the set $u_{\mathcal{N}_i,t-1}$ 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 utilization of deterministic automatons from a statistical modeling perspective is challenging, however, because the rule space quickly becomes cumbersome as the size of the neighborhood inceases, 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 creating a rule that is 8 dimensional; however, using the same neighborhood and one-dimensional spatial domain, there are $2^{(2^3)}$ possible rules. In general, for a neighborhood of dimension $d_{\mathcal{N}}$, there will be $2^{(2^{d_{\mathcal{N}}})}$ possible rules. 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 $1.34 \times 10^{154}$. Searching the space of rules for CA systems with large neighborhoods is not feasible, therefore it is necessary to either simplify the space of rules 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:

$$u_{i,t}|u_{\mathcal{N}_i,t-1} \sim h(u_{\mathcal{N}_i,t-1}), \tag{1}$$

whereas, the evolution equation of a deterministic automata is written:

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

In the case of the probabilistic automata, the function $h$ is a probability distribution. For the binary automaton (i.e., $u_{i,t} \in \{0, 1\}$), the probability distribution $h$ could be a Bernoulli distribution with the Bernoulli probability a function of the neighborhood at the previous time and some parameters. Notice that the specification in (1) is only different from Gaussian state-space models 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 2007).

The advantage of the probabilistic specification (1) over the determinstic 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.

### 2.2 Anisotropic Stationary Dispersal on Homogeneous Discrete Domains

One approach to modeling a spatio-temporal process such as the rabies process discussed in Section 1.1 is to estimate the dynamics of spread in terms of probability of presence. That is, assume the phenomenon in question is present ($y_{i,t} = 1$) at the spatial location $s_i$ and time $t$ with probability $\theta_{i,t}$ and absent ($y_{i,t} = 0$) with probability $1 - \theta_{i,t}$. This is common form for the classes of models known as occurrence and occupancy models (e.g., MacKenzie et al. 2002, 2003; Royle and Dorazio 2006). A more convenient notation for this data model is: $y_{i,t}|\theta_{i,t} \sim \text{Bern}(\theta_{i,t})$. Rather than model the transformed probability (either as a logit or probit transform) at the next level in the hierarchy, consider the following specification for $\theta_{i,t}$:

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

where,

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

and, $\mathcal{N}_i$ is the spatial neighborhood of area $i$. In a regularly gridded 2-dimensional domain with a "Queen's

neighborhood,"

$$\mathcal{N}_i = \begin{bmatrix} \mathrm{N}_{j=i-4} & \mathrm{N}_{j=i-1} & \mathrm{N}_{j=i+2} \\ \mathrm{N}_{j=i-3} & \mathrm{N}_{j=i} & \mathrm{N}_{j=i+3} \\ \mathrm{N}_{j=i-2} & \mathrm{N}_{j=i+1} & \mathrm{N}_{j=i+4} \end{bmatrix} = \begin{bmatrix} \mathrm{N}_1 & \mathrm{N}_4 & \mathrm{N}_7 \\ \mathrm{N}_2 & \mathrm{N}_5 & \mathrm{N}_8 \\ \mathrm{N}_3 & \mathrm{N}_6 & \mathrm{N}_9 \end{bmatrix} \tag{4}$$

In (3) the probability $\phi$ corresponds to persistence of the phenomena; that is, once the phenomena is present in spatial area $i$, the probability it will be present at the next time step is $\phi$. The probability $\psi$ corresponds to out of neighborhood dispersal; that is, the phenomena will occur in an area outside of those occupied with probability $\psi$. 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 $\psi$ 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. Many possible specifications for $\bar{p}_{i,t}$ are possible depending on the spatial domain, level of realistic detail desired in the model, and estimability of $\bar{p}_{i,t}$. Assuming a regularly gridded spatial domain and Queen's neighborhood dispersal, let,

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

where, $\mathbf{y}_{\mathcal{N}_i,t-1} = [y_{N_{i+4},t-1}, \ldots, y_{i,t-1}, \ldots, y_{N_{i-4},t-1}]'$ is a vector of the data (i.e., ones and zeros) corresponding to the neighborhood of the $i^{\text{th}}$ spatial area. The probabilities in $\mathbf{p}$ ($\mathbf{p} = [p_1, p_2, \ldots, p_{d_\mathcal{N}})]$'), where $d_\mathcal{N}$ is the dimensionality of the neighborhood, refer to the transition probability from spatial area $i$ to its $j^{\text{th}}$ neighbor. 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:

$$\begin{aligned} &\mathrm{P}(\cup_j E_{N_j,t};\ \text{for all occupied neighbors } j) = \\ &\quad = 1 - \mathrm{P}(\cap_j E^c_{N_j,t};\ \text{for all occupied neighbors } j) \\ &\quad = 1 - \prod_j \mathrm{P}(E^c_{N_j,t}) \\ &\quad = 1 - \prod_j (1 - \mathrm{P}(E_{N_j,t})) \\ &\quad = 1 - \exp(\sum_j \log(1 - \mathrm{P}(E_{N_j,t}))), \end{aligned}$$

where the last equality is shown only for computational efficiency and 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 ($E_{N_j,t}$) which would be required to calculate the probability of the union.

