Spatial Statistics 20 (2017) 206–220

Contents lists available at ScienceDirect

Spatial Statistics

journal homepage: www.elsevier.com/locate/spasta

Dynamic spatio-temporal models for spatial data

Trevor J. Hefley [a,*], Mevin B. Hooten [b], Ephraim M. Hanks [c], Robin E. Russell [d], Daniel P. Walsh [d]

[a] *Department of Statistics, Kansas State University, United States*

[b] *U.S. Geological Survey, Colorado Cooperative Fish and Wildlife Research Unit, Department of Fish, Wildlife, and Conservation Biology, Department of Statistics, Colorado State University, United States*

[c] *Department of Statistics, Pennsylvania State University, United States*

[d] *U.S. Geological Survey, National Wildlife Health Center, United States*

ARTICLE INFO

Article history:
Received 28 September 2016
Accepted 16 February 2017
Available online 12 April 2017

Keywords:
Ecological diffusion
Generalized linear mixed model
Homogenization
Partial differential equations
Spatial confounding
Spatial statistics

ABSTRACT

Analyzing spatial data often requires modeling dependencies created by a dynamic spatio-temporal data generating process. In many applications, a generalized linear mixed model (GLMM) is used with a random effect to account for spatial dependence and to provide optimal spatial predictions. Location-specific covariates are often included as fixed effects in a GLMM and may be collinear with the spatial random effect, which can negatively affect inference. We propose a dynamic approach to account for spatial dependence that incorporates scientific knowledge of the spatio-temporal data generating process. Our approach relies on a dynamic spatio-temporal model that explicitly incorporates location-specific covariates. We illustrate our approach with a spatially varying ecological diffusion model implemented using a computationally efficient homogenization technique. We apply our model to understand individual-level and location-specific risk factors associated with chronic wasting disease in white-tailed deer from Wisconsin, USA and estimate the location the disease was first introduced. We compare our approach to several existing methods that are commonly used in spatial statistics. Our spatio-temporal approach resulted in a higher predictive accuracy when compared to methods based on optimal spatial prediction, obviated confounding among the spatially indexed covariates and the spatial

* Corresponding author.
E-mail address: thefley@ksu.edu (T.J. Hefley).

http://dx.doi.org/10.1016/j.spasta.2017.02.005

random effect, and provided additional information that will be important for containing disease outbreaks.

1. Introduction

The generalized linear mixed model (GLMM) is an important tool for spatially continuous and areal data (Besag et al., 1991; Gotway and Stroup, 1997; Diggle et al., 1998), is widely used in fields such as econometrics, epidemiology, and ecology (e.g., Waller and Gotway, 2004; Bolker et al., 2009), and is the building block for more sophisticated hierarchical models (Cressie and Wikle, 2011). In many applications of the spatial GLMM, the goal is to obtain inference about the regression coefficients by accounting for spatial dependence while simultaneously obtaining optimal predictions at locations that were not sampled. Although the spatial GLMM is widely used in discipline-specific domains, there are surprisingly few examples where the model component that accounts for spatial dependence is specified by incorporating subject matter knowledge (Hanks, in press). For example, the spatial pattern of disease occurrences in wildlife may be driven by host species movement patterns across a landscape (e.g., Smith et al., 2002; Wheeler and Waller, 2008) and incorporating this information into model components that account for spatial dependence may be useful.

In contrast, subject matter knowledge is commonly used in spatio-temporal applications of the GLMM, often by specifying a partial differential equation (PDE) that incorporates a priori knowledge of the data generating process (Wikle, 2003; Wikle and Hooten, 2010). The discrepancy in how models for spatial and spatio-temporal data are specified might be related to the notion that spatial data, with respect to a spatio-temporal generating process, are information poor relative to spatio-temporal data. Although spatial data generated by an underlying spatio-temporal process contain no information about the temporal dynamics, incorporating subject matter knowledge of the spatio-temporal process is needed to understand the spatial dynamics. For example, the phenomenon of spatial confounding that occurs when the spatial random effect is collinear with the covariates (Reich et al., 2006; Hodges and Reich, 2010), can negatively influence inference and might be obviated by considering subject matter knowledge (Hanks et al., 2015; Hanks, in press). In what follows, we briefly review the spatial GLMM as well as common practices in spatial data analysis. We then introduce an approach for modeling spatial data that incorporates subject matter knowledge of spatio-temporal data generating process.

The GLMM for spatial processes can be written as

$$\mathbf{y} \sim [\mathbf{y}|\boldsymbol{\mu}, \psi] \tag{1}$$

$$g(\boldsymbol{\mu}) = \beta_0 + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{u} \tag{2}$$

$$\boldsymbol{u} \sim \mathrm{N}\left(\mathbf{0}, \sigma_u^2 \mathbf{C}(\boldsymbol{\phi})\right), \tag{3}$$

where $\mathbf{y} \equiv (y_1, \ldots, y_n)'$ is a set of n observations at fixed spatial locations $\mathbf{s}_i$ $(i = 1, 2, \ldots, n)$ that arise from an arbitrary distribution belonging to the exponential family denoted by $[\cdot]$. The parameters $\boldsymbol{\mu}$ and ψ are the conditional expected value and dispersion of $[\cdot]$ respectively. The inverse of the link function $g(\cdot)$ is applied element-wise and transforms the linear predictor $\beta_0 + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{u}$ so that it has the same support as $\boldsymbol{\mu}$. The linear predictor has four components: an intercept term β_0, a $n \times p$ matrix that contains covariates ($\mathbf{X}$), a $p \times 1$ vector of fixed-effect regression coefficients ($\boldsymbol{\beta}$), and a zero-mean Gaussian random effect ($\boldsymbol{u} \equiv (u_1, \ldots, u_n)'$) with variance σ_u^2 and spatial correlation matrix $\mathbf{C}(\boldsymbol{\phi})$ with parameters $\boldsymbol{\phi}$. In many applications, however, subject matter knowledge of the underlying spatio-temporal data generating process is not explicitly considered when specifying the correlation matrix for spatial data (Hanks, in press). Rather, a semiparametric approach is typically taken, with the class of spatial covariance chosen based on the support of the data. For example, when using point

referenced data, the spatial random effect $\boldsymbol{u}$ is most commonly modeled using a covariance function of the Matérn class (e.g., Cressie, 1993; Stein, 1999).

Spatial data almost always arise from a spatio-temporal generating process observed over a fixed interval of time. For example, spatial patterns observed in binary data that capture the presence of an infectious disease can be explained by a dynamic transmission process occurring in space and time (e.g., Smith et al., 2002; Hooten and Wikle, 2010). Incorporating scientific knowledge about the spatio-temporal data generating process is critical to specifying an appropriate random effect in (1)–(3) that adequately models the spatial dependence. For example, the popular Matérn covariance function can be motivated by considering the spatial process to be a stationary solution to a spatio-temporal stochastic PDE (Whittle, 1954, 1963; Hanks, in press). Relating stochastic PDEs to covariance functions has aided in linking the spatial random effect to a spatio-temporal data generating process (Hanks, in press) and for developing computationally efficient implementations of the spatial GLMM (Lindgren et al., 2011).

When implementing the spatial GLMM in (1)–(3), covariates associated with a spatial location (hereafter location-specific covariates) are sometimes collinear with the spatial random effect when both have smooth spatial structure (Reich et al., 2006; Hodges and Reich, 2010; Paciorek, 2010; Hughes and Haran, 2013; Hanks et al., 2015; Murakami and Griffith, 2015). When using model specifications that included location-specific covariates with smooth spatial patterns and spatially structured random effects, both the regression coefficients and the random effect compete to explain variability in the response, which can produce surprising and counterintuitive results (e.g., Hodges and Reich, 2010; Hefley et al., 2016). The term "spatial confounding" has been used to describe collinearity among location-specific covariates in $\mathbf{X}$ and the spatial random effect $\boldsymbol{u}$ in (2) and may be a consequence of model misspecification (Hanks et al., 2015). This is in contrast to observation-specific covariates, which are the covariates associated with a particular observation, but not necessarily anchored to any specific spatial location (e.g., sex or age of a mobile individual). Observation-specific covariates may lack spatial structure and, therefore, are less likely to result in spatial confounding, although the potential for confounding does exist.

