Spatial Statistics 14 (2015) 382–399

Contents lists available at ScienceDirect

Spatial Statistics

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

Transformed Gaussian Markov random fields and spatial modeling of species abundance

Marcos O. Prates [a,*], Dipak K. Dey [b,c,d], Michael R. Willig [c,e], Jun Yan [b,c,d]

[a] Department of Statistics, Universidade Federal de Minas Gerais, Avenida Antôni Carlos, 6627, Belo Horizonte, Brazil

[b] Department of Statistics, University of Connecticut, Storrs, CT 06269, USA

[c] Center for Environmental Sciences & Engineering, University of Connecticut, Storrs, CT 06269, USA

[d] Institute for Public Health Research, University of Connecticut Health Center, East Hartford, CT 06108, USA

[e] Department of Ecology & Evolutionary Biology, University of Connecticut, Storrs, CT 06269, USA

ARTICLE INFO

Article history:
Received 1 August 2014
Accepted 11 July 2015
Available online 29 July 2015

Keywords:
Bayesian inference
Beta field
Gamma field
Gaussian copula
Generalized linear mixed model

ABSTRACT

Gaussian random field and Gaussian Markov random field have been widely used to accommodate spatial dependence under the generalized linear mixed models framework. To model spatial count and spatial binary data, we present a class of transformed Gaussian Markov random fields, constructed by transforming the margins of a Gaussian Markov random field to desired marginal distributions that accommodate asymmetry and heavy tail, as needed in many empirical circumstances. The Gaussian copula that characterizes the dependence structure facilitates inferences and applications in modeling spatial dependence. This construction leads to new models such as gamma or beta Markov fields with Gaussian copulas, that are used to model Poisson intensities or Bernoulli rates in hierarchical spatial analyses. The method is naturally implemented in a Bayesian framework. To illustrate our methodology, abundances of variety of gastropod species were collected as counts or presence versus absence from a network of spatial locations in the Luquillo Mountains of Puerto Rico. Gastropods are of considerable ecological importance in terrestrial ecosystems because of their species richness, abundances, and critical roles in ecosystem processes such as decomposition and nutrient cy-

* Corresponding author.
E-mail address: marcosop@est.ufmg.br (M.O. Prates).

http://dx.doi.org/10.1016/j.spasta.2015.07.004

cling. The new models outperform the traditional models based on Bayesian model comparison with conditional predictive ordinate. The validity of Bayesian inferences and model selection were assessed through simulation studies for both spatial Poisson regression and spatial Bernoulli regression.

1. Introduction

Spatial count or binary data are generally analyzed with a generalized linear mixed model (GLMM), where spatial dependence is captured by Gaussian random field (GRF) effects (e.g., Breslow and Clayton, 1993). When data are point-referenced or geostatistical, and prediction at unobserved sites is of main concern, Diggle et al. (1998) extended the kriging method to the spatial GLMM (SGLMM) with GRF random effects to predict the surface of the spatial random effects. Under this scheme, Christensen and Waagepetersen (2002) developed predictions for the count of weeds at unobserved sites over a region. For lattice or areal data, as is the case in our application, a Markov dependence with an appropriate neighborhood structure is often imposed on the GRF random effects, which offers both intuitive interpretation and computational advantages. A Gaussian Markov random field (GMRF) is represented by an undirected graph, and is more naturally defined through its precision matrix. The (i, j)th entry of the precision matrix is nonzero if and only if i and j are connected in the graph (Rue and Held, 2005). GLMMs with random effects of GMRF have been used in many fields. Because of the public concerns regarding global change and public health, recent applications have surged in environmental sciences (e.g., Wikle et al., 1998; Rue et al., 2004) and epidemiology (e.g., Besag et al., 1991; Schmid and Held, 2004).

We propose a hierarchical spatial generalized linear model (GLM) that is subtly different from the GLMM with GRF random effects. At the first level, the observed data are independent Poisson or Bernoulli variables given the Poisson intensities or Bernoulli rates. At the second level, the Poisson intensities or Bernoulli rates are modeled by a transformed GRF (TGRF) such that the marginal distributions are of any desired form. Similarly, a transformed GMRF (TGMRF) can be defined if the GRF is a GMRF, and the Markov property is retained regardless of the transformations. With gamma or beta margins, this leads to gamma fields or beta fields for modeling Poisson intensities or Bernoulli rates, respectively. Our specification offers new avenues to construct hierarchical spatial GLMs and a fresh look at common SGLMMs with GRF random effects. Clearly, the new framework will facilitate the definition of an adequate marginal distribution for the mean parameters that is not necessarily a simple task in the conditional modeling framework. Moreover, the dependence structure is kept unchanged in the TGMRF because of the use of the Gaussian copula and, therefore, the interpretation of the β parameters are kept unchanged. A limitation of the new methodology in comparison to the traditional conditional approach is that, although it can be done, extension of the model to include more random effects, e.g. temporal effects, is not as trivial as it is in additive models. Inferences are conducted in the Bayesian framework with a general purpose, easy-to-implement Gibbs sampling algorithm.

The essence of TGRF or TGMRF is the Gaussian copula (Nelsen, 2006; Song, 2000; Masarotto and Varin, 2012), which has been used under other terminologies in various contexts. For multivariate data, it is equivalent to the Gaussian copula regression model (Pitt et al., 2006), where the response vector may be a combination of discrete and continuous variables. Under a graphical model framework it is similar to the copula Gaussian graphical model of Dobra and Lenkoski (2011), where the dependence structure determined by the precision matrix is of specific interest. In some fields such as hydrology, it is named as meta-Gaussian distribution (e.g., Guillot and Lebel, 1999; Schaake et al., 2007). In geostatistics with point-referenced data, it is called the anamorphosis Gaussian field (Chilès and Delfiner, 1999) or Gaussian copula model (Bárdossy, 2006; Kazianka and Pilz, 2010), where the main interest of these models has being interpolation and prediction at unmeasured locations. For this setup, a Matlab toolbox implementation is available (Kazianka, 2013). Most geostatistical applications

deal with continuous variables, but adaptation to discrete data is possible (Madsen, 2009). Berrocal et al. (2008) used it with binary and gamma margins, respectively, to construct random fields for precipitation occurrence and precipitation amount. For lattice or areal data, which is the context of our application, the TGRF or TGMRF is related to the general Gaussian graphical model of Dobra et al. (2011), except that they make inferences about the graph structure. We apply the TGMRF to model parameters that are continuous, to avoid the complexities associated with discrete data (Genest and Neslehova, 2007); this is in contrast to the existing works where it is applied to the data directly. Another important advantage of the new formulation is that the dependence parameters do not interfere with the marginal models. This is in contrast to traditional SGLMMs, where the spatial random effects are confounded with the fixed effects (Reich et al., 2006). Therefore, more reliable inferences about the regression coefficients can be made, which is critical in identifying important explanatory variables in the presence of spatial variation in our application.

To illustrate our methodology we present an ecological study. Understanding the causes of variation in species abundances is a central concern of ecology, conservation biology, and biodiversity science. Moreover, distinguishing pure environmental effects from those with a strong spatial signature has received increased interest from both theoretical and applied perspectives (e.g. Peres-Neto et al., 2012). In this context, gastropods are of considerable ecological importance in terrestrial ecosystems because of their species richness, abundance, and critical roles in ecosystem processes such as decomposition and nutrient cycling (Mason, 1970). The forest ecosystems of the Luquillo Mountains of Puerto Rico have a long history of environmental study (e.g., Brown et al., 1983; Reagan and Waide, 1996), resulting in deep understanding of the spatial and temporal dynamics of populations, communities, and biogeochemical processes, especially as they relate to natural and human disturbances (Brokaw et al., 2012). In the Luquillo Mountains, abundance data of various gastropod species were collected in wet seasons over a lattice of sites known as the Luquillo Forest Dynamics Plot (LFDP). Among the gastropod taxa, *Nenia tridens* and *Gaeotis nigrolineata* are two common terrestrial species (Willig et al., 1998; Bloch and Willig, 2006; Willig et al., 2011). *N. tridens* is one of the most abundant and widely distributed species in tabonuco forest, with abundance data in count format. *G. nigrolineata*, however, often occurs in low numbers that are more suitable for analysis as presence/absence data. The main goal of this project was to find habitat characteristics that affect abundances of different gastropod species in the presence of spatial dependence and variation.