Assuming the transition probabilities in $\mathbf{p}$ sum to one and 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 $\log(\mathbf{a}) \sim \mathrm{N}(\boldsymbol{\mu}_a, \boldsymbol{\Sigma}_a)$.

It is important to point out several consequences concerning the effect of probabilities, $\mathbf{p}$, on the propagating phenomenon:

- This specification assumes that the neighborhood-based dispersal of the process is stationary. However, because the process also depends on non-neighborhood dispersal and persistence, it appears non-stationary; though, because of the exchangeable specification for $\phi$ and $\psi$ in space and time, it could be argued that the overall process model is still stationary in the sense that given the neighborhood-based dispersal, the probabilities for non-neighborhood dispersal and persistence are uniformly distributed in space and time.

- By allowing $\mathbf{p}$ to be a vector, the neighborhood-based dispersal model is explicitly anisotropic. A simpler form of dispersal would have a phenomena propagating to all neighboring areas with an equal probability. Thus, in the case with a regularly gridded spatial domain and Queen's neighborhood, the multivariate nature of $\mathbf{p}$ allows for differing rates of dispersal in different directions (8 directions in this case).

- 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 equal to 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, since the probability of persistence is being modeled in $\theta$, by $\phi$, 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 down the propagating phenomena without affecting the anisotropy of the process, hence providing the model with additional flexibility.

Priors for $\phi$ and $\psi$ 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

$[\psi] = \text{Beta}(\alpha_\psi, \beta_\psi)$ 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:

$$[\mathbf{p}, \mathbf{a}, \phi, \psi | \{y_{i,t}, \text{ for } t = 1, \ldots, T \text{ and } i = 1, ..., m\}] \propto$$
$$\propto \prod_{t=1}^{T}\prod_{i=1}^{m}[y_{i,t}|\theta(\mathbf{p}, \phi, \psi)_{i,t}][\mathbf{p}|\mathbf{a}][\mathbf{a}][\phi][\psi] \quad (6)$$

It can be shown that the full-conditional distributions for the model parameters, necessary for implementation in an MCMC setting, are:

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

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

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

where, the $\mathcal{A}$ are sets of indices defined by:

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

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

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

### 2.3 Anisotropic Non-Stationary Dispersal on Heterogeneous Discrete Domains

The model in the previous section is very robust in that it allows for anisotropic dynamic behavior in the process which is capable of dispersing at varying rates. Like the matrix model utilized 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 covariance 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 regards to the successful propagation of the phenomenon under study. For example, if the phenomenon were a wildfire, it might "prefer" an exposed topography with steep slopes; whereas if it were a songbird, it might "prefer" forest edges.

Distinguishing areas of suitable or desirable habitat from those that are unsuitable is a long studied subject in ecology. One of the simplest and most common approaches being taken to map habitat suitability is to find the associations between organism abundance (or presence) and environmental covariates (e.g., Hooten et al. 2003). Due to limitations imposed though 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, the terms "suitability" and "preference" are used interchangeably. Consider the suitability ($\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \cdots, \alpha_m]'$) of a spatial domain, partitioned into $m$ areas, in terms of phenomenon preference. 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 $f$, parameters $\boldsymbol{\beta}$, and some error structure $\boldsymbol{\varepsilon}$ accounting for any unknown covariates and (or) error related to the choice of $f$. Assuming a linear function $f$, and Gaussian error $\boldsymbol{\varepsilon}$, the standard linear model follows:

$$\boldsymbol{\alpha} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \quad (7)$$

Making the assumption that the error contains residual spatial dependence, and letting $\boldsymbol{\varepsilon} \sim \text{N}(\mathbf{0}, \boldsymbol{\Sigma}_\alpha)$, 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 areas, 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 $\boldsymbol{\Sigma}_\alpha$.

Since the propagating phenomena are attracted to areas of higher suitability, an "attraction" model is useful to explain probability. 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 phenom-

ena is present at location $i$, the probability that it will propagate to location $j$ depends, at least partially, 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 regards 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 2-D 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 ($\boldsymbol{\alpha}$):

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

where, $N_j$ refers to the $j^{\text{th}}$ neighbor of $i$ in the Queen's neighborhood ($\mathcal{N}_i$ in (4)) and $d_{N_j,i}$ is the Euclidean distance between area $N_j$ and $i$.