The standard spatial GLMM does not differentiate location-specific covariates from observation-specific covariates, both of which are included in $\mathbf{X}$. For point referenced data, one could specify a covariance function that depends on location-specific covariates as an alternative to specifying location-specific covariates as fixed effects (e.g., Calder, 2008; Reich et al., 2011; Gladish et al., 2014; Ingebrigtsen et al., 2014; Poppick and Stein, 2014; Vianna Neto et al., 2014; Risser and Calder, 2015). Specifying a covariance function that depends on location-specific covariates has the obvious advantage of obviating collinearity between covariates and the spatial random effect in (1)–(3), but also allows for the covariates to influence the properties of the spatial random effect. For example, Ingebrigtsen et al. (2014) extended the work of Lindgren et al. (2011), developing a flexible class of non-stationary models where location-specific covariates can be specified to influence parameters of a stochastic PDE and, thus, influence the spatial random effect. Although it is possible to specify the spatial GLMM so that location-specific covariates influence the spatial random effect, a principled approach to determine the role of location-specific covariates is needed. For example, what subject matter knowledge should be used to determine if a location-specific covariate should be included in $\mathbf{X}$ and, therefore, influence the expected value of (1) or included in the covariance function and, therefore, influence the spatial random effect in (3)?

One way to view the stochastic PDE approach of Lindgren et al. (2011) is that the spatial random effect, $\boldsymbol{u}$, is the stationary (time-limiting) solution to a spatio-temporal stochastic PDE (Hanks, in press). One disadvantage of this approach is that assuming $\boldsymbol{u}$ is the stationary solution to a spatio-temporal stochastic PDE results in a loss of information about the initial conditions that gave rise to the data. For environmental and ecological applications, the spatio-temporal data generating process might demonstrate transient (non-equilibrium) dynamics and, therefore, could be used to make critical inference. For example, spatial data pertaining to disease outbreaks are a result of a dynamic spatio-temporal transmission process that has not reached equilibrium and is dependent on the initial conditions (e.g., location of introduction). Likewise, when a species invades a novel environment, the underlying spatio-temporal population dynamics are transient (e.g., Wikle, 2003; Hooten et al., 2007; Williams et al., 2017). When attempting to control disease outbreaks or halt the invasion of a pest

species, identifying the initial conditions, such as the location of introduction, is critical to containing the incident and preventing future invasions or outbreaks (e.g., Snow, 1855; Nash et al., 2001).

Our goal is to develop an approach for modeling spatial random effects as the transient distribution of the spatio-temporal data generating process. This approach can be applied to large data sets, incorporates scientific knowledge of the spatio-temporal data generating process, allows for inference on initial conditions, eliminates spatial confounding, and has competitive predictive ability when compared to methods based on optimal spatial prediction (e.g., Gotway and Stroup, 1997; Diggle et al., 1998). To implement our method for large data sets, we apply the homogenization technique that is used in applied mathematics to obtain approximate solutions to PDEs, but only recently has been used in statistical applications (Garlick et al., 2011; Holmes, 2012; Hooten et al., 2013). We then use our spatial modeling approach to understand individual-level and location-specific risk factors associated with chronic wasting disease in white-tailed deer from Wisconsin, USA. We compare our approach to the methods developed by Lindgren et al. (2011) and Ingebrigtsen et al. (2014) and show that spatial confounding can be obviated by specifying a spatio-temporal component that depends on location-specific covariates.

2. Methods

We propose replacing the spatial random effect in (1)–(3), or any other appropriate hierarchical model, with the transient solution to a spatio-temporal PDE. Given a model for which a spatial effect is desired, we propose constructing the spatial effect as follows:

1. Define a deterministic spatio-temporal generating model for the spatio-temporal effect $u(\mathbf{s}, t)$

$$\frac{\partial}{\partial t}u(\mathbf{s}, t) = \mathcal{F}\left(u(\mathbf{s}, t); \mathbf{z}(\mathbf{s})\right) \tag{4}$$

   where $\mathbf{s}$ indexes space, t indexes time, and $\mathcal{F}$ depends on the location-specific covariates $\mathbf{z}(\mathbf{s})$. For example, use the differential operator such that $\mathcal{F}\left(u(\mathbf{s}, t); \mathbf{z}(\mathbf{s})\right) \equiv \left(\frac{\partial^2}{\partial s_1^2} + \frac{\partial^2}{\partial s_2^2}\right)\delta(\mathbf{s})u(\mathbf{s}, t)$ where s_1 and s_2 are the spatial coordinates contained within the vector $\mathbf{s}$, $\delta(\mathbf{s})$ is the spatially varying diffusion rate that depends on $\mathbf{z}(\mathbf{s})$ and regression coefficients $\boldsymbol{\alpha}$ (see example in Section 3.4).
2. Solve (4) to obtain $u(\mathbf{s}, t)$ at the desired t and use the solution to model the spatial process in a hierarchical model (e.g., (1)–(3)).

Solving (4) requires the initial state $u(\mathbf{s}, t_0), \forall\, \mathbf{s}$. ($\mathbf{s}_i \in \mathcal{S} \subset \mathcal{R}^2$), where t_0 is the time the process was initiated. In many cases, the initial state $u(\mathbf{s}, t_0)$ is identifiable if an informative constraint or an informative prior under a Bayesian specification are justifiable. The specification of such informative constraint or prior will be problem specific (see Section 3.4). Our method also assumes that the time elapsed since initiation (t_0) and when the spatial data were collected (t) is known. The implication of the assumption that the elapsed time is known will depend on the specific PDE that is used. As such, we address the impact of this assumption in our data example.

2.1. Homogenization

Obtaining the spatio-temporal effect $u(\mathbf{s}, t)$ in (4) requires solving the PDE. Although several numerical methods exist to solve PDEs (e.g., finite element methods, finite difference methods), they become computationally prohibitive at the discretization required for fine-scale inference over a large spatial domain. To address this computational challenge, we apply the homogenization technique (Garlick et al., 2011; Holmes, 2012; Hooten et al., 2013). Homogenization is an analytical approach that takes advantage of multiple scales to solve a particular PDE on small scales and derives a related PDE at the large scale that accurately represents the integrated consequences of the small-scale solution behavior (Holmes, 2012). Homogenization of PDEs leads to more efficient numerical solutions when the goal is to capture large scale dynamics with rapid fine-scale spatial variability (Hooten et al., 2013; Powell and Bentz, 2014; Hefley et al., 2017b).

2.2. Hierarchical modeling framework

Hierarchical models with PDE process components are flexible and can be tailored to match the specifics of the study (e.g., Wikle et al., 2001; Wikle, 2003; Hooten et al., 2007; Hooten and Wikle, 2008; Wikle and Hooten, 2010; Zheng and Aukema, 2010; Williams et al., 2017). We describe a hierarchical modeling framework consistent with the terminology common to statisticians who use GLMMs, but modify the linear predictor in (2) to appropriately capture the dynamics of the process. Our modeling approach replaces the linear predictor (2) with

$$g(\mu_i) = h\left(\mathbf{x}_i'\boldsymbol{\beta}, u(\mathbf{s}_i, t_i)\right) \tag{5}$$

where the inverse of the link function $g(\cdot)$ transforms the (possibly) nonlinear predictor to the support of μ_i. The nonlinear predictor (5) has three components: a vector of observation-specific covariates ($\mathbf{x}_i$) associated with each observation y_i, a vector of regression coefficients ($\boldsymbol{\beta}$), and the spatio-temporal effect ($u(\mathbf{s}_i, t_i)$) that is a solution to a PDE. The function $h(\cdot)$, in (5), combines the influence of $\mathbf{x}_i'\boldsymbol{\beta}$ and the spatio-temporal effect $u(\mathbf{s}_i, t_i)$. For example, the linear form of $h(\cdot)$ used in the standard GLMM results in: $g(\mu_i) = \mathbf{x}_i'\boldsymbol{\beta} + u(\mathbf{s}_i, t_i)$. Although useful in some contexts, linear forms of $h(\cdot)$ may not be the most appropriate (e.g., Hooten et al., 2007; Cangelosi and Hooten, 2009; Williams et al., 2017). Determining how $\mathbf{x}_i'\boldsymbol{\beta}$ and $u(\mathbf{s}_i, t_i)$ are combined via $h(\cdot)$ will be problem specific and require knowledge of the process being modeled, which we demonstrate in the next section with an example.