The rest of this article is organized as follows. The sampling design and abundance data for two gastropod species, *N. tridens* and *G. nigrolineata*, are introduced in Section 2. In Section 3, TGMRFs are applied to the parameters of marginal distributions in a hierarchical spatial GLM framework to accommodate spatial dependence; the general framework is then applied to the specific context of Poisson regression and Bernoulli regression. Computational issues of the Bayesian inference and model selection are summarized in Section 4. Simulation studies mimicking the empirical count data and binary data are reported in Section 5. The abundance data are analyzed in Section 6. A discussion concludes the paper in Section 7.

2. Gastropods abundance data

The LFDP ($18^\circ\,20'$ N, $65^\circ\,49'$ W) is a 16-ha grid located in the northwest of the Luquillo Experimental Forest (LEF) in the Luquillo Mountains of northeastern Puerto Rico (e.g., Willig et al., 1998). The LEF includes tabonuco forest, a subtropical wet forest type (Ewel and Whitmore, 1973) found below 600 m of elevation. Precipitation is substantial throughout the year. A modestly drier period typically extends from January to April, but rainfall generally remains higher than 20 cm in all months (Brown et al., 1983). Abundances of gastropod species were censused during the wet season of 1995 at each of 160 circular sites (3 m radius) on a lattice. As shown in Fig. 1, there were 40 major sites in dark, 60 m apart, and 120 supplementary sites in gray, 20 m apart, placed inside the squares formed by the 40 major sites.

The abundance data are in count format for *N. tridens*, one of the most abundant and widely distributed species in tabonuco forest. Each count was the minimum number known alive from four nocturnal surveys based on well established protocols (e.g., Willig et al., 1998; Bloch and Willig, 2006).

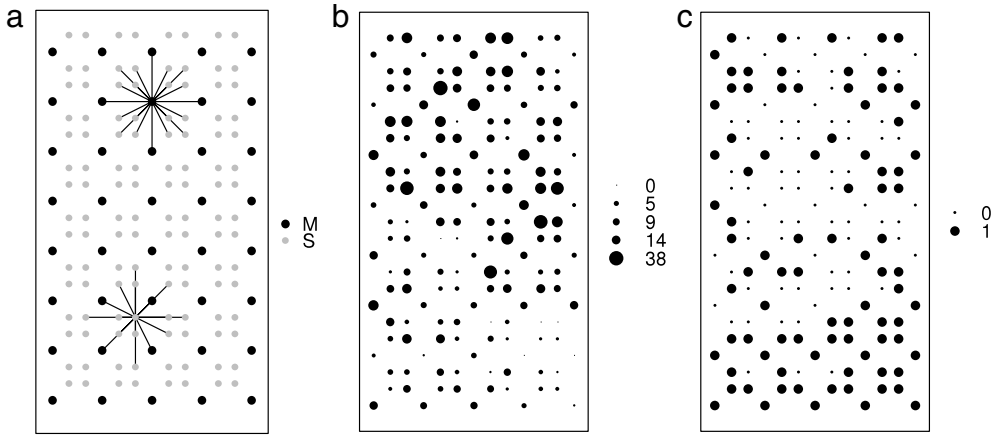

Fig. 1. (a) Lattice of sampling sites at the LFDP in 1995 and the neighbor structure for an internal major site and an internal supplementary site, labeled by M and S, respectively. (b) Abundance of *N. tridens*. (c) Presence/absence of *G. nigrolineata*.

The observed counts over the lattice are displayed in Fig. 1(b). Covariates included topographic (i.e., elevation and slope) and habitat characteristics (i.e., quantity of litter, canopy openness, apparency of sierra palm, and plant apparency). Quantity of litter was estimated as the mean number of leaves on the forest floor at each of four locations that were sampled at each site along mid-points of the radii from the center of a site, arranged along cardinal compass directions (cardinal points). Canopy openness was the amount of light that penetrates to the understory (1.5 m above the forest floor) based on the mean number of open cross-hairs on a gridded densiometer, quantified at the four cardinal points. Plant apparency measured the volume of space in the understory that was occupied by vegetation using a device at each of the four cardinal points, that captured the number of foliar intercepts along each of two perpendicular 1.0 m dowels placed at 0.5 m intervals from ground level to 3 m of height. Apparency of sierra palm specifically measured the apparency of *Prestoea acuminata*, a preferred substrate and food of both *N. tridens* and *G. nigrolineata*.

G. nigrolineata often occurs in low numbers and its abundance is suitably analyzed as presence/absence data. Presence/absence data were obtained by dichotomizing the abundance of *G. nigrolineata*, which were determined in the same manner as described for *N. tridens*. The distribution of incidences for *G. nigrolineata* is heterogeneous with spatial clustering across the lattice (Fig. 1(c)). Because *G. nigrolineata* does not live or feed in the leaf litter, quantity of litter will not be included as a covariate in analyses of its abundance but all other habitat characteristics were retained.

3. A new class of hierarchical spatial models

3.1. Transformed Gaussian Markov random fields

For ease of notation, we present the definition of TGRF and TGMRF in the context of finite dimension n in the sequel. For random fields indexed by elements in some space, the definition applies to n-dimensional marginal distributions for any n. Suppose that $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_n)^\top$ follows a standard multivariate normal distribution with mean $\mathbf{0}$ and positive definite correlation matrix, $\boldsymbol{\Psi}$, denoted as $N_n(\mathbf{0}, \boldsymbol{\Psi})$. Define a random vector $\boldsymbol{Z} = (Z_1, \ldots, Z_n)^\top$ through $Z_i = F_i^{-1}\{\Phi(\varepsilon_i)\}$, $i = 1, \ldots, n$, where F_i is the distribution function of an absolutely continuous variable and Φ is the distribution function of $N(0, 1)$. Then, each Z_i has a marginal distribution F_i. We call $\boldsymbol{Z}$ a TGRF and denoted by $\text{TGRF}_n(\boldsymbol{F}, \boldsymbol{\Psi})$, where $\boldsymbol{F} = (F_1, \ldots, F_n)$. It is completely specified by its marginal distributions $\boldsymbol{F}$ and a Gaussian copula with a dispersion matrix $\boldsymbol{\Psi}$. Note that $\boldsymbol{\Psi}$ is not the correlation matrix of $\boldsymbol{Z}$ and that the TGRF is not affected by the scales of the original GRF.

A $\text{TGRF}_n(\boldsymbol{F}, \boldsymbol{\Psi})$ is a TGMRF if the GRF before the transformations is a standard GMRF with correlation matrix $\boldsymbol{\Psi}$. As commonly used for GMRFs, it is more convenient to present a TGMRF using the precision matrix $\boldsymbol{Q} = \boldsymbol{\Psi}^{-1}$, since it leads to an intuitive interpretation of conditional distributional properties. Since the transformations are marginal-wise, the Markov property is inherited by the TGMRF: for $i \neq j$, $Z_i \perp Z_j | \boldsymbol{Z}_{(-ij)}$ if and only if $Q_{ij} = 0$, where $\boldsymbol{Z}_{(-ij)}$ is $\boldsymbol{Z}$ without the ith and jth observations (Rue and Held, 2005). The sparseness of the precision matrix completely determines the graph of the TGMRF and, hence, its conditional dependence structure. For convenience, we denote a TGMRF with marginal distributions $\boldsymbol{F}$ and precision matrix $\boldsymbol{Q}$ in the original scale by $\text{TGMRF}_n(\boldsymbol{F}, \boldsymbol{Q})$. Matrix $\boldsymbol{Q}$ is, again, not to be interpreted as precision on the post-transformation scale; we simply call it dependence matrix. As copulas are invariant with respect to scales, we constrain $\boldsymbol{Q}$ such that $\boldsymbol{Q}^{-1}$ is a correlation matrix for identifiability.