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. Since the transition probabilities $\mathbf{p}$ are the focal point of the neighborhood-based dynamics in this class of models, allowing them to depend hierarchically upon the approximate gradients $\tilde{\mathbf{A}}$ is critical to the utilization of covariate information and non-stationary anisotropic dynamics in the process. Therefore, let $\mathbf{A}$ be a fuction 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. Consider, for example, the probit or standard normal CDF function (the exponential is another option) which 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 where the parameters $\mathbf{A}$ are only proportional 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). 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:

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

Recall, (from (3)) that the actual probability of presence in area $i$ when area $i$ has been 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. Utilizing 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 non-stationary model, let $\bar{p}_{i,t}$ be defined as:

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

where, $\mathbf{y}_{\mathcal{N}_i,t-1}$ is defined as in the previous section, $\mathbf{p}_{\mathcal{N}_i} = [p_{N_9,1}, p_{N_8,2}, \ldots, p_{N_1,9}]'$, and $N_j$ refers to the $j^{\text{th}}$ neighboring area or location of $i$. In this way, the probability of presence in area $i$ at the current time is the probability of the unions of transitions into area $i$ from occupied neighboring areas 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 non-neighborhood based probability of presence be defined as before, $\phi \sim \text{Beta}(\alpha_\phi, \beta_\phi)$ and $\psi \sim \text{Beta}(\alpha_\psi, \beta_\psi)$. The spatial covariance parameters on the habitat preference spatial field, $\sigma^2_\alpha$ and $\theta_\alpha$, can be modeled in the conventional manner where $\sigma^2_\alpha \sim \text{InvGamma}(q, r)$ and $\theta_\alpha \sim \text{Gamma}(\alpha_\theta, \beta_\theta)$. The spatial range parameter $\theta_\alpha$ may be difficult to identify if there is minimal residual spatial structure in the $\boldsymbol{\alpha}$ process. Additionally, due to the formulation of the model, a large $\theta_\alpha$ (implying greater spatial structure) may be counterproductive considering that the model is trying to find sharp differences between neighboring areas (i.e., due to the directional gradient fields in $\tilde{\mathbf{A}}$).

The posterior distribution of interest can now be written as:

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

The full-conditional distributions for the model parameters necessary for implementation of this non-stationary anisotropic model can be found in a similar fashion as those in the previous section.

## 3. Results

This section contains a description of the results of fitting the two previously introduced models (i.e., (2.2) and

(2.3)) to the Connecticut rabies dataset (Figure 1); a full discussion of the results follows in the next section (4).

The rabies dataset introduced in Section 1.1 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 models. The models proposed were presented with sufficient generality to model the data in their original format however, and the implementation of such generalizations is the focus of ongoing research. Moreover, as previously discussed, the rabies data were originally available in "time since first appearance" in each township (i.e., spatial area), and were converted into presence / absence data for use with this model. This form of data is more common in epidemiological and studies where more traditional biostatistical models are utilized (e.g., longitudinal data analysis and survival analysis). The models developed here, though not specifically developed to address such data, could be modified to allow time-delay data to enter through a different data model (e.g., an exponential or Wiebull distribution). Such generalizations to other data structures are also the focus of ongoing research.

Another important note, regarding the persistence of the process, is that epidemic behavior is often characterized by a travelling epidemic "wave," where 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 ($\phi$) 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 where the disease can be spread readily and is largely unmanaged, the assumption is not unreasonable; however, such an assumption would not be appropriate in all epidemiological situations. In either event, the persistence term ($\phi$) should accommodate differing levels of dispersal.

The models from Sections 2.2 and 2.3 were implemented by sampling from the posterior distributions via the full conditional distributions. Non-conjugate full-conditional distributions were sampled via Metropolis-Hastings updates. Convergence in the MCMC samples was reached relatively quickly (i.e., $< 1000$ iterations) and the remaining 19,000 realizations were resampled to remove autocorrelation induced by the algorithm and then used to calculate posterior statistics.

### 3.1 Stationary Model

This section contains the results of fitting the anisotropic stationary model introduced in Section 2.2 to the Connecticut rabies data (Figure 1). The estimated posterior distributions for the neighborhood-based transition probabilities $\mathbf{p}$ are presented as boxplots in Figure 2. Recall that $\mathbf{p}$ is comprised of 9 transition probabilities, one representing neighborhood based dispersal in each of the first 8 cardinal directions (i.e., Northwest, West, Southwest, North, South, Northeast, East, Southeast), also $p_5$ corresponds to the probability of no transition. Also shown in Figure 2 is the expected posterior 95% credible interval of $\mathbf{p}$ given only the information from the prior distribution of $\mathbf{a}$ (also the expected posterior mean). Due to the hierarchical structure of the model, the neighborhood-based transition probabilities $\mathbf{p}$ have no prior distribution, however, if the data were uninformative, the posterior distribution for $\mathbf{p}$ would resemble the credible interval shown in dashed lines in Figure 2. The posterior mean and standard deviation for persistence ($\phi$) were 0.9995 and 0.0001, while for non-neighborhood dispersal ($\psi$) were 0.01 and 0.0015, respectively. The posterior mean of the probability of presence ($\theta_{i,t}$) is shown for each time step $t$ and spatial area $i$ in Figure 3. The sequence of maps is shown in the same order as the data in Figure 1, where the first time step appears in the upper left corner of the figure and downward, then to the next column from top to bottom, and so on until reaching the final time step in the lower right corner of the figure. The posterior standard deviation of the probability of presence ($\theta_{i,t}$) is shown in the same manner as the posterior mean in Figure 4. Deviance Information Criterion (DIC) was used as a form of model comparison for the stationary versus non-stationary models implemented here is applicable especially in this case because the data models used are specified the same. DIC, in this case, provides a measure of how well the top level parameters in the hierarchy fit the data. Here, the top level parameters (i.e., $\mathbf{p}$, $\phi$, and $\psi$) are the most important because they provide an insight into the dynamics of dispersal in spreading phenomena. DIC also penalizes (likewise for AIC and BIC) for the number of effective parameters in the model so that overfit models can be recognized and inference can be made cautiously. The effective number of parameters