3. Example: predicting the geographic distribution of chronic wasting disease

3.1. Data

The spread of infectious diseases is inherently a spatio-temporal process (e.g., Waller et al., 1997; Hooten and Wikle, 2010). During the initial stages of disease outbreak, researchers might have access to spatial data from only a single biologically relevant time period (e.g., one season). In what follows, we describe a spatial data set that exemplifies many of the challenges faced when modeling the initial stages of a disease outbreak.

Chronic wasting disease (CWD) is a fatal transmissible spongiform encephalopathy that occurs in cervids (Williams and Young, 1980). In the state of Wisconsin, CWD was first detected in white-tailed deer in 2001 as a result of the Wisconsin Department of Natural Resources surveillance efforts. After initial detection of CWD in 2001, increased sampling in 2002 resulted in a large spatial data set (Joly et al., 2009, Fig. 1). Using this data set, we analyzed records from 14,648 tested deer (168 positive deer) that were sampled within a 15,539 km^2 region in the southwestern portion of Wisconsin and had location information collected at the public land survey system section-level or better (Fig. 1). Although dynamics of CWD within a year might be important, most deer (>90%) were collected during the fall hunting season (October–December); therefore, we lack sufficient resolution in the data to model a finer temporal scale (e.g., monthly) and have only a single snapshot in time of the spatio-temporal data generating process.

We used the 2011 National Land Cover Dataset to calculate landscape risk factors (Homer et al., 2015). Although many different landscape risk factors could be used, we considered only the proportion of hardwood forest within a grid cell that was the same area of a section of land (2.6 km^2; Fig. 2(a)), because this location-specific covariate resulted in a high level of spatial confounding in previous analyses (Hefley et al., 2017a). In addition to the location-specific covariate, we also used the observation-specific covariates sex and age of the tested deer.

For initial model fitting and statistical inference, we used all 14,648 tested deer within our study area. To evaluate the predictive ability of our model, we fit our model using 50% of the data and calculated the logarithmic score for the remaining 50% of the data using predictions from our model. In section four, we compare the predictive ability of our method to existing methods. To facilitate comparison, we calculate the logarithmic score evaluated using the posterior mean of the probability

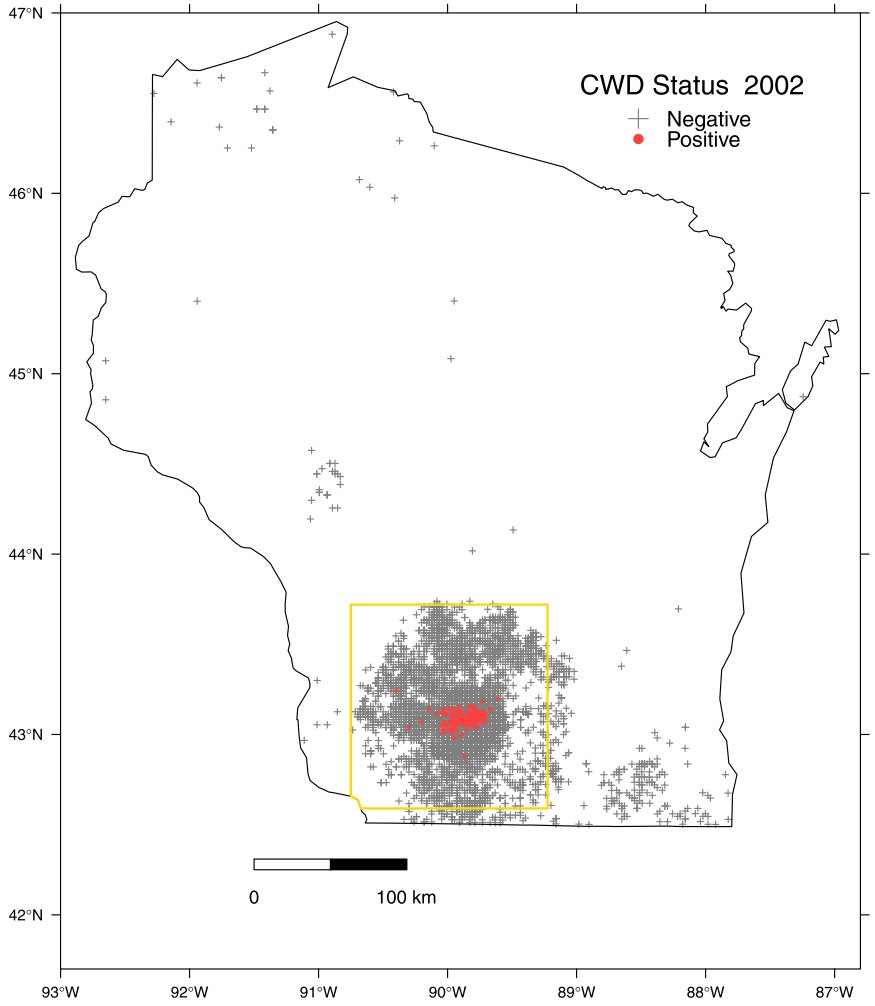

Fig. 1. Location and infection status of white-tailed deer tested for chronic wasting disease in Wisconsin, USA in 2002. Our models were applied to 14,648 tested deer with locations contained within the yellow box.

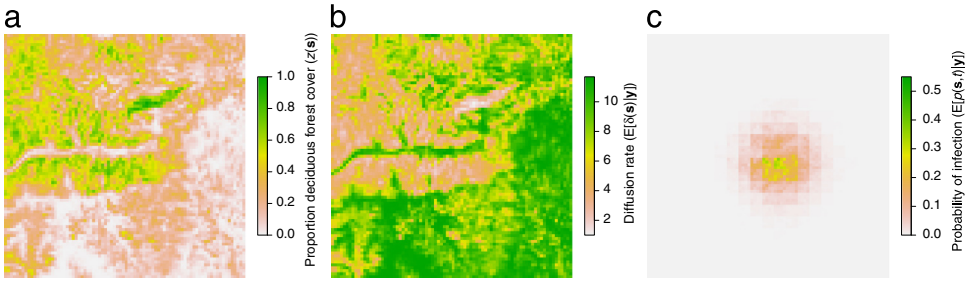

Fig. 2. A location-specific covariate used to model the geographic distribution of CWD in white-tailed deer (panel a; proportion of deciduous forest within each 2.6 km^2 grid cell). Posterior expectation of the scaled diffusion rate $\delta(\mathbf{s})$ (panel b) from the ecological diffusion model presented in (10)–(14) and fitted to the data in Fig. 1. Posterior expectation of the probability of CWD infection in male white-tailed deer that are 4 years of age or older (panel c; see Fig. 1 for study area).

of infection (i.e., $y_i \log(p_i) + (1 - y_i) \log(1 - p_i)$ where y_i is the ith out-of-sample observation and p_i is the posterior mean probability of infection; Gneiting and Raftery, 2007; Gneiting, 2011).