3.2. Hierarchical spatial model

Consider the traditional GLMM with a GRF random effect. Suppose that we observe $(Y_i, \boldsymbol{X}_i)$ at sites $i = 1, \ldots, n$, where Y_i is the response variable and $\boldsymbol{X}_i$ a $q \times 1$ covariate vector. Let $\boldsymbol{Y} = (Y_1, \ldots, Y_n)$ and $\boldsymbol{X} = (\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n)^\top$. Let $\boldsymbol{e} = (e_1, \ldots, e_n)^\top$ be a vector of unobserved random effects with joint distribution H, which introduces spatial dependence. A spatial GLMM assumes that, given $(\boldsymbol{X}_i, e_i)$, $i = 1, \ldots, n$, the observations Y_i's are independent with a distribution from the exponential family with mean $\mu_i = E(Y_i | \boldsymbol{X}, \boldsymbol{e})$. The conditional expectation μ_i is connected to the covariate $\boldsymbol{X}_i$ and random effect e_i through a fixed link function g, $g(\mu_i) = \eta_i + e_i$, where $\eta_i = \boldsymbol{X}_i^\top \boldsymbol{\beta}$ is the fixed effect, and $\boldsymbol{\beta}$ is a $q \times 1$ vector of regression coefficients of covariates $\boldsymbol{X}_i$. The dependence among $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)^\top$ is determined by the link function g and the joint distribution H of $\boldsymbol{e}$. With GRF random effect, H is the multivariate normal distribution with mean zero and covariance matrix $\boldsymbol{\Sigma}$.

We propose to specify a random field directly for $\boldsymbol{\mu}$:

$$\boldsymbol{\mu} \sim \text{TGRF}_n(\boldsymbol{F}, \boldsymbol{\Psi}), \tag{1}$$

where $\boldsymbol{F} = (F_1, \ldots, F_n)$, F_i is the marginal distribution of μ_i, and $\boldsymbol{\Psi}$ is the dispersion matrix characterizing the dependence structure of the underlying Gaussian copula. The marginal distribution F_i is specified by linear predictor η_i and possibly some other parameter ν shared across $i = 1, \ldots, n$. If F_i is chosen to be the distribution function of $\mu_i = g^{-1}(\eta_i + e_i)$, $i = 1, \ldots, n$, where $\boldsymbol{e}$ is multivariate normal with mean zero and correlation matrix $\boldsymbol{\Psi}$, then model (1) is the same as the GLMM of Diggle et al. (1998) with link function g and H being the distribution function of $N_n(\boldsymbol{0}, \sigma^2 \boldsymbol{\Psi})$ for some scale parameter σ. The TGRF specification with desired marginal distributions provides random field models such as gamma field, beta field, and their Markov versions, which can be incorporated into a spatial hierarchical GLM framework. For instances, one can use gamma margins for Poisson intensities and beta margins for Bernoulli rates, that can in turn be used, respectively, to model spatial count data or spatial binary data. The spatial dependence is completely characterized by the Gaussian copula, parameterized by a dispersion matrix $\boldsymbol{\Psi}$.

Assuming a Markov structure, we can replace the TGRF in model (1) with a TGMRF

$$\boldsymbol{\mu} \sim \text{TGMRF}_n(\boldsymbol{F}, \boldsymbol{Q}), \tag{2}$$

where the spatial dependence is characterized by $\boldsymbol{Q}$, the precision matrix of the Gaussian copula. For identifiability, we want the dependence matrix $\boldsymbol{Q}$ to be scale-free in the sense that $\boldsymbol{Q}^{-1}$ is a valid correlation matrix. We propose to parametrize $\boldsymbol{Q}$ with the "structure" of the precision matrix of a conditional autoregressive (CAR) model (Besag, 1974). In a CAR specification with precision matrix $\boldsymbol{\Omega}/\sigma^2$, the structure $\boldsymbol{\Omega}$ is defined such that $\boldsymbol{\Omega}_{ij}$ is nonzero if and only if site i and site j are neighbors of each other. To assure symmetry and positive definiteness, $\boldsymbol{\Omega}$ is defined as $\boldsymbol{\Omega} = \boldsymbol{M}^{-1}(\boldsymbol{I} - \rho \boldsymbol{W})$, where $\boldsymbol{M}^{-1}$ is a diagonal matrix whose ith diagonal elements equal to n_i, the number of neighbors of site i, $\boldsymbol{I}$ is the identity matrix, ρ is a spatial dependence parameter, and $\boldsymbol{W}$ is a weight matrix providing contrasts of all neighbors to each site. Weight matrix $\boldsymbol{W}$ is determined by the neighboring structure such that $W_{ij} = 1/n_i$ if site i and site j are neighbors, and zero otherwise.

Nevertheless, $\boldsymbol{\Omega}^{-1}$ is not a correlation matrix; its diagonals are in general not one and unequal. How do we use the "structure" of $\boldsymbol{\Omega}$ to define $\boldsymbol{Q}$? Since copulas are scale invariant, we could always

obtain $\boldsymbol{\Omega}^{-1}$ first and then standardize it to make all the diagonals one. Specifically, let γ_i^2 be the ith diagonal element of $\boldsymbol{\Omega}^{-1}$ and let $\boldsymbol{V} = \text{diag}\,(\gamma_1^2, \ldots, \gamma_n^2)$. We could parametrize $\boldsymbol{Q}$ by

$$\boldsymbol{Q} = \boldsymbol{V}^{1/2}\boldsymbol{\Omega}\boldsymbol{V}^{1/2}. \tag{3}$$

This way, $\boldsymbol{Q}$ is a valid dependence matrix whose inverse, $\boldsymbol{V}^{-1/2}\boldsymbol{\Omega}^{-1}\boldsymbol{V}^{-1/2}$, is a correlation matrix. The only dependence parameter in this specification is ρ. In our implementation of a Gibbs sampling algorithm with site-wise updating, the rescaling is not needed in the calculation.

In summary, a hierarchical model with a TGMRF component for the observed response Y_i, $i = 1, \ldots, n$, is of the form

$$\begin{aligned} Y_i|\mu_i &\sim \pi\,(y|\mu_i), \quad i = 1, \ldots, n, \\ \boldsymbol{\mu} &\sim \text{TGMRF}_n(\boldsymbol{F}_{\boldsymbol{\beta},\nu,\boldsymbol{X}}, \boldsymbol{Q}_\rho), \end{aligned}$$

where $\pi\,(y|\mu_i)$ is the distribution of Y_i given μ_i, $\boldsymbol{F}_{\boldsymbol{\beta},\nu,\boldsymbol{X}}$ is specified by η_i and possibly another parameter vector ν, and $\boldsymbol{Q}_\rho$ is specified by a neighboring structure and a dependence parameter ρ. The model parameters of interest are $\boldsymbol{\theta} = (\boldsymbol{\beta}, \nu, \rho, \boldsymbol{\mu})$.

3.3. Spatial poisson regression with gamma field

Consider count data observed at n sites in a spatial domain. Let Y_i be the count at site i, and with a covariate vector $\boldsymbol{X}_i$, $i = 1, \ldots, n$. Poisson models are widely used for count data and the Poisson intensities are often modeled by gamma distributions. Few choices of gamma fields are available in the literature. An exception is Wolpert and Ickstadt (1998), where a doubly stochastic process is used to construct positively auto-correlated intensity measures for spatial Poisson point processes which are then used to model spatial count data. The TGMRF model provides a gamma Markov random field that can be used as intensities of Poisson count in a hierarchical model.