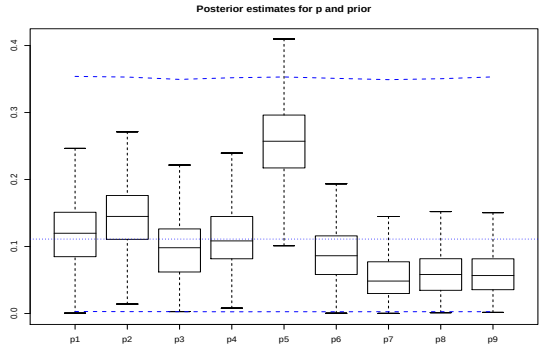



Figure 2: Posterior estimates for the parameters in $\mathbf{p}$ corresponding to: NW, W, SW, N, None, S, NE, E, SE in terms of directional probability of neighborhood-based propagation. Dashed lines correspond to the 95% credible interval based on only the prior distribution of $\mathbf{a}$ in the case where no information is gained from the data about $\mathbf{p}$. Dotted line corresponds to the mean directional propagation probability under the prior.

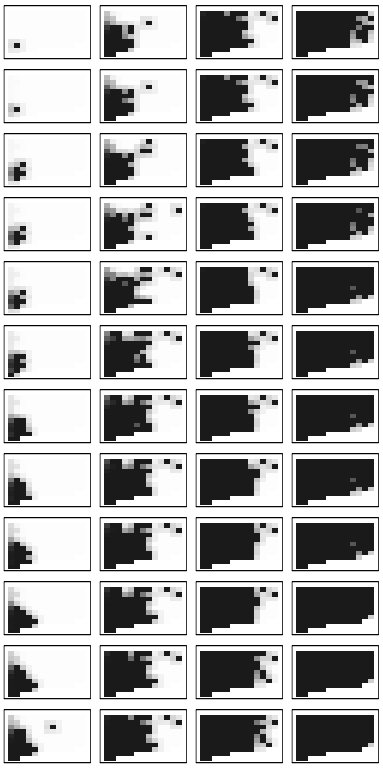


Figure 3: Posterior mean of $[\theta_{i,t}|\mathbf{y}]$ from stationary model in the same sequence as the data in Figure 1. Values of intensity range between zero (white) and one (black).

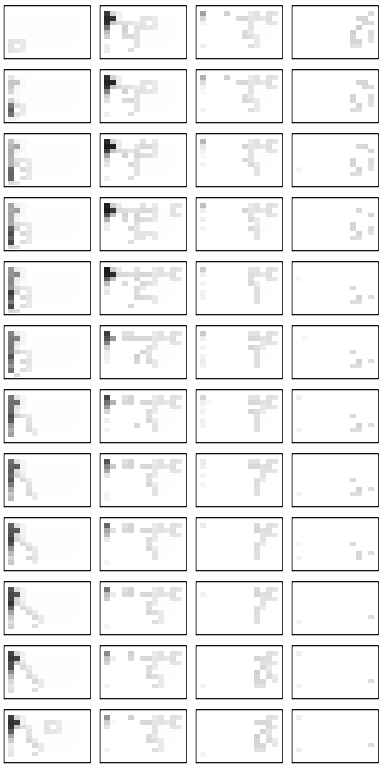


Figure 4: Posterior standard deviation of $[\theta_{i,t}|\mathbf{y}]$ from stationary model in the same sequence as the data in Figure 1. Values of intensity range between zero (white) and one (black).

for the stationary model was estimated to be $p_{D_1} = 36.5$, while the DIC value was 723.9 (after correcting for the effective number of parameters).