3.2. Ecological diffusion

Ecological diffusion has been motivated by tracking the location of a single individual that is undertaking a random walk in discrete time (e.g., Turchin, 1998; Hooten et al., 2017, pp. 189–191; see Fig. 1 in Hefley et al., 2017b). Motivated by the Lagrangian (individual-based) perspective of movement, the Eulerian (population-level) description of movement is captured by the ecological diffusion PDE

$$\frac{\partial}{\partial t} u(\mathbf{s}, t) = \left(\frac{\partial^2}{\partial s_1^2} + \frac{\partial^2}{\partial s_2^2} \right) \delta(\mathbf{s}) u(\mathbf{s}, t), \tag{6}$$

where $u(\mathbf{s}, t)$ is the density of the dispersing population (or disease), s_1 and s_2 are the spatial coordinates contained within the vector $\mathbf{s}$, and t is the time. The diffusion coefficient (or motility coefficient), $\delta(\mathbf{s})$, is inversely related to the amount of time an individual spends in a particular area (i.e., residence time) and could depend on location-specific covariates (Garlick et al., 2011). There are two other common types of diffusion, Fickian and constant diffusion, which differ in their motivation, derivation, and interpretation (Garlick et al., 2011). Mathematically, the difference between Fickian and constant diffusion is the position of the diffusion coefficient within (Fickian) or outside (constant) the partial derivatives in (6). Ecologically, Fickian diffusion describes a smooth process where the equilibrium state is a homogeneous (uniform) spatial distribution of $u(\mathbf{s}, t)$, even when the diffusion coefficient (defined as motion down density gradients) is spatially heterogeneous (see Fig. 1. in Garlick et al., 2011 or Hooten et al., 2013). In contrast, ecological diffusion describes the consequences of random walks with varying residence times and results in a spatially heterogeneous equilibrium distribution of $u(\mathbf{s}, t)$ that captures local variability and abrupt transitions.

From a Lagrangian (individual-based) perspective, we expect that the geographic spread in the prevalence of CWD among white-tailed deer is driven by the movement of individuals from a central location during the initial stages of an outbreak. For example, white-tailed deer demonstrate habitat preferences that, from a Eulerian (population-level) perspective, result in a heterogeneous geographic distribution in the abundance of deer due to variable residence times. When viewed from a Lagrangian perspective, the heterogeneous geographic distribution of a species is a result of individuals moving quickly through habitat of poor quality and congregating in more favorable habitat. During the initial outbreak of a disease such as CWD, the transmission and spread may be driven by the movement of individuals; thus, ecological diffusion (6) is a realistic model for a population-level spatio-temporal process representing the latent prevalence of CWD (Garlick et al., 2014).

3.3. Homogenization

Homogenization is a useful approach for ecological diffusion models because the implementation is conceptually simple and involves solving a piecewise constant diffusion PDE over a much coarser discretization (Garlick et al., 2011; Hooten et al., 2013). Following Hooten et al. (2013), homogenization for the ecological diffusion PDE in (6) results in

$$\frac{\partial}{\partial t} c(\mathbf{s}, t) = \overline{\delta}(\mathbf{s}) \left(\frac{\partial^2}{\partial s_1^2} + \frac{\partial^2}{\partial s_2^2} \right) c(\mathbf{s}, t), \tag{7}$$

where $c(\mathbf{s}, t)$ is the underlying process that results from diffusion with a locally constant diffusion rate $\overline{\delta}(\mathbf{s})$. The locally constant diffusion rate $\overline{\delta}(\mathbf{s})$ is

$$\overline{\delta}(\mathbf{s}) = \left(\frac{1}{|\mathcal{A}|} \int_{\mathcal{A}} \frac{1}{\delta(\mathbf{s})} d\mathbf{s} \right)^{-1}, \tag{8}$$

where $\mathcal{A}$ is a polygon (with area $|\mathcal{A}|$) at the broad-scale discretization and $\delta(\mathbf{s})$ is the diffusion coefficient in (6). Approximating the locally constant diffusion rate $\overline{\delta}(\mathbf{s})$ in (7) involves calculating

the harmonic mean of $\delta(\mathbf{s})$ using values contained within $\mathcal{A}$ from the fine-scale discretization. The $u(\mathbf{s}, t)$ from the original ecological diffusion PDE in (6) is related to $c(\mathbf{s}, t)$ in (7) by

$$u(\mathbf{s}, t) \approx \frac{c(\mathbf{s}, t)}{\delta(\mathbf{s})}. \tag{9}$$

3.4. Hierarchical Bayesian model

To understand the geographic pattern in the prevalence of CWD, we used the model:

$$y_i \sim \text{Bernoulli}\left(p(\mathbf{s}_i, t_i)\right) \tag{10}$$

$$g(p(\mathbf{s}_i, t_i)) = u(\mathbf{s}_i, t_i)e^{\mathbf{x}_i'\boldsymbol{\beta}} \tag{11}$$

$$\frac{\partial}{\partial t}u(\mathbf{s}, t) = \left(\frac{\partial^2}{\partial s_1^2} + \frac{\partial^2}{\partial s_2^2}\right)\delta(\mathbf{s})u(\mathbf{s}, t) \tag{12}$$

$$\log\left(\delta(\mathbf{s})\right) = \alpha_0 + \mathbf{z}(\mathbf{s})'\boldsymbol{\alpha}, \tag{13}$$

where y_i is 1 if the ith deer is CWD-positive and 0 otherwise. The probability that a deer is CWD-positive ($p(\mathbf{s}_i, t_i)$) depends on the spatio-temporal effect $u(\mathbf{s}_i, t_i)$ defined by the PDE in (12), and the effect of $\mathbf{x}_i$, a vector that includes observation-specific covariates sex and age of the tested deer. For the ecological diffusion PDE, $u(\mathbf{s}_i, t_i) \geq 0 \; \forall \; \mathbf{s}, \; t$; the quantity $e^{\mathbf{x}_i'\boldsymbol{\beta}}$ scales $u(\mathbf{s}_i, t_i)$ depending on characteristics of the individual deer. We used the multiplicative form in (11), rather than the more common additive form, to maintain the nonnegative nature of the latent prevalence process $u(\mathbf{s}, t)$. Scaling $u(\mathbf{s}_i, t_i)$ by a quantity that depends on covariates specific to the infected and non-infected individuals mimics the dynamics of CWD because individual-level risk factors should not influence the probability of infection unless the disease process has spread to the location where the individual is located (i.e., $u(\mathbf{s}_i, t_i) > 0$). The product, $u(\mathbf{s}_i, t_i)e^{\mathbf{x}_i'\boldsymbol{\beta}} \geq 0, \forall \; \mathbf{s}, \; t$, and therefore, the inverse of the link function $g(\cdot)$ must map the positive real line $[0, \infty)$ to the unit line $(0, 1)$. In our example, we used a cumulative normal distribution truncated below zero as a link function (i.e., truncated probit link function). The diffusion rate $\delta(\mathbf{s}) > 0, \forall \; \mathbf{s}, \; t$, which motivated the log link function in (13), and depends on an intercept term (α_0) and the location-specific covariate proportion deciduous forest cover ($z(\mathbf{s})$; Fig. 2(a)). For regression coefficients $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}$, we used the following priors: $\boldsymbol{\beta} \sim \text{N}(0, 10\mathbf{I})$ and $\boldsymbol{\alpha} \sim \text{N}(0, 10\mathbf{I})$. For the intercept term, we used the prior $\alpha_0 \sim \text{N}(0, 10)$.

At the boundary, we assume $u(\mathbf{s}, t) = 0, \forall \, t$. For the initial state, we assume

$$u(\mathbf{s}, t_0) = \begin{cases} \theta_j & \text{if } \mathbf{s} = \boldsymbol{\omega}_j \\ 0 & \text{if } \mathbf{s} \neq \boldsymbol{\omega}_j, \end{cases} \tag{14}$$

where θ_j is the unknown initial concentration of the process and $\boldsymbol{\omega}_j$ are the unknown coordinates of the jth location of introduction ($j = 1, 2, \ldots, J$). Ecologically, (14) assumes that CWD was introduced at J unknown locations within the study area at time t_0. For our data, we assumed $J = 1$ and assigned θ_1 the priors $\theta_1 \sim \text{TN}(0, 10^6)$ (TN refers to a normal distribution truncated below zero). We assigned $\boldsymbol{\omega}_1$ a conditional HPP prior (HPP refers to a homogeneous Poisson point process). A conditional HPP implies that the number of points J in (14) is known, thus the location density of the HPP is the necessary probability density function (Cressie, 1993, p. 651; Illian et al., 2008, p. 119–120). Ecologically, the conditional HPP prior is equivalent to assuming CWD was introduced with equal chances at J locations within the study area. In cases where J is unknown, a HPP prior could be specified and J could be estimated.