A GLMM introduces spatial dependence through a spatial random effect. Conditioning on $\boldsymbol{\mu}$, the observed spatial count data Y_i's are assumed to be independent, and each Y_i is Poisson with mean μ_i, $i = 1, \ldots, n$. The most commonly used GLMM for spatial count data uses the canonical log link on the Poisson intensities:

$$\log \mu_i = \boldsymbol{X}_i^\top \boldsymbol{\beta} + e_i, \tag{4}$$

where $\boldsymbol{e}$ follows a CAR model with precision matrix $\boldsymbol{\Omega}/\nu$, $\nu > 0$. Let the ith component of $\boldsymbol{F}_{\boldsymbol{\beta},\nu,\boldsymbol{X},\rho}$ be the distribution function of

$$\text{LN}(\boldsymbol{X}_i^\top \boldsymbol{\beta}, \quad \nu\gamma_i^2), \quad \nu > 0,\ i = 1, \ldots, n, \tag{5}$$

where $\text{LN}(a, b)$ denotes a log-normal distribution with mean a and variance b on the log scale. Then model (4) could be equivalently specified by (2) with $\boldsymbol{F} = \boldsymbol{F}_{\boldsymbol{\beta},\nu,\boldsymbol{X},\rho}$ and $\boldsymbol{Q} = \boldsymbol{Q}_\rho$ given by (3). Note that the distributional parameters of the spatial random effects ν and ρ both affect the marginal distribution $\boldsymbol{F}$.

The TGMRF framework provides a new way to construct models for $\boldsymbol{\mu}$ that incorporate spatial dependence and covariates. The Gaussian copula of TGMRFs captures the spatial dependence. Any positive continuous distribution can be used to specify the marginal distribution of $\boldsymbol{\mu}$, and covariate effects can be accommodated into its parameters. Changing $\boldsymbol{F}$ in model (2) from log-normal to other distribution functions with positive support leads to new models. Gamma distribution is a natural choice for the margins. Covariates can be incorporated into either one of the two parameters, resulting in two different gamma models as long as there is at least one covariate. The gamma scale model, hereafter the GSC model, incorporates covariates into the scale parameter and defines the marginal distribution F_i as

$$\Gamma\left(1/\nu, \quad \nu \exp(\boldsymbol{X}_i^\top \boldsymbol{\beta})\right), \quad \nu > 0,\ i = 1, \ldots, n. \tag{6}$$

The gamma shape model, hereafter the GSH model, incorporates covariates into the shape parameter and defines the marginal distribution F_i as

$$\Gamma\left(\exp(\boldsymbol{X}_i^\top \boldsymbol{\beta})/\nu, \quad \nu\right), \quad \nu > 0,\ i = 1, \ldots, n. \tag{7}$$

Under both models, the expectation of μ_i is the same, $\exp(\boldsymbol{X}_i^\top \boldsymbol{\beta})$, but the parameter ν has different interpretations and should not be compared directly. TGMRF models with other marginal distribution for μ_is can be constructed similarly.

There is an additional subtle difference between the log-normal model (5), hereafter the LN model, and the two gamma models (6) and (7). Unlike the gamma models, where the dependence structure does not interfere with the marginal models, the dependence structure $\boldsymbol{\Omega}$ enters the marginal distributions of μ_is through γ_i^2 in the LN model. This difference makes the new formulation more attractive in the sense that the marginal parameters and the dependence parameters are easier to identify, avoiding the confounding between spatial random effects and fixed effects in the traditional model as pointed out in Reich et al. (2006). Consequently, more precise inference about the regression coefficients is possible.

3.4. Spatial Bernoulli regression with beta field

Consider presence/absence data at n sites in a spatial domain. Let Y_i be 1 if presence is observed and 0 otherwise at site i, with a covariate vector $\boldsymbol{X}_i$, $i = 1, \ldots, n$. Conditioning on $\boldsymbol{\mu}$ the observed data Y_i's are assumed to be independent, and each Y_i is Bernoulli with mean μ_i, $i = 1, \ldots, n$. The traditional spatial GLMM for binary data is

$$\text{logit}(\mu_i) = \boldsymbol{X}_i^\top \boldsymbol{\beta} + e_i, \tag{8}$$

where β is the regression coefficient vector, $\boldsymbol{e}$ follows a CAR model with mean zero and precision matrix $\boldsymbol{\Omega}/\nu$, $\nu > 0$. Let the ith component of $\boldsymbol{F}_{\boldsymbol{\beta},\nu,\boldsymbol{X},\rho}$ be the distribution function of $\mu_i = \text{logit}^{-1}(\boldsymbol{X}_i^\top \boldsymbol{\beta} + e_i)$, $i = 1, \ldots, n$, which depends on both ν and ρ. Then, model (8) is equivalent to model (2) with $\boldsymbol{F}_{\boldsymbol{\beta},\nu,\boldsymbol{X},\rho}$ and $\boldsymbol{Q}$ given in (3). Of course, this equivalence is just of mathematical interest; specifying model (8) is much more intuitive.

Changing $\boldsymbol{F}$ in model (2) to any distribution function defined over the (0, 1) support leads to new models. Covariate effects can be accommodated into the marginal parameters. Spatial dependence is modeled through the Gaussian copula with dispersion matrix $\boldsymbol{Q}^{-1}$. The beta distribution is a natural choice for the margins. Let $\text{Beta}(\nu p, \nu(1-p))$ represent a beta distribution with mean parameter p and dispersion parameter ν. Covariates can be incorporated into the mean parameter p using any transformation function from the real line to (0, 1) (e.g., Ferrari and Cribari-Neto, 2004). We propose a beta-logit model that incorporates covariates into the mean parameter p using an inverse logit transformation and defines marginal distribution F_i as

$$\text{Beta}\left[\nu \frac{\exp(\boldsymbol{X}_i^\top \boldsymbol{\beta})}{\exp(\boldsymbol{X}_i^\top \boldsymbol{\beta}) + 1}, \quad \nu\left\{1 - \frac{\exp(\boldsymbol{X}_i^\top \boldsymbol{\beta})}{\exp(\boldsymbol{X}_i^\top \boldsymbol{\beta}) + 1}\right\}\right], \quad \nu > 0,\ i = 1, \ldots, n. \tag{9}$$

Model (9) is, again, subtly different from the traditional spatial logit model (8) in that the parameters in the dependence structure $\boldsymbol{\Omega}$ do not enter the marginal distributions.

4. Bayesian inference with MCMC

The proposed models fit naturally into the Bayesian framework. With carefully chosen priors for the parameters, Markov chain Monte Carlo (MCMC) algorithms can be developed to draw samples from the posterior distribution of the parameters of interests (e.g., Gelman et al., 2003). The joint density of $\text{TGMRF}_n(\boldsymbol{F}, \boldsymbol{Q})$ is easily derived from the marginal distributions and the Gaussian copula,

$$h(\boldsymbol{\mu}|\boldsymbol{F}, \boldsymbol{Q}) = (2\pi)^{-\frac{n}{2}} |\boldsymbol{Q}|^{\frac{1}{2}} \exp\left(-\frac{1}{2}\boldsymbol{\varepsilon}^\top \boldsymbol{Q} \boldsymbol{\varepsilon}\right) \prod_{i=1}^{n} \frac{f_i(\mu_i; \boldsymbol{\beta}, \nu)}{\phi(\varepsilon_i)},$$

where f_i is the density of F_i, ϕ is the density of $N(0, 1)$, $\varepsilon_i = \Phi^{-1}\{\text{F}_i(\mu_i)\}$, $i = 1, \ldots, n$, and $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_n)^\top$. Let $\pi(\beta_j)$, $j = 1, \ldots, q$, be independent priors for $\boldsymbol{\beta}$. Let $\pi(\nu)$ and $\pi(\rho)$ be

independent prior for ν and ρ, respectively, independent of the prior for $\boldsymbol{\beta}$. The joint posterior density of $\boldsymbol{\theta}^\top = (\boldsymbol{\beta}^\top, \nu, \rho)$ is

$$\pi(\boldsymbol{\theta}|\boldsymbol{Y}, \boldsymbol{X}) \propto \left[\prod_{i=1}^{n} \pi(y_i|\mu_i)\right] h(\boldsymbol{\mu}|\boldsymbol{F}_{\beta,\nu,\boldsymbol{X}}, \boldsymbol{Q}_\rho) \left[\prod_{j=1}^{q} \pi(\beta_j)\right] \pi(\nu)\pi(\rho). \tag{10}$$