### 3.2 Non-Stationary Model

Figure 5 shows the covariates ($\mathbf{X}$) used in fitting the anisotropic non-stationary model from Section 2.3 to the Connecticut rabies dataset given in Figure 1. All covariates used in this analysis were binary with the black regions in Figure 5 corresponding to the occurrence of the environmental characteristic. Specifically, the $X_1$ and $X_2$ covariates correspond 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. A covariate corresponding to human population density was omitted from this analysis because of lack of availability. It is important to note, however, that major population centers in Connecticut coincide with the Connecticut River ($X_1$ and $X_2$) and coastline ($X_4$) covariates and is therefore implicitly accounted for in $\mathbf{X}$. An explicit continuous human population covariate may prove to be helpful in future analyses. The posterior distributions of the neighborhood-based

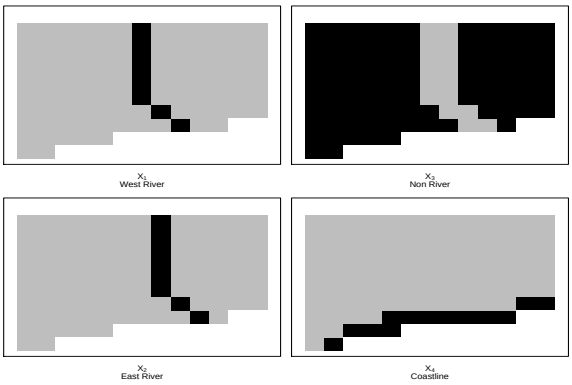



Figure 5: The covariates used in the non-stationary model: $\mathbf{x}_1$ and $\mathbf{x}_2$ represent the west and east sides of the Connecticut River, respectively, while $\mathbf{x}_3$ represents all land area not adjacent to the Connecticut River, and $\mathbf{x}_3$ represents the Connecticut coastline.

dispersal for the anisotropic non-stationary model could be displayed as in Figure 2, however they would have to shown for each spatial location $i$ since $\mathbf{p}_i$ now varies in space. Instead, consider the posterior mean of each spatial field of the neighborhood-based dispersal probabilities in Figure 6. The $j^{\text{th}}$ map in Figure 6 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 non-dispersal. For example, the map in Figure 6 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

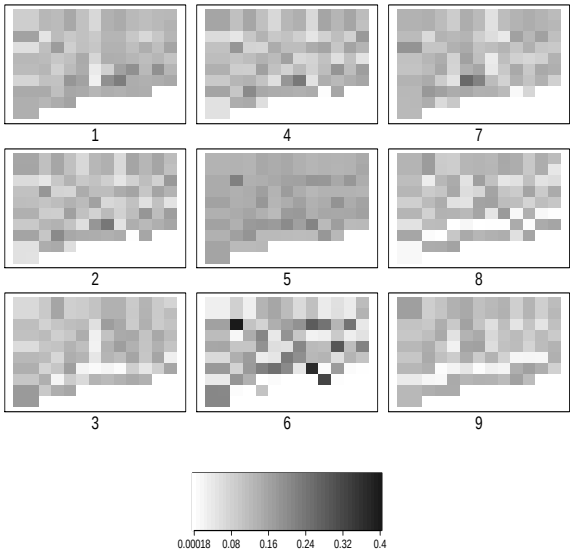
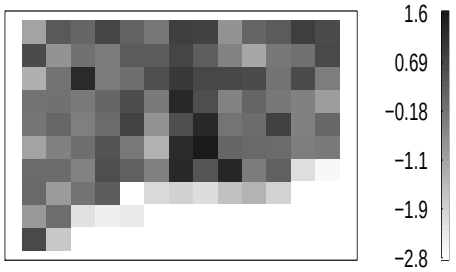



Figure 6: Posterior mean of $[\mathbf{P}|\mathbf{y}]$ from non-stationary model with respect to the Queen's neighborhood arrangement. The $j^{\text{th}}$ map refers to the posterior mean spatial field of the neighborhood-based transition probabilities in the $j^{\text{th}}$ direction.

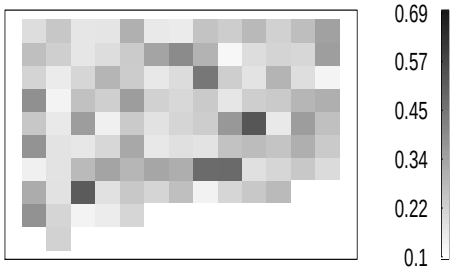
preference ($\boldsymbol{\alpha}$) given in Figure 7a, in this case corresponds to areas (shown as darker grid cells) where the rabies epidemic spreads most quickly towards. The posterior standard deviation corresponding to habitat preference is provided in Figure 7b and illustrates where areas of high uncertainty regarding habitat preference are over the spatial domain. The posterior distributions for the covariate effect parameters ($\boldsymbol{\beta}$) are provided in Figure 8 and, when premultiplied by the covariates ($X$), correspond to the mean of habitat preference ($\boldsymbol{\alpha}$). The posterior distributions for persistence ($\phi$) and non-neighborhood dispersal ($\psi$) for the anisotropic non-stationary model were very similar to that of the stationary model.

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

Deviance Information Criterion was also calculated for this anisotropic non-stationary model and had a value of 702.6, while the effective number of parameters was estimated as $p_{D_2} = 1.92$.




(a) $\mathrm{E}(\boldsymbol{\alpha}|\mathbf{y})$




(b) $\sqrt{V(\boldsymbol{\alpha}|\mathbf{y})}$

Figure 7: Posterior mean and standard deviation of habitat preference ($\boldsymbol{\alpha}$), respectively. Large values in the $\mathrm{E}(\boldsymbol{\alpha}|\mathbf{y})$ correspond to areas of higher suitability.