Solving the PDE in (12) also requires that the time since introduction is known (i.e., $t - t_0$); that is, when solving for $u(\mathbf{s}, t)$ in (12), we need to know the limits of integration. For most disease outbreaks, the time of introduction (t_0) is unknown, although prior knowledge may be available in some situations (see Discussion). In Appendix A, we show that the units associated with the diffusion rate are nonidentifiable because only the product $\delta(\mathbf{s})(t - t_0)$ is identifiable. As a result, the diffusion rate $\delta(\mathbf{s})$ should be interpreted as a relative rate, with the intercept (α_0) in (13) being unidentifiable,

but the proportional effect of covariates on the spatially-varying diffusion rate $\delta(\mathbf{s})$ being identifiable (i.e., $\boldsymbol{\alpha}$). Given that the intercept (α_0) is unidentifiable, we report $\delta(\mathbf{s})$ scaled such that the minimum value within the study area is one (Appendix A).

3.5. Implementation

The location accuracy of the observation (section level) results in a natural spatial grid with 2.6 km^2 cells within the study area (Fig. 1). Using traditional finite-difference methods to solve (12) results in a spatial grid with 10,000 cells and would be computationally challenging to implement (note that an enlarged study area of 25,900 km^2 was used for implementation of the finite-difference method; see Appendix B). Instead, we use homogenization and solved (7) over a spatial grid with 65 km^2 cells (400 total cells). The homogenization technique allows this solution to be analytically downscaled to the 2.6 km^2 cells using (9). We solved (7) to obtain $c(\mathbf{s}, t)$ using a finite-difference approximation with eight iterations and $\Delta t = 1$. For a fixed Δt, the number of iterations chosen to obtain $c(\mathbf{s}, t)$ in (7) using the finite-difference approximation implies a specific value for t_0, namely $t_0 = t - m\Delta t$ where m is the number of iterations. As a result, the number of iterations will be inversely proportional to the estimated value of α_0 in (13), which can be illustrated by substituting $m\Delta t$ for $t - t_0$ in (S1). To maximize computational efficiency, we chose the number of iterations of the finite-difference approximation of (7) to be as small as possible, but large enough so that the approximation was stable and convergent.

We implemented the model presented in (10)–(14) using a Markov Chain Monte Carlo algorithm (MCMC) coded in R (R Core Team, 2016, Appendix B). We obtained 400,000 samples from a single chain using our MCMC algorithm, resulting in an effective sample size of $>1{,}000$ for all parameters in the model (Givens and Hoeting, 2012). For this example, fitting the model at the 2.6 km^2 scale using homogenization required approximately 5 min to acquire 1,000 MCMC samples on a laptop computer with a 2.8 GHz quad-core processor, 16 GB of RAM, and optimized basic linear algebra subprograms. We include computational details for this example and the necessary computer code to reproduce all results and figures in program R (R Core Team, 2016, Appendix B).

3.6. Results

The prevalence of CWD was highest in the center of the study region and the relative rate of diffusion was influenced by the deciduous forest covariate (Figs. 2 and 3). The posterior mean of the coefficient for the deciduous forest covariate was -0.29 and the 95% equal-tailed credible interval (CI) was -0.46 to -0.11. The prevalence of CWD increased with age and within an age class was highest for male deer (Fig. 4). The posterior distribution of the location of introduction ($\boldsymbol{\omega}_1$) shows that the point of introduction likely occurred near the center of the study area (Fig. 4). The logarithmic score was -384. In the following section, we fit two alternative models to the data and compare the inference and predictions obtained using our approach to existing approaches.

4. Spatial confounding and comparison

For our CWD data example, one might consider using a binary GLMM with a Gaussian spatial random effect. Several studies in the applied literature have used a GLMM with a Gaussian spatial random effect to identify risk factors associated with the landscape that might influence the prevalence of CWD (e.g., Walter et al., 2011; Evans et al., 2016; Hefley et al., 2017a). For purposes of comparison, we fit a binary GLMM with a probit link that uses the stochastic PDE approach of Lindgren et al. (2011) and the generalization proposed by Ingebrigtsen et al. (2014). The spatial random effects of Lindgren et al. (2011) and Ingebrigtsen et al. (2014) can be viewed as the stationary (time-limiting) solution to the spatio-temporal stochastic PDE (see Hanks, in press)

$$\frac{\partial}{\partial t}u(\mathbf{s}, t) = \left(\frac{\partial^2}{\partial s_1^2} + \frac{\partial^2}{\partial s_2^2} - \kappa(\mathbf{s})\right)\tau(\mathbf{s})u(\mathbf{s}, t) + \mathcal{W}(\mathbf{s}), \tag{15}$$

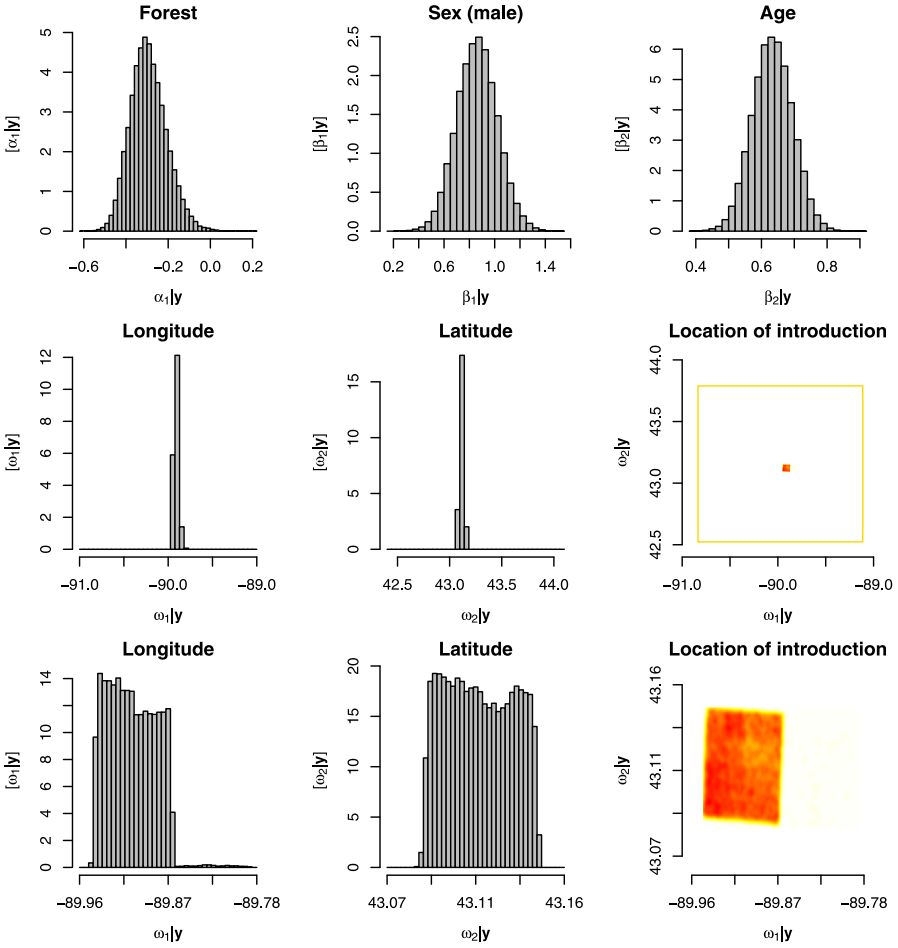

Fig. 3. Posterior distribution for parameters of the ecological diffusion model (10)–(14). The first row shows the posterior distribution for the effect for the proportion of deciduous forest covariate as well as the effect of sex and age of the deer. The bottom two rows show the posterior distribution of the location of introduction at two different spatial scales. The panels with colored heat maps are a kernel density smooth of the location of introduction where red represents area of higher density. Note that the yellow box corresponds to the study area in Fig. 1.