A general Gibbs sampling algorithm with Metropolis–Hasting update can be devised to draw from $\pi(\boldsymbol{\theta}|\boldsymbol{Y}, \boldsymbol{X})$. As the full conditionals from (10) are not known distributions (see for details, Appendix A), we use the adaptive rejection Metropolis sampling (ARMS) method (Gilks et al., 1995) in every update. ARMS is a general-purpose method for efficiently sampling from complicated, possibly non-logconcave or multi-modal univariate densities, as what are typically encountered in Gibbs sampling. It makes the implementation of MCMC straightforward for general application of TGMRF in hierarchical models, including the Poisson and Bernoulli cases for the snail abundance modeling. The computing burden is similar to that in a SGLMM. Both models can be viewed as hidden GMRF models: a GMRF is hidden behind the link functions in a SGLMM, whereas it is hidden behind the marginal quantile functions in a TGMRF. The quantile functions could be viewed as a class of new link functions (Prates et al., 2013). Therefore, the computation burden and the convergence speed of the two models are very similar, as long as the quantile functions are not much more expensive to evaluate than the link functions. If the quantile functions were available in BUGS, an implementation would be very easy, similar to that of SGLMM in model description. Our implementation was in C (using Gilk's C function for ARMS) and interfaced to R. Similar to MCMC algorithms for CAR models, auto-correlation in the MCMC sample can be high and the sample has to be thinned. A block Metropolis algorithm may be possible if we think of the model as a hidden GMRF model (Rue and Held, 2005, ch. 5), but a fuller study and comparison are worth a separate project.

If prediction is of interest for non observed areas, Y_{nobs}, an extra step can be added in the MCMC scheme to sample the missing data from the likelihood given the parameters at iteration m of the MCMC, generating $Y_{nobs}^{(m)}$. This way, the posterior sample $(Y_{nobs}^{(1)}, \ldots, Y_{nobs}^{(M)})$ can be used to find point and interval estimates of the predicted data.

In our study, the prior distributions of regression coefficients β_i, $i = 1, \ldots, q$, were set to be independent $N(0, 1/\tau)$ with $\tau = 0.01$. The additional parameter ν for the marginal distributions turns out to be either scale or shape parameter in our applications. Its prior distribution was set to be $\Gamma(\kappa_1, \kappa_2)$, a gamma distribution with shape κ_1 and scale κ_2. The hyperparameters were set to be $\kappa_1 = 0.01$ and $\kappa_2 = 100$. These priors were chosen to be proper but vague to allow the posterior estimates to be mainly data driven. A uniform prior over (0, 1) was put on the dependence parameter ρ to ensure positive spatial dependence as intuitively expected. Its support is well within $(1/\lambda_{\min}, 1/\lambda_{\max})$, where $\lambda_{\min} < 0$ and $\lambda_{\max} > 0$ are the minimum and maximum eigenvalues of $\boldsymbol{W}$, respectively, guaranteeing its propriety (Banerjee et al., 2004).

To compare different models for the same data, we propose to use the conditional predictive ordinate (CPO) criterion (e.g., Gelfand et al., 1992; Dey et al., 1997). The summary statistic is the logarithm of the pseudo-marginal likelihood (LPML), which is the summation of the log density of leave-one-out marginal predictive posterior distribution. The performance of the CPO criterion in selecting the right models for count data and binary data are studied through simulations. The LPML is the same as the logarithmic score of Good (1952) except that a leaving-one-out cross-validation is built in. Using it as a predictive measure in model selection dates back to Geisser and Eddy (1979). It is a special case of the general scoring rule studied in a very general framework by Gneiting and Raftery (2007), and specifically for count data by Czado et al. (2009). The performance of the CPO criterion in selecting the right models will be studied through simulations. Notice that a higher value of the LPML indicates a better fit of the model.

The deviance information criterion (DIC) is an alternative Bayesian model selection criterion (Spiegelhalter et al., 2002). In our simulation studies, however, DIC had much higher variation than LPML and was outperformed in selecting the correct models. This might be explained by the fact that the DIC measures are highly dependent on the marginalization of the random effects, and become unstable when the distributions are nonnormal.

The Integrate Nested Laplace Approximation (INLA) approach (Rue et al., 2009) initially seems as an alternative to implement the TGMRF approach since it works for Latent Gaussian Models (LGM) where the latent field is additive and normally distributed. However, our modeling framework can be interpreted as a link function (Prates et al., 2013) with the dependence structure defined by the Gaussian copula, in such way that a multivariate distribution is defined for $\boldsymbol{\mu}$. For this reason, this modeling setup is not suitable for the LGM class since our representation does not have a latent field with additive distribution. Moreover, the parameter $\boldsymbol{\theta}$ belongs to the marginal distribution $\boldsymbol{F}_{\boldsymbol{\beta},\nu,\boldsymbol{X},\rho}$, which can be viewed as a part of the link function. Therefore, from an INLA perspective all parameters would be treated as hyperparameters and it is known that INLA works well for large latent fields but a small number of hyperparameters.

5. Simulation study

5.1. Poisson regression

To assess the fitting capacity of the TGMRF models, the properties of the Bayesian inferences, and the effectiveness of LPML as a model comparison criterion in this context, we conducted a simulation study using the lattice and neighbor structure in Fig. 1(a). Each of the three models was used as data generating models. In addition to the intercept, one covariate was generated from $N(0, 1)$, and the true covariate coefficient vector was $\boldsymbol{\beta} = (1.0, 0.7)$. The precision matrix of the TGMRF took the form of (3) for the CAR model, with $\rho = 0.8$. The parameter ν, which is related to the variance in all models, was set at $\nu = 2$, although it has completely different meanings. With $\nu = 2$, the gamma scale model and the gamma shape model appeared to be more similar to each other than to the log-normal model. To make a more interesting comparison, a second log-normal model was also used to generate data, where $\nu = 6.5$ was chosen because it provides good approximation to the gamma scale model with $\nu = 2$. In summary, we had a total of four data generating models: two LN models LN1 and LN2, one GSC model, and one GSH model.

For each data generating model, we generated 100 datasets, and fit each dataset with all three proposed TGMRF models. In each fitting process, a vague prior, $\Gamma(0.01, 100)$, was set for the dispersion parameter ν, and an uninformative $U(0, 1)$ prior was set for the spatial dependence parameter ρ. Independent $N(0, 100)$ priors were set on the regression coefficients β_j, $j = 0, 1$. Table 1 summarizes the mean and standard deviations of the Bayesian estimate of the parameters and LPML from the 100 replicates.

When the model was correctly specified, the true values of the regression coefficients were recovered very well. The estimates seem to be upward biased for the dispersion parameter ν but downward biased for the dependence parameter ρ, suggesting that spatial dependence and spatial heterogeneity are hard to identify. When the model was misspecified, the regression coefficient estimates were still recovered reasonably well, especially in the GSC model and the GSH model, probably because the mean of $\boldsymbol{\mu}$ was still correctly specified, regardless of the misspecified model. In all cases, the average of the LPML statistic was higher for correctly specified models than for the misspecified models, with similar variation under different models. For the true models the coverage rate was close to 95% as expected, even the misspecified models in general had a high coverage rate for the regression coefficients and the dependence parameters. The scale parameter has very different interpretation and scales between the models, so in the true model we can verify that the coverage rate was close to 95%, for the misspecified models the coverage were omitted since they are not comparable.