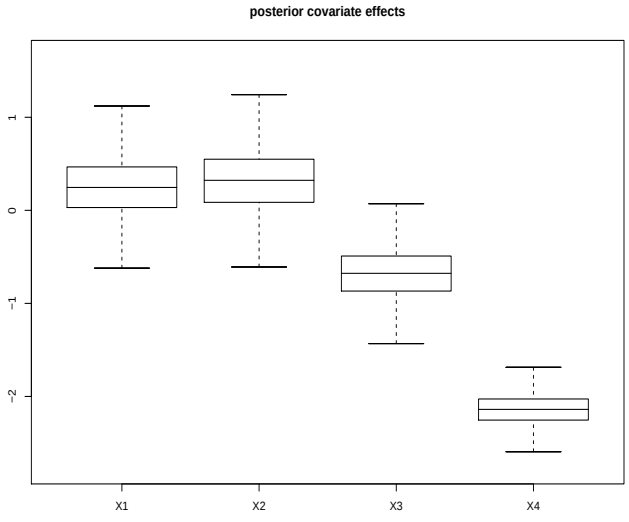



Figure 8: Posterior distributions of the covariate effects ($\boldsymbol{\beta}$), corresponding to the covariates in Figure 5.

### 3.3 Discussion

The results of fitting the stationary and non-stationary models to the Connecticut rabies dataset indicate that both methods have strengths and weaknesses. Beginning with the results from the anisotropic stationary model: Figure 2 provides a description of the neighborhood-based transition probabilities in each direction. More specifically, Figure 2 shows that in addition to all of the parameter estimates being much more precise than they would have been had the data been uninformative, they also suggest that neighborhood based dispersal is overall more likely in the North, Northwest, and West directions than in other directions due to the larger posterior values for $p_1$, $p_2$, and $p_4$ and smaller $p_7$, $p_8$, and $p_9$ values. Another notable feature of the posterior distributions for $\mathbf{p}$ is that $p_5$ is significantly greater than the other transition probabilities. Recall that $p_5$ is the parameter that controls the rate of spread. The fact that $p_5$ is greater suggests that rabies is propagating across Connecticut at a slower rate than it is capable of, given the time step. In sum, the neighborhood dynamics, when estimated with a stationary model, suggest that the spread of rabies in Connecticut during 1991–1995 was likely an anisotropic spatio-temporal process given the degree of separation in the posterior distributions of $\mathbf{p}$.

The posterior estimates of the persistence and non-neighborhood dispersal parameters $\phi$ and $\psi$ suggest that the process is extremely persistent due to the fact that $\phi$ is very close to one, while non-neighborhood dispersal (or long-distance dispersal) does indeed occur and has a mean near 0.01. Recall that in the transformation of the dataset for use with these models, the process was forced to persist, thus the posterior of $\phi$ is only indicating a known characteristic of the process. Had there not been persistence in the dataset, the $\phi$ parameter would have indicated this with a posterior significantly less than one. The posterior mean for $\psi$ indicates that dispersal events out of the neighborhood of areas with rabies present occurs only 1% of the time. It may be that if $\psi$ were allowed to vary with distance from active areas, it may be possible to make more specific inference regarding the likely distance of non-neighborhood dispersal. For example, perhaps dispersal events at 3–4 cells away from active cells are much more probable than other distances. This type of extension to the model is the focus of ongoing research.

Close inspection of the sequence of maps (Figure 3) showing the posterior mean of the probability of dispersal ($\theta_{i,t}$) in space and time illustrates the effect of the neighborhood-based dispersal parameters ($\mathbf{p}$) from Figure 2 on the overall propagation of the process. Specifically, the first map in the sequence (top left, Figure 3) shows that in the neighborhood of the active cell, the neighbors in the North, Northwest, and West directions are slightly darker than the neighbors in the other directions. This directly corresponds to the posterior mean values of $p_1$, $p_2$, and $p_4$ being greater than the remaining parameter values in $\mathbf{p}$. Notice that rabies propagates first up the left hand side of the map (i.e., western edge of Connecticut) and then eastward into the rest of the state. Again, this is the effect suggested by the posterior distribution of $\mathbf{p}$.

The posterior standard deviation of the probability of presence (Figure 4) provides information about the uncertainty regarding the propagation of rabies throughout Connecticut. Notice first that in the first column of maps in Figure 4 there are large values of standard deviation along the western edge of Connecticut. This is due to the lack of data from New York (Connecticut's western neighbor). The uncertainty is higher in New York because the parameters are being estimated in those locations without the aid of data. This high uncertainty in areas of limited data is a common artifact in spatial and spatio-temporal statistical models. Another feature evident in the sequence of maps in Figure 4 is the traveling wave of high uncertainty moving from the Southwest corner of Connecticut to the Northeast corner and then finally to the Southeast region. Comparing this figure with with Figure 3 indicates that the propagating uncertainty corresponds to the wave front of the propagating rabies epidemic and suggests that there is less certainty regarding those areas in flux than in the persisting areas and yet to be infected areas.