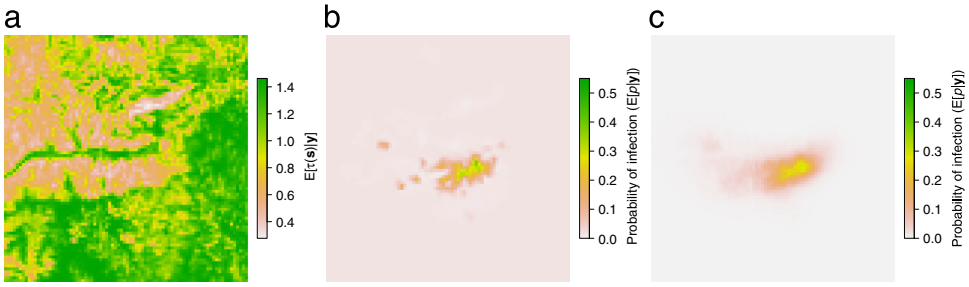

Fig. 4. Posterior expectation of $\tau(\mathbf{s})$ from (1)–(3), (15) and (17) that depends on the proportion of deciduous forest covariate (panel a in Fig. 2). Posterior expectation of the probability of chronic wasting disease infection in male white-tailed deer that are 4 years of age or older from a binary GLMM where the location-specific forest covariate was specified to influence the spatial random effect (Ingebrigtsen et al., 2014, panel b) and treated as a traditional covariate (Lindgren et al., 2011, panel c).

where $\kappa(\mathbf{s})$ is a scaling parameter, $\tau(\mathbf{s})$ is parameter that rescales the field $u(\mathbf{s}, t)$, and $\mathcal{W}(\mathbf{s})$ is spatial Gaussian white noise. As discussed in Ingebrigtsen et al. (2014), both $\kappa(\mathbf{s})$ and $\tau(\mathbf{s})$ could vary as a function of location-specific covariates (e.g., proportion deciduous forest cover). The scaling parameter $\kappa(\mathbf{s})$ can be linked to the range parameter ($\rho(\mathbf{s})$) of a Matérn correlation function using the empirically derived relationship $\rho(\mathbf{s}) \approx \frac{\sqrt{8}}{\kappa(\mathbf{s})}$. As shown by Ingebrigtsen et al. (2014), both $\kappa(\mathbf{s})$ and $\tau(\mathbf{s})$ control the non-stationary marginal variance of the Gaussian process, where the marginal variance ($\sigma^2(\mathbf{s})$) can be approximated with

$$\sigma^2(\mathbf{s}) \approx \frac{1}{4\pi\kappa(\mathbf{s})^2\tau(\mathbf{s})^2}. \tag{16}$$

We first consider a binary GLMM with a Gaussian random effect letting $\tau(\mathbf{s})$ in (15) be modeled as a log linear function

$$\log(\tau(\mathbf{s})) = \gamma_0 + \mathbf{z}(\mathbf{s})'\boldsymbol{\gamma}, \tag{17}$$

where γ_0 is an intercept term and $\mathbf{z}(\mathbf{s})$ is a vector that contains the location-specific covariates and $\boldsymbol{\gamma}$ is the associated coefficients. The linear predictor (2) of the GLMM contains the spatial random effect as well as the covariates sex and age of the tested deer. As in the ecological diffusion example (10)–(14), we used the location-specific covariate proportion deciduous forest cover ($z(\mathbf{s})$; Fig. 2(a)). In this model, $\kappa(\mathbf{s})$ was assumed to be spatially constant. Our motivation for fitting a model where $\tau(\mathbf{s})$ depended on the deciduous forest covariate and $\kappa(\mathbf{s})$ was constant is that this specification is similar to the ecological diffusion PDE in (6).

We also consider an alternative specification of the binary GLMM where $\tau(\mathbf{s})$ is constant and the deciduous forest covariate is included in $\mathbf{X}$ of (2). In this model specification, there is potential for collinearity between the location-specific covariate proportion deciduous forest and the spatial random effect. To assess the potential for collinearity, we calculated the Pearson product-moment correlation coefficient between the forest covariate and the posterior mean of the spatial random effect at the location of each observation (Hanks et al., 2015; Hefley et al., 2017a). To assess the predictive performance, we fit both models using 50% of the data and calculated the predictive score using the remaining 50% of the data (using the same partitions as in Section 3). We calculated the logarithmic score using the projected posterior mean of each variable and the CWD data (Krainski et al., 2016). We implemented the models of Lindgren et al. (2011) and Ingebrigtsen et al. (2014), presented in this section, using the R-INLA software (Appendix B; Rue et al., 2009).

4.1. Results

Using the GLMM specification from Ingebrigtsen et al. (2014), $\tau(\mathbf{s})$ in (15)–(17) depended on the proportion deciduous forest covariate and had coefficient (γ_1) with a posterior mean of −1.70 (−2.63, −0.86; 95% CI; Fig. 4(a)). The spatial map of $\tau(\mathbf{s})$ from (15)–(17) was similar in appearance to the spatial map $\delta(\mathbf{s})$ of our ecological diffusion model (12; cf. Figs. 2(b) and 4(a)). Using the GLMM specification from Lindgren et al. (2011), where location-specific covariates are included with the observation-specific covariates in $\mathbf{X}$ of (2), the posterior mean of the regression coefficient for the deciduous forest covariate was 0.45 (−0.27, 1.18; 95% CI; Fig. 4(a)) and the correlation between the forest covariate and spatial random effect was $r = 0.49$. The projected posterior expectation of the probability of CWD infection across the study area obtained from the Ingebrigtsen et al. (2014) and Lindgren et al. (2011) models are shown in Fig. 4(b and c). The logarithmic score for the GLMM with $\tau(\mathbf{s})$ in (15)–(17) that depended on the forest covariate (i.e., Ingebrigtsen et al., 2014) was −408 and −387 for the standard GLMM of Lindgren et al. (2011). Our ecological diffusion model (10)–(14) had the best predictive score of −384.

5. Discussion

Statisticians and researchers face many choices when specifying models for spatial data. One important choice, as illustrated in our study, is incorporating the spatio-temporal data generating

process into a model for spatial data. Although the inference for the location-specific covariate (proportion deciduous forest cover) from our ecological diffusion model cannot be compared directly to the stochastic PDE approach of Ingebrigtsen et al. (2014), if one were to infer an influence of the percent of deciduous forest cover on the spatio-temporal process (based on 95% CIs that did not overlap zero), then both models would result in a "significant" effect. Such inference is in contrast to the standard GLMM, which relies on treating observation-specific covariates (e.g., sex, age) and location-specific covariates similarly, where the 95% CI for the forest covariate overlapped zero. Based on the predictive score, our ecological diffusion model was the best predictive model with the stochastic PDE approach of Lindgren et al. (2011) a close second.

The term "spatial confounding" has been used to describe collinearity among covariates in $\mathbf{X}$ and the spatial random effect (2) in the spatial GLMM (Reich et al., 2006; Hodges and Reich, 2010). For the analysis of spatial data, confounding can present major challenges to inference and the interpretation of location-specific covariates (Hodges and Reich, 2010; Hanks et al., 2015). One approach that can obviate traditional spatial confounding is to use scientific knowledge of the process when specifying how location-specific covariates are incorporated into the model. Incorporating the spatio-temporal data generating process into our model yields a principled approach to determine how location-specific covariates should be included. Our approach, and that of Ingebrigtsen et al. (2014), allow the location-specific covariate to influence the spatio-temporal process directly and offer two viable options when considering model specifications other than the standard spatial GLMM.

The scientific approach we used to model the CWD data is borrowed from the approaches used for modeling spatio-temporal data (e.g., Wikle, 2003; Hooten and Wikle, 2008; Wikle and Hooten, 2010; Hefley et al., 2017b), but is rarely used for spatial data that arise as a single snapshot in time of a spatio-temporal data generating process. For spatial data, the spatio-temporal approach offers three main advantages over existing spatial models, while providing comparable or better predictive performance for our example. The first advantage of the scientific approach is that models, such as the ecological diffusion PDE in (6), have interpretable parameters related to the spatio-temporal process. For example, the diffusion rate in the ecological diffusion PDE has physically interpretable units that are inversely related to the residence time of a spreading process (e.g., transmittable disease; Garlick et al., 2011, 2014; Hefley et al., 2017b). Although the absolute diffusion rate is unidentifiable from purely spatial data, the proportional effect of covariates on the diffusion rate is identifiable and provides useful information about the process (Hanks, in press).