To gain a clearer picture on model comparison using LPML, we summarize the frequencies of the models selected with the highest LPML from all 100 replicates under each of the four models (Table 2). The criterion seems to be very effective when the true model was the GSH model, correctly selecting the true model 89 times. When the true model was LN1 or GSC, the correct model was selected 59 times in either case with our sample size, while the alternative GSC model or LN model was selected 29 and 34 times, respectively; the GSH model was selected only 12 and 7 times, respectively. This indicates that the LN model and the GSC model provide good approximation to each other, similar to their well known similarity in univariate modeling without covariates and spatial concerns; a larger

Table 1
Summaries of posterior means, standard deviations (SD), 95% HPD coverage rates (Cov) and LPML from 100 replicates in the simulation of spatial Poisson regression.

True model	Param	True value	Specified model					
			LN		GSC		GSH	
			Mean (SD)	Cov	Mean (SD)	Cov	Mean (SD)	Cov
LN1	β_0	1.00	0.99 (0.10)	97%	1.08 (0.09)	90%	1.11 (0.14)	90%
	β_1	0.70	0.70 (0.06)	95%	0.70 (0.05)	95%	0.67 (0.06)	92%
	ρ	0.80	0.53 (0.26)	96%	0.55 (0.26)	98%	0.55 (0.26)	96%
	ν	2.00	2.32 (0.73)	98%	6.35 (1.53)	–	0.84 (0.46)	–
	LPML		−331.90 (10.76)		−332.55 (10.87)		−335.57 (11.11)	
LN2	β_0	1.00	0.98 (0.16)	96%	1.33 (0.18)	60%	1.46 (0.23)	50%
	β_1	0.70	0.70 (0.08)	96%	0.70 (0.07)	92%	0.58 (0.08)	59%
	ρ	0.80	0.60 (0.22)	94%	0.61 (0.21)	88%	0.64 (0.22)	93%
	ν	6.50	6.97 (1.42)	95%	2.00 (0.41)	–	3.50 (1.16)	–
	LPML		−362.90 (15.06)		−365.69 (14.78)		−371.35 (15.77)	
GSC	β_0	1.00	0.76 (0.18)	63%	0.99 (0.13)	94%	1.09 (0.19)	94%
	β_1	0.70	0.70 (0.08)	96%	0.70 (0.07)	96%	0.61 (0.07)	71%
	ρ	0.80	0.59 (0.23)	95%	0.59 (0.23)	93%	0.60 (0.23)	94%
	ν	2.00	6.34 (1.30)	–	2.24 (0.53)	98%	2.18 (0.73)	–
	LPML		−336.34 (14.96)		−335.38 (14.47)		−340.95 (15.12)	
GSH	β_0	1.00	0.71 (0.20)	64%	1.00 (0.14)	90%	0.99 (0.18)	94%
	β_1	0.70	0.77 (0.08)	84%	0.71 (0.08)	94%	0.70 (0.07)	95%
	ρ	0.80	0.64 (0.21)	95%	0.62 (0.21)	94%	0.63 (0.21)	96%
	ν	2.00	7.09 (1.31)	–	1.95 (0.48)	–	2.17 (0.60)	99%
	LPML		−335.86 (17.27)		−335.96 (16.60)		−328.81(17.35)	

Table 2
Frequencies of selected model using the LPML statistics for the 100 simulated datasets.

True model	Frequency selected		
	LN	GSC	GSH
LN1	59	29	12
LN2	77	16	7
GSC	34	59	7
GSH	6	5	89

sample would be necessary to distinguish them effectively. With our sample size, when the true model was LN2, the LPML was able to differentiate the LN model better from the GSC model, correctly selecting the LN model 77 times. Therefore, the similarity between the GSC model and the LN model appear to be different under different scenarios. The GSH model seems to have specific characteristics that make it further from the LN model and the GSC model in the model space.

A closer look at the difference in LPML across models is through box plots. Fig. 2 presents the box plots of the difference in LPML between the correct model and two misspecified models for each true model. The magnitude of the differences provides guidance in practice on what models are similar to each other and on how big a difference should be to be considered important. In the spatial setting we considered, the LN model and the GSC model were very similar, as seen from the boxes centered near zero. The majority of each box plot is well above −5, suggesting that if the LPML of one model is observed to be higher than that of another model by 5, then it is very unlikely that the other model is the true model.

5.2. Bernoulli regression

A simulation study was conducted for the spatial Bernoulli regressions. Both the logit model and the beta-logit model with the CAR dependence structure were used to generate data. Except for the response variable, the simulation setup was the same as that in Section 5.1 with model parameters

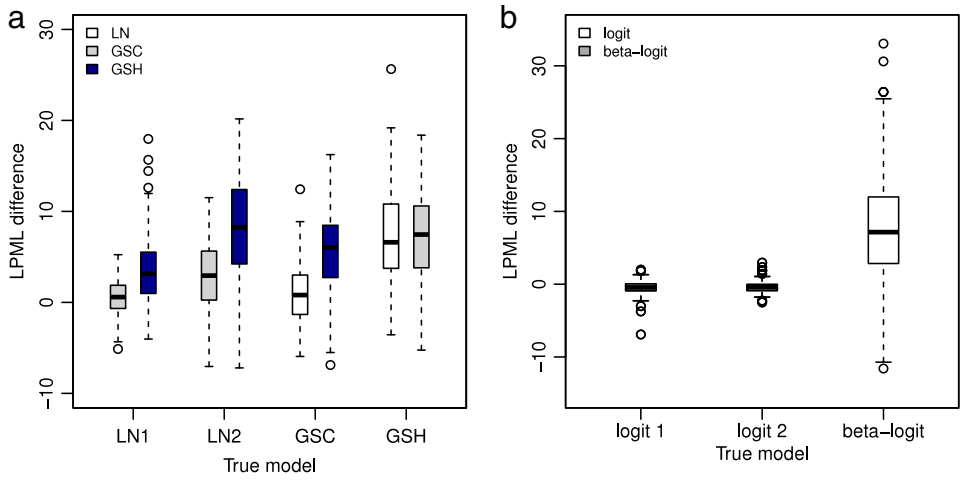

Fig. 2. LPML difference between the correct model and misspecified models. (a) Poisson simulation. (b) Bernoulli simulation.

Table 3
Summaries of posterior means, standard deviations (SD), 95% HPD coverage rates (Cov) and LPML from 100 replicates in the simulation of spatial Bernoulli regression.

True model	Param	True value	Specified model			
			logit		beta-logit	
			Mean (SD)	Cov	Mean (SD)	Cov
logit 1	β_0	1.00	1.02 (0.22)	98%	0.95 (0.23)	98%
	β_1	0.70	0.72 (0.23)	95%	0.65 (0.20)	95%
	ρ	0.80	0.50 (0.29)	100%	0.47 (0.27)	100%
	ν	2.00	2.33 (6.03)	86%	4.24 (2.49)	–
	LPML		−92.30 (5.59)		−91.77 (5.77)	
logit 2	β_0	1.00	1.03 (0.22)	97%	0.97 (0.23)	98%
	β_1	0.70	0.73 (0.22)	96%	0.66 (0.20)	94%
	ρ	0.80	0.50 (0.29)	100%	0.46 (0.27)	100%
	ν	1.00	1.56 (3.57)	95%	3.79 (2.52)	–
	LPML		−91.67 (5.53)		−91.31 (5.58)	
beta-logit	β_0	1.00	1.08 (0.22)	84%	1.01 (0.27)	92%
	β_1	0.70	0.78 (0.23)	93%	0.68 (0.20)	97%
	ρ	0.80	0.52 (0.29)	100%	0.56 (0.27)	100%
	ν	2.00	1.10 (1.09)	–	3.99 (2.37)	100%
	LPML		−96.16 (6.27)		−88.16 (8.21)	

$\boldsymbol{\beta} = (1.0, 0.7)$, $\rho = 0.8$ and $\nu = 2$. Again, since ν has different interpretation in the two models, a second logit model with $\nu = 1$ was also used to generate data in attempt to approximate the beta-logit model with $\nu = 2$. For each of three true models, we generated 100 datasets, and fit each dataset with each of two TGMRF models. The priors were chosen in the same manner as Section 5.1. Table 3 summarizes the posterior mean and standard deviations estimates from 100 replicates.