The results pertaining to the anisotropic non-stationary model provide some additional insight into the dynamics of the rabies epidemic in Connecticut. The posterior mean propagation of probability of presence ($\theta_{i,t}$) for the non-stationary model appeared similar to that of the stationary model (Figure 3). However, the posterior standard deviation for probability of presence ($\theta_{i,t}$) indicated 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 non-stationary model is more precisely estimating the neighborhood-based component of the probability of presence ($\theta_{i,t}$).

In fact, the maps corresponding to the spatial fields of the directional neighborhood-based dispersal parameters ($\mathbf{p}_j$) in Figure 6 indicate that there are indeed differences in directional dispersal depending on location. For example, in map 1 (Northwest direction) and 3 (Southwest direction) of Figure 6 there is a distinct column of low dispersal corresponding to the west side of the Connecticut River covariate ($X_1$ in Figure 5). From Figure 7a, it is evident that the Connecticut river is an area of high habitat preference; this is also substantiated by the posterior distribution of $\boldsymbol{\beta}$ in Figure 8. Thus, as the rabies epidemic spreads from the western region of Connecticut it exhibits a different pattern of neighborhood-based dispersal than in the other 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. This characteristic in dispersal is also evident

in Figure 1, where the second and third columns of maps in the sequence show the presence of rabies slowing in the eastward direction and instead spreading out to the North and South as the epidemic approaches the Connecticut River. Also evident from Figure 6 is that in areas just North of the Connecticut coastline, rabies is dispersing very slowly in the Southwest, Southeast, and East directions. One interpretation of this result is that the rabies epidemic in areas near the coast is propagating more inward (toward land) rather than outward (toward the Long Island Sound). This seems like an obvious characteristic of a land based phenomenon, yet such differences in neighborhood-based dispersal probabilities over a heterogeneous environment provides the added flexibility needed for a better model fit.

The posterior probability of persistence resulting from fitting the non-stationary model is again very close to one, suggesting that, given the available data (which were constructed to be persistent), the rabies epidemic in Connecticut is very persistent. The ramifications of this inference were previously discussed. The posterior mean probability of non-neighborhood dispersal is again estimated to be near 0.01, suggesting a 1% chance of long-distance rabies dispersal for non-neighboring areas.

One way to compare, *a posteriori*, the stationary model with the non-stationary model is to assess the difference in effective number of parameters as well as DIC for the two models. The effective number of parameters provides an indication of the degree to which parameters in the model are informed by the data (as opposed to informed by the prior). DIC provides a measure of the likelihood given the posterior estimates and corrects for the number of effective parameters. Therefore, given models with equal numbers of parameters, a higher $p_D$ value and a lower DIC value is desirable. In models with differing numbers of parameters (as with the two models considered here), the effective number of parameters are difficult to compare; however, the DIC value can still serve as a basis for model comparison. Based on the results of the two models fitted here, the non-stationary model has a lower DIC and thus suggests a better fitting model. On the other hand, parsimony is an important factor in model comparison so there is certainly still value in the stationary model as it is more computationally efficient and retains more degrees of freedom. The non-stationary model, like the stationary model has a dynamical component but also has an explicit connection to scientifically meaningful environmental covariates and is thus 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.

## 4. Conclusion

In this paper, two methods for modeling a spreading phenomena over a partitioned landscape, given only binary data, were developed and applied to the rabies epidemic in raccoon populations in Connecticut during 1991–1995. Both models take a probabilistic cellular automata approach, while the first model was specified so that the focus was on the estimation of the dynamics governing the process in terms of stationary neighborhood-based dispersal. The second model utilized environmental covariates to help determine how patterns of dispersal might differ given heterogeneous environmental conditions.

Although model 1 assumed stationary dynamics, it was able to accommodate anisotropic dispersal behavior based on transition probabilities into neighboring areas. Model 2 allowed for anisotropic and non-stationary dynamics while accounting for associations between dispersal and environmental factors. Both models allowed for the possibility of long-distance dispersal and varying levels of persistence, although in this case the persistence parameter was not informative.

The models presented could be extended to address other issues as well. For example, an assumption one must make when applying these models is that of perfect detection. That is, 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 a latent process model would be employed to model the "true" presence or absence of rabies in an area. Royle and Kery (2006) and Royle (2006) have taken this approach in a logit-transformed framework and are able to make use of repeated measurements to estimate such parameters. In the situation considered here repeated measurements are not available and without another means to estimate detectability, the employment of 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 3) absorbs the detection probability in $\theta_{i,t}$ at locations where rabies has yet to be observed.

Overall, the methods presented here place the simulation models presented by Smith et al. (2002) into a rigorous statistical framework where the parameters governing dispersal can be estimated and inference can be made about the effects of environmental covariates on dispersal. Models such as these could be used to identify important factors in the spread of epidemics (e.g., river corridors, shorelines, and population centers) and ultimately utilized in policy making and management decisions.