The second advantage of the spatio-temporal approach is that inference for the initial conditions can be obtained and may provide critical insight. For example, in our CWD data analysis, the posterior distribution of the location of introduction ($\boldsymbol{\omega}_1$ in (14)) showed that spatial data can provide precise estimates regarding where a disease was introduced. Estimating the location of introduction is critical in situations of disease outbreak or when attempting to halt the invasion of a pest species (e.g., Wikle, 2003; Zheng and Aukema, 2010). Inference about the initial conditions cannot be obtained when using standard spatial models that account for dependence by specifying a semiparametric random effect or when using the stationary solution to a stochastic PDE (e.g., Lindgren et al., 2011; Ingebrigtsen et al., 2014).

The third advantage of the spatio-temporal approach, is that with the use of informative priors, one can obtain dynamic forecasts rather than static spatial predictions (e.g., Figs. 2(c), 4(b) and (c)). For example, because α_0 in (13) is unidentifiable, the units associated with the diffusion rate are unidentifiable from spatial data without prior knowledge about the time of introduction (t_0). In many situations, experts may have knowledge about when a disease or pest was first introduced. Expert knowledge could be used to obtain informative priors (Kuhnert et al., 2005; Choy et al., 2009) or auxiliary data about the process (Lele and Allen, 2006), which could help identify the units associated with diffusion rate (e.g., using an informative prior for t_0). In ongoing research, we are incorporating the use of expert knowledge into the analysis of spatial data so that dynamic forecasts of the geographic spread of diseases such as CWD can be obtained.

Acknowledgments

We thank Chris Wikle, two anonymous reviewers, and the editor for their constructive comments. We thank the Wisconsin Department of Natural Resources for collecting deer tissue samples, the Wisconsin hunters who provided them, and Erin Larson for maintaining the CWD sample data base. The data set that contains the location and disease status of the white-tailed deer used in our example is owned by the Wisconsin Department of Natural Resources. Please contact Tamara Ryan (Tamara.Ryan@wisconsin.gov) to request access to the data used in our example. We acknowledge Dennis Heisey for his early contributions to development of this research endeavor. Funding for this project was provided by the USGS National Wildlife Health Center via grant USGS G14AC00366 and the National Science Foundation via grant DMS 1614392. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Appendix A

Consider a diffusion PDE from (6) with a spatially constant diffusion rate defined on the unit line $0 < s < 1$,

$$\frac{\partial}{\partial t}u(s,t) = \left(\frac{\partial^2}{\partial s^2}\right)\delta u(s,t). \tag{S1}$$

Specify the boundary conditions such that $u(0,t) = 0$ and $u(1,t) = 0, \forall\, t$ and the initial state $u(s,t_0)$ as a continuous function $f(s), \forall\, s$. Given the boundary conditions and initial state, the solution to (18; Farlow, 1993 pp. 35-41) is

$$u(s,t) = \sum_{n=1}^{\infty} A_n e^{-\delta(t-t_0)(n\pi)^2} \tag{S2}$$

where the coefficients A_n are given by

$$A_n = 2\int_0^1 f(s)\sin(n\pi s)ds.$$

The solution to (S1) depends on both the diffusion rate δ and $t - t_0$ (the time elapsed since introduction). Both δ and t_0 are unknown and only the product is identifiable from data collected at a single time point. Analogously, we expect that the diffusion rates from (12) will have unknown temporal scaling units and should be interpreted as a relative rate of diffusion. The assumption that $\delta(\mathbf{s})$ in (12) should be interpreted as a relative rate can be checked numerically by assessing the sensitivity of estimates of $\delta(\mathbf{s})$ in (12) to changes in t_0. For example, the map shown in Fig. 2(b) is invariant to choice of t_0, because α_0 in (13) is unidentifiable and we scaled $\delta(\mathbf{s})$ such that the minimum value within the study area is one.

Appendix B. Supplementary data

Supplementary material related to this article can be found online at http://dx.doi.org/10.1016/j.spasta.2017.02.005.

References

Besag, J., York, J., Mollié, A., 1991. Bayesian image restoration, with two applications in spatial statistics. Ann. Inst. Statist. Math. 43 (1), 1–20.

Bolker, B.M., Brooks, M.E., Clark, C.J., Geange, S.W., Poulsen, J.R., Stevens, M.H.H., White, J.S., 2009. Generalized linear mixed models: a practical guide for ecology and evolution. Trends Ecol. Evol. 24 (3), 127–135.

Calder, C.A., 2008. A dynamic process convolution approach to modeling ambient particulate matter concentrations. Environmetrics 19 (1), 39–48.

Cangelosi, A.R., Hooten, M.B., 2009. Models for bounded systems with continuous dynamics. Biometrics 65 (3), 850–856.