Similar to the results from Section 5.1, when the model was specified correctly, the true values of regression coefficients are recovered very well; the dispersion parameter estimate tended to be bigger than true value; and the dependence parameter estimate appeared to be downward biased. The coverage rate also had similar behavior to the ones observed in Section 5.1. However, the models tend to provide higher coverage rate for the dispersion and dependence parameters, this is due to the fact that the model has a large standard deviation in these parameters estimation. When the true model was the beta-logit model, the average LPML value of the beta-logit model was 8 higher than that of the logit model. When the true model was the logit 1 or logit 2, however, the average LPML value of the

Table 4
Frequencies of model selection using the LPML statistics for the 100 simulated datasets.

True model	Frequency selected	
	logit	beta-logit
logit 1	46	54
logit 2	49	51
beta-logit	16	84

beta-logit model was very close to (actually slightly higher than) that of the logit model in both cases. This implies that the beta-logit model is quite accommodating and can provide close approximation to the logit model; with the sample size in our simulation, they are hard to distinguish.

Table 4 summarizes the frequencies of the models selected with the highest LPML from all 100 datasets generated under each scenario. When the true model was the beta-logit model, the LPML criterion worked effectively, correctly selecting the true model 84 times. When the true model was logit 1 or logit 2, however, the logit model and the beta-logit model were selected with almost equal frequency, indicating that the beta-logit model provides very good approximation of the logit model with our sample size.

Box plots of the difference in LPML between the correct model and the misspecified model are shown in Fig. 2(b). The boxes are surprisingly tight around zero when the true model is the logit model, indicating that the beta-logit model approximates the logit model very closely in terms of LPML. When the true model was the beta-logit model, however, the LPML value of the logit model was very unlikely to be higher than that of the correctly specified model. The majority of all box plots were well above −5. A difference of 4.2 between the two models as observed in the analysis of presence/absence of *G. nigrolineata* seems to be quite strong evidence in favor of the beta-logit model.

6. Analysis of experimental data

6.1. Abundance of Nenia tridens

We fitted hierarchical Poisson regressions to the count data of *N. tridens* with four models: the random effect log linear model (4), the LN model, and two TGMRF models, GSC and GSH. An intercept model with the available covariates were fit for each model where the precision matrix $\boldsymbol{Q}$ of the Gaussian copula was specified with (3) from the CAR model. Any two sites within 60 m were considered to be neighbors. This neighborhood structure results in different numbers of neighbors for major sites and for supplementary sites. As shown in Fig. 1(a), an internal major site is connected to 20 neighbors and an internal supplementary site is connected to 16 neighbors. The priors were chosen as presented in Section 4. For each model, two chains with 60,000 iterations each were generated. We discarded the first 30,000 iterations as burn-in and thinned the rest by 10, resulting in 6,000 posterior samples. Convergence was verified using Geweke (1992) and Gelman and Rubin (1992) criteria.

The GSH model had the largest LPML (−482.12), followed by the GSC model (−483.90) and the LN model (−491.15). These results suggest that the GSH model and the GSC model performed similarly, with the former being slightly preferred. The GSH model provides considerably better fit than did the traditional LN model with a 9.1 difference in LPML. This difference is quite strong evidence in favor of the GSH model over the LN model (Kass and Raftery, 1995). As to be seen in our simulation study, when the true model was the GSH model, 38 of 100 replicates had LPML differences of greater than 9.1; when the true model was either one of the two LN considered models, however, this rate became 0 out of 100.

The posterior point estimates and 95% highest posterior density (HPD) credible intervals of the parameters from the GSH model and the traditional LN model are summarized in Table 5. The two models lead qualitatively to the same conclusions. Neither elevation nor slope was found to have a significant effect on the abundance of *N. tridens*. Of the habitat characteristics, only canopy openness was significant and negatively so. More canopy openness implies fewer trees and dryer soil, which

Table 5
Posterior point estimates and 95% HPD credible intervals of the parameters in the Poisson regression for the abundance of *N. tridens* with the GSH model and the traditional LN model. The regression coefficients are in the order of intercept, elevation, slope, quantity of litter, canopy openness, plant apparency, and apparency of sierra palm.

Parameters	Specified model			
	GSH		LN	
	Estimate	95% HPD	Estimate	95% HPD
Regression coefficients				
β_0	1.889	(0.561, 2.656)	2.112	(1.520, 2.839)
β_1	−0.056	(−0.322, 0.189)	−0.028	(−0.301, 0.252)
β_2	−0.051	(−0.184, 0.091)	−0.044	(−0.156, 0.077)
β_3	−0.108	(−0.250, 0.027)	−0.101	(−0.213, 0.027)
β_4	−0.149	(−0.304, −0.002)	−0.127	(−0.249, −0.003)
β_5	0.028	(−0.109, 0.170)	0.022	(−0.106, 0.145)
β_6	0.035	(−0.094, 0.165)	0.023	(−0.104, 0.142)
Scale and spatial dependence parameters				
ν	4.823	(2.857, 7.075)	5.001	(3.461, 6.773)
ρ	0.951	(0.840, 0.997)	0.954	(0.858, 0.999)

do not constitute the preferred habitat conditions of *N. tridens*. The marginal scale parameter ν is estimated to be 4.823. The spatial dependence parameter ρ is estimated as 0.951, with a HPD interval that does not include zero, indicating that higher spatial dependence is needed in the model.

To study the sensitivity of neighborhood choice, two other scenarios were fitted. First, where the major sites have only supplementary sites as neighbors, such that two sites within 57 m were considered to be neighbors; and second where two major sites were considered neighbors, thus two sites within 120 m were considered to be neighbors. The posterior estimates of the parameters were very similar and thus are not shown. The estimate of the spatial dependence parameter decreases in the scenario with neighborhood structure encompassing 120 m as ecologically expected, since these gastropods should not often move such great distances during over the period of days to weeks.

6.2. Presence of *Gaeotis nigrolineata*

We fitted Bernoulli regressions with an intercept for the presence/absence data of *G. nigrolineata* with two models: the random effect logit model (8) and the TGMRF model (9), denoted by the logit model and the beta-logit model, respectively. Prior distributions for model parameters were selected in the same way as those described in Section 4. As in Section 6.1, two chains of size 60,000 were generated, with the first 30,000 discarded and the rest thinned by 10. The convergence of the resulting 6,000 posterior samples was checked again with Geweke's and Gelman and Rubin criteria.

The LPML values were −99.31 and −103.51 for the beta-logit model and the logit model, respectively. Therefore, using a CAR dependence structure, the beta-logit model fits better than does the traditional logit model. The difference of 4.2 is quite strong evidence that the beta-logit model is superior to the logit model in this scenario. To be seen in our simulation study, when the true model was beta-logit, 68 out of 100 replicates had LPML differences between the fitted beta-logit model and the fitted logit model that were greater than 4.2, but the rate was only 1 or 0 out 100 when the true model was a logit model.

The posterior point estimates and 95% HPD credible intervals for parameters in both models are summarized in Table 6. The conclusions of the two models are virtually the same. As in the case for *N. tridens*, neither elevation nor slope had a significant effect on the incidence of *G. nigrolineata*. Of the habitat characteristics, only plant apparency had a significant negative effect on the incidence of *G. nigrolineata*. That is, the greater the volume of vegetation in the understory of the forest, the lower the abundance of *G. nigrolineata*. The apparency of sierra palm, which measures the preferred substrate for the *G. nigrolineata*, was almost significant and positively so, with the 95% HPD credible interval barely including zero. The negative effect of plant apparency was surprising but the paradox may be resolved if high plant apparency in the understory indicates the presence of an opening in the canopy, and attendant temperatures (high) with humidities (low) outside of the fundamental niche

Table 6
Posterior point estimates and 95% HPD credible intervals of the parameters in the Bernoulli regressions for the presence/absence of *G. nigrolineata* with the beta-logit model and the traditional logit model using the CAR dependence structure. The regression coefficients are in the order of intercept, elevation, slope, canopy openness, plant apparency, and apparency of sierra palm.