## Acknowledgments

The authors would like to thank Lance Waller for providing data in addition to many helpful comments and suggestions.

## References


Banerjee, S., B. Carlin, and A. Gelfand. 2004. *Hierarchical Modeling and Analysis for Spatial Data*. Chapman & Hall/CRC, New York.

Besag, J. 1974. Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistics Society*, 36:192–236.

Clark, J. 2003. Uncertainty and variability in demography and population growth: A hierarchical approach. *Ecology*, 84(6):1370–1381.

Cressie, N. 1993. *Statistics for Spatial Data: Revised Edition*. John Wiley and Sons, New York, New York, USA.

Grimm, V., E. Revilla, U. Berger, F. Jeltsch, W. Mooij, S. Railsback, H.-H. Thulke, J. Weiner, T. Wiegand, and D. DeAngelis. 2005. Pattern-oriented modeling of agent-based complex systems: Lessons from ecology. *Science*, 310:987–991.

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

Heikkinen, J. and H. Hogmander. 1994. Fully bayesian approach to image restoration with an application in biogeography. *Applied Statistics*, 43(4):569–582.

Hirzel, A., J. Hausser, D. Chessel, and N. Perrin. 2002. Ecological-niche factor analysis: How to compute habitat-suitability maps without absence data? *Ecology*, 83:2027–2036.

Hoeting, J., M. Leecaster, and D. Bowden. 2000. An improved model for spatially correlated binary responses. *Journal of Agricultural, Biological, and Environmental Statistics*, 5(1):102–114.

Hogeweg, P. 1988. Cellular automata as a paradigm for ecological modeling. *Applied Mathematics and Computation*, 27:81–100.

Hogmander, H. and J. Moller. 1995. Estimating distribution maps from atlas data using methods of statistical image analysis. *Biometrics*, 51:393–404.

Hooten, M., D. Larsen, and C. Wikle. 2003. Predicting the spatial distribution of ground flora on large domains using a hierarchical Bayesian model. *Landscape Ecology*, 18:487–502.

Hooten, M. and C. Wikle. 2007. 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*, DOI: 10.1007/s10651-007-0040-1.

Hooten, M., C. Wikle, R. Dorazio, and J. Royle. 2007. Hierarchical spatio-temporal matrix models for characterizing invasions. *Biometrics*, 63:558–567.

Krebs, C. 1978. *Ecology: The Experimental Analysis of Distribution and Abundance*. Harper & Row Publishers Inc., New York, New York.

Lee, Y., S. Qian, R. Jones, C. Barnes, G. Flake, M. O'Rourke, K. Lee, H. Chen, G. Sun, Y. Zhang, and D. Chen. 1990. Adaptive stochastic cellular automata: Theory. *Physica D*, 45:159–180.

MacKenzie, D., J. Nichols, J. Hines, M. Knutson, and A. Franklin. 2003. Estimating site occupancy, colonization, and local extinction when a species is detected imperfectly. *Ecology*, 84:2200–2207.

MacKenzie, D., J. Nichols, G. Lachman, S. Droege, J. Royle, and C. Langtimm. 2002. Estimating site occupancy rates when detection probabilities are less than one. *Ecology*, 83:2248–2255.

Nettles, V., J. Shaddock, R. Sikes, and C. Reyes. 1979. Rabies in translocated raccoons. *American Journal of Public Health*, 69:601–602.

Pace, R., R. Barry, O. Gilley, and C. Sirmans. 2000. A method for spatial-temporal forecasting with an application to real estate prices. *International Journal of Forecasting*, 16:229–246.

Royle, J. 2006. Modeling spatio-temporal dynamics in occupancy: Nest site occupancy in a Kittiwake colony. *In Review*.

Royle, J. and R. Dorazio. 2006. Hierarchical models of animal abundance and occurrence. *Journal of Agricultural, Biological, and Environmental Statistics*, In Press.

Royle, J. and M. Kery. 2006. A Bayesian state-space formulation of dynamic occupancy models. *In Review*.

Smith, D., B. Lucey, L. Waller, J. Childs, and L. Real. 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.

Waller, L. and C. Gotway. 2004. *Applied Spatial Statistics for Public Health Data*. John Wiley & Sons, New Jersey.

Wolfram, S. 1983. Statistical mechanics of cellular automata. *Reviews of Modern Physics*, 55:601–644.

Wolfram, S. 1984. Cellular automata as models for complexity. *Nature*, 311:419.

Wolfram, S. *In: High-Speed Computing: Scientific Applications and Algorithm Design*, chapter Cellular Automaton Supercomputing, pages 40–48. University of Illinois Press, 1988.

Zhu, J., H.-C. Huang, and C.-T. Wu. 2005. Modeling spatial-temporal binary data using Markov random fields. *Journal of Agricultural, Biological, and Environmental Statistics*, 10:212–225.