Choy, S.L., O'Leary, R., Mengersen, K., 2009. Elicitation by design in ecology: using expert opinion to inform priors for Bayesian statistical models. Ecology 90 (1), 265–277.
Cressie, N., 1993. Statistics for Spatial Data. John Wiley & Sons.
Cressie, N., Wikle, C.K., 2011. Statistics for Spatio-temporal Data. John Wiley & Sons.
Diggle, P.J., Tawn, J., Moyeed, R., 1998. Model-based geostatistics. J. R. Stat. Soc. Ser. C. Appl. Stat. 47 (3), 299–350.
Evans, T.S., Kirchgessner, M.S., Eyler, B., Ryan, C.W., Walter, W.D., 2016. Habitat influences distribution of chronic wasting disease in white-tailed deer. J. Wildl. Manage 80 (2), 284–291.
Farlow, S., 1993. Partial Differential Equations for Scientists and Engineers. Dover Publicaitons.
Garlick, M.J., Powell, J.A., Hooten, M.B., McFarlane, L.R., 2011. Homogenization of large-scale movement models in ecology. Bull. Math. Biol. 73 (9), 2088–2108.
Garlick, M.J., Powell, J.A., Hooten, M.B., MacFarlane, L.R., 2014. Homogenization, sex, and differential motility predict spread of chronic wasting disease in mule deer in southern Utah. J. Math. Biol. 69 (2), 369–399.
Givens, G.H., Hoeting, J.A., 2012. Computational Statistics. John Wiley & Sons.
Gladish, D., Wikle, C., Holan, S., 2014. Covariate-based cepstral parameterizations for time-varying spatial error covariances. Environmetrics 25 (2), 69–83.
Gneiting, T., 2011. Making and evaluating point forecasts. J. Amer. Statist. Assoc. 106 (494), 746–762.
Gneiting, T., Raftery, A.E., 2007. Strictly proper scoring rules, prediction, and estimation. J. Amer. Statist. Assoc. 102 (477), 359–378.
Gotway, C.A., Stroup, W.W., 1997. A generalized linear model approach to spatial data analysis and prediction. J. Agric. Biol. Environ. Stat. 2 (2), 157–178.
Hanks, E.M., 2017. Modeling spatial covariance using the limiting distribution of spatio-temporal random walks. J. Amer. Statist. Assoc. (in press).
Hanks, E.M., Schliep, E.M., Hooten, M.B., Hoeting, J.A., 2015. Restricted spatial regression in practice: geostatistical models, confounding, and robustness under model misspecification. Environmetrics 26 (4), 243–254.
Hefley, T.J., Hooten, M.B., Drake, J.M., Russell, R.E., Walsh, D.P., 2016. When can the cause of a population decline be determined? Ecol. Lett. 19 (11), 1353–1362.
Hefley, T.J., Hooten, M.B., Hanks, E.M., Russell, R.E., Walsh, D.P., 2017a. The Bayesian group lasso for confounded spatial data. J. Agric. Biol. Environ. Stat. 22 (1), 42–59.
Hefley, T.J., Hooten, M.B., Russell, R.E., Walsh, D.P., Powell, J., 2017b. When mechanism matters: Bayesian forecasting using models of ecological diffusion. Ecol. Lett. 20 (5), 640–650.
Hodges, J.S., Reich, B.J., 2010. Adding spatially-correlated errors can mess up the fixed effect you love. Amer. Statist. 64 (4), 325–334.
Holmes, M.H., 2012. Introduction to Perturbation Methods. Springer Science & Business Media.
Homer, C.G., Dewitz, J.A., Yang, L., Jin, S., Danielson, P., Xian, G., Coulston, J., Herold, N.D., Wickham, J.D., Megown, K., 2015. Completion of the 2011 National Land Cover Database for the conterminous United States-representing a decade of land cover change information. Photogramm. Eng. Remote Sens. 81 (5), 345–354.
Hooten, M.B., Garlick, M.J., Powell, J.A., 2013. Computationally efficient statistical differential equation modeling using homogenization. J. Agric. Biol. Environ. Stat. 18 (3), 405–428.
Hooten, M.B., Johnson, D.S., McClintock, B.T., Morales, J.M., 2017. Animal Movement: Statistical Models for Telemetry Data. CRC Press.
Hooten, M.B., 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. Environ. Ecol. Stat. 15 (1), 59–70.
Hooten, M.B., Wikle, C.K., 2010. Statistical agent-based models for discrete spatio-temporal systems. J. Amer. Statist. Assoc. 105 (489), 236–248.
Hooten, M.B., Wikle, C.K., Dorazio, R.M., Royle, J.A., 2007. Hierarchical spatiotemporal matrix models for characterizing invasions. Biometrics 63 (2), 558–567.
Hughes, J., Haran, M., 2013. Dimension reduction and alleviation of confounding for spatial generalized linear mixed models. J. R. Stat. Soc. Ser. B Stat. Methodol. 75 (1), 139–159.
Illian, J., Penttinen, A., Stoyan, H., Stoyan, D., 2008. Statistical Analysis and Modelling of Spatial Point Patterns. John Wiley & Sons.
Ingebrigtsen, R., Lindgren, F., Steinsland, I., 2014. Spatial models with explanatory variables in the dependence structure. Spat. Stat. 8, 20–38.
Joly, D.O., Samuel, M.D., Langenberg, J.A., Rolley, R.E., Keane, D.P., 2009. Surveillance to detect chronic wasting disease in white-tailed deer in wisconsin. J. Wildl. Dis. 45 (4), 989–997.
Krainski, E., Lindgren, F., Simpson, D., Rue, H., 2016. The R-INLA tutorial on SPDE models. http://www.math.ntnu.no/inla/r-inla.org/tutorials/spde/spde-tutorial.pdf.
Kuhnert, P.M., Martin, T.G., Mengersen, K., Possingham, H.P., 2005. Assessing the impacts of grazing levels on bird density in woodland habitat: a Bayesian approach using expert opinion. Environmetrics 16 (7), 717–747.
Lele, S.R., Allen, K.L., 2006. On using expert opinion in ecological analyses: a frequentist approach. Environmetrics 17 (7), 683–704.
Lindgren, F., Rue, H., Lindström, J., 2011. An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. J. R. Stat. Soc. Ser. B Stat. Methodol. 73 (4), 423–498.
Murakami, D., Griffith, D.A., 2015. Random effects specifications in eigenvector spatial filtering: a simulation study. J. Geogr. Syst. 17 (4), 311–331.
Nash, D., Mostashari, F., Fine, A., Miller, J., O'Leary, D., Murray, K., Huang, A., Rosenberg, A., Greenberg, A., Sherman, M., et al., 2001. The outbreak of West Nile virus infection in the New York City area in 1999. New Engl. J. Med. 344 (24), 1807–1814.
Paciorek, C., 2010. The importance of scale for spatial-confounding bias and precision of spatial regression estimators. Statist. Sci. 25 (1), 107–125.
Poppick, A., Stein, M.L., 2014. Using covariates to model dependence in nonstationary, high-frequency meteorological processes. Environmetrics 25 (5), 293–305.
Powell, J.A., Bentz, B.J., 2014. Phenology and density-dependent dispersal predict patterns of mountain pine beetle (*Dendroctonus ponderosae*) impact. Ecol. Model. 273, 173–185.

R Core Team 2016. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, URL https://www.R-project.org/.
Reich, B.J., Eidsvik, J., Guindani, M., Nail, A.J., Schmidt, A.M., 2011. A class of covariate-dependent spatiotemporal covariance functions. Ann. Appl. Stat. 5 (4), 2265–2687.
Reich, B.J., Hodges, J.S., Zadnik, V., 2006. Effects of residual smoothing on the posterior of the fixed effects in disease-mapping models. Biometrics 62 (4), 1197–1206.
Risser, M.D., Calder, C.A., 2015. Regression-based covariance functions for nonstationary spatial modeling. Environmetrics 26 (4), 284–297.
Rue, H., Martino, S., Chopin, N., 2009. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. J. R. Stat. Soc. Ser. B Stat. Methodol. 71 (2), 319–392.
Smith, D.L., Lucey, B., Waller, L.A., Childs, J.E., Real, L.A., 2002. Predicting the spatial dynamics of rabies epidemics on heterogeneous landscapes. Proc. Natl. Acad. Sci. 99 (6), 3668–3672.
Snow, J., 1855. On the Mode of Communication of Cholera. John Churchill.
Stein, M.L., 1999. Interpolation of Spatial Data: Some Theory for Kriging. Springer Science & Business Media.
Turchin, P., 1998. Quantitative Analysis of Movement: Measuring and Modeling Population Redistribution in Animals and Plants. Sinauer Associates Sunderland.
Vianna Neto, J.H.V., Schmidt, A.M., Guttorp, P., 2014. Accounting for spatially varying directional effects in spatial covariance structures. J. R. Stat. Soc. Ser. C. Appl. Stat. 63 (1), 103–122.
Waller, L.A., Carlin, B.P., Xia, H., Gelfand, A.E., 1997. Hierarchical spatio-temporal mapping of disease rates. J. Amer. Statist. Assoc. 92 (438), 607–617.
Waller, L.A., Gotway, C.A., 2004. Applied Spatial Statistics for Public Health Data. John Wiley Sons.
Walter, W.D., Walsh, D.P., Farnsworth, M.L., Winkelman, D.L., Miller, M.W., 2011. Soil clay content underlies prion infection odds. Nature Commun. 2, 200.
Wheeler, D.C., Waller, L.A., 2008. Mountains, valleys, and rivers: the transmission of raccoon rabies over a heterogeneous landscape. J. Agric. Biol. Environ. Stat. 13 (4), 388–406.
Whittle, P., 1954. On stationary processes in the plane. Biometrika 41 (3/4), 434–449.
Whittle, P., 1963. Stochastic-processes in several dimensions. Bull. Int. Stat. Inst. 40 (2), 974–994.
Wikle, C.K., 2003. Hierarchical Bayesian models for predicting the spread of ecological processes. Ecology 84 (6), 1382–1394.
Wikle, C.K., Hooten, M.B., 2010. A general science-based framework for dynamical spatio-temporal models. Test 19 (3), 417–451.
Wikle, C.K., Milliff, R.F., Nychka, D., Berliner, L.M., 2001. Spatiotemporal hierarchical Bayesian modeling tropical ocean surface winds. J. Amer. Statist. Assoc. 96 (454), 382–397.
Williams, P.J., Hooten, M.B., Womble, J.N., Esslinger, G.G., Bower, M.R., Hefley, T.J., 2017. An integrated data model to estimate spatio-temporal occupancy, abundance, and colonization dynamics. Ecology 98 (2), 328–336.
Williams, E., Young, S., 1980. Chronic wasting disease of captive mule deer: a spongiform encephalopathy. J. Wildl. Dis. 16 (1), 89–98.
Zheng, Y., Aukema, B.H., 2010. Hierarchical dynamic modeling of outbreaks of mountain pine beetle using partial differential equations. Environmetrics 21 (7–8), 801–816.