Parameters	Specified model			
	beta-logit		logit	
	Estimate	95% HPD	Estimate	95% HPD
Regression coefficients				
β_0	0.269	(−0.312, 0.816)	0.295	(−0.896, 1.571)
β_1	0.326	(−0.163, 0.792)	0.514	(−0.437, 1.467)
β_2	0.067	(−0.268, 0.418)	0.096	(−0.462, 0.686)
β_3	−0.031	(−0.364, 0.327)	−0.051	(−0.600, 0.558)
β_4	−0.513	(−0.936, −0.165)	−0.775	(−1.538, −0.111)
β_5	0.287	(−0.077, 0.649)	0.447	(−0.134, 1.220)
Scale and spatial dependence parameters				
ν	1.035	(0.740, 1.378)	52.960	(0.001, 186.300)
ρ	0.666	(0.179, 0.988)	0.707	(0.135, 0.999)

of *G. nigrolineata*, precluding its presence even though its preferred substrate may be common. The spatial dependence parameter ρ is estimated as 0.666 and 0.707 in the two models, respectively, with 95% HPD credible intervals excluding zero, indicating spatial dependence within neighborhood areas.

Although the beta-logit model agrees with the logit model in the directions of covariate effects, it gives smaller widths in the HPD credible interval than does the logit model. This indicates better precision in the estimation of coefficients. Nonetheless, the ν parameter does not have the same interpretation in the two models, and, hence, they are not directly comparable. For the logit model, the ν parameter controls the overall variation level of spatial random effects, whereas for the beta-logit model, the ν parameter controls the dispersion of the marginal distribution.

7. Discussion

7.1. Statistical interpretations

Our hierarchical spatial model with TGMRF provides a new avenue for modeling lattice data, especially those that are discrete, under the generalized linear model setting. Unlike Gaussian-copula spatial models where the Gaussian copula is applied to the data, our model applies the Gaussian copula to parameters of marginal distributions in the exponential family. This avoids the complexities and identification issues in handling discrete data, as cautioned in Genest and Neslehova (2007). The random field for the marginal parameters retains the robustness to misspecification of a general hierarchical model. Our new formulation covers the traditional SGLMM with GRF random effects as a special case, as we demonstrated in the context of the Poisson regression and Bernoulli regression. An important advantage, however, is that in the new formulation, the dependence parameters do not affect the marginal models. This is in contrast to the traditional SGLMM, where the spatial random effects are confounded with the fixed effects, making significant predictors less significant (Reich et al., 2006). This has important practical implications in making inferences about the regression coefficients in the marginal models through narrower credible intervals of the regression coefficients and, hence, more powerful detection of covariate effects. The scope of the modeling framework could be extended to multivariate spatial data, which may or may not have the same support at each site. It would be equivalent to applying the general Gaussian graphical model of Dobra et al. (2011) to parameters instead of data in a hierarchical model. Such a framework would then enable jointly modeling of the count data and binary data in our snail abundance application.

There is a considerable ongoing discussion in the ecological literature about how to identify the link between a trait of interest to environmental variables from the spatial variation of the trait itself (e.g., Peres-Neto et al., 2012). A simple GLM without spatial effects led to misleading results. The slope, which is apparently a spatially dependent covariate, would have been found to have a significant positive effect, but it effectively compensated for the spatial variation that is missing in the model.

From a statistical perspective, we need valid inferences about the regression coefficients in the presence of spatial variation. Both the traditional SGLMM and our hierarchical TGMRF model are capable of doing so, making inference on marginal regression coefficients more reliable after accounting for spatial variation. Application of both models helps to understand what environmental variables influence the variation in the gastropod abundances. Moreover, since our model overcomes the confounding between the spatial random effects and the fixed effects in a traditional SGLMM (Reich et al., 2006), it led to more reliable cherry picking the "pure" environmental variables that have an effect on the trait of interest in the presence of spatial variation. In particular, in the analysis of the count data of *N. tridens*, our GSH model was able to select canopy openness as a significant variable, while the traditional Poisson regression with GMRF random effect was not. In the analysis of the presence/absence data of *G. nigrolineata*, although the conclusions are the same as those from the traditional logit model with GMRF random effect, our beta-logit model gave much tighter credible intervals. The R-C code for the TGMRF regression is provided as supplementary material (see Appendix B).

7.2. Ecological interpretations

The empirical data on the spatial distribution of *N. Tridens* and *G. nigrolineata* was obtained for a time period (1995) that was only six years after the impact of Hurricane Hugo (1989)—a category 4 storm with catastrophic wind speeds (>144 km/h) that caused considerable damage to the Luquillo Mountains, including mortality or defoliation to 7% and 56% of trees, respectively, in tabonuco forest near the LFDP (Scatena et al., 2012). Within six months of the hurricanes impact, the mean densities of both gastropods significantly declined to essentially zero (i.e., locally extinct) in some areas of tabonuco forest (Willig and Camilo, 1991). By 1995, mean abundance of *N. tridens* had increased appreciably while those of *G. Nigrolineata* initially increased but then decreased (Willig et al., 1998). Moreover, spatial variation in abundance of each species was changing significantly over time in response to the hurricane (Bloch and Willig, 2006). These temporal patterns suggest that the spatial organization of gastropod abundance was dynamic in complex ways after the hurricane in response to the interplay between the site-specific effects of canopy opening and debris deposition, immediate effects of hurricane disturbance that initiate a sequence of responses (secondary succession) by the biota (Willig and McGinley, 1999; Willig et al., 2012). Thus, results from our modeling (i.e., few environmental characteristics were significant) are consistent with the expectation that many environmental characteristics of sites on the LFDP may not appreciably predict gastropod abundance or presence while this ecological system is undergoing rapid reorganization and not likely in an equilibrial state (i.e., habitat characteristics are dynamic over time, as are abundances of gastropods, potentially reflecting a lag-time in response to changing conditions (Willig et al., 2011)). Similarly, cross-scale interactions between environmental conditions at each site (fine scale) and the surrounding characteristics of the landscape (broad scale), may produce complex patterns in abundance or incidence that are not strongly related to contemporary environmental conditions at a site (Willig et al., 2007). The spatial relationship (negative) between canopy openness and abundance of *N. tridens* does correspond to a priori ecological expectations that this snail would avoid sites with high canopy openness because of the desiccation effects and the poor quality of forage in such environs. Similarly, a priori expectations and statistical results of analyses correspond with regard to the presence of *G. Nigrolineata*, a gastropod with a vestigial and internal shell, that prefers areas in which its primary substrate or food source are pervasive (i.e., those with high apparency of the palm). The association may be weak because the response variable is binary: presence does not distinguish between situations in which a single individual or 100 individuals are present. The likelihood of absence being associated with high plant apparency arises because appreciable vegetation in the understory suggests that canopy closure is not yet complete, and abiotic characteristics (higher temperatures and lower moisture) may create hostile environments for this species.

Acknowledgments

This research was partially supported by a Multidisciplinary Environmental Research Award for Graduate Students to M.O. Prates from the Center for Environmental Sciences and Engineering at the

University of Connecticut. M.O. Prates also acknowledges FAPEMIG and CNPq for partial financial support. In addition, this research was facilitated by grant numbers BSR-8811902, DEB-9411973, DEB-0080538, and DEB-0218039 from the National Science Foundation to the Institute of Tropical Ecosystem Studies, University of Puerto Rico, and the International Institute of Tropical Forestry as part of the Long-Term Ecological Research Program in the Luquillo Experimental Forest. Additional support was provided by the USDA Forest Service, the University of Puerto Rico, the Department of Biological Sciences at Texas Tech University, and the Center for Environmental Sciences and Engineering at the University of Connecticut. The staff of El Verde Field Station provided valuable logistical support in Puerto Rico. Finally, we thank the mid-sized army of students and colleagues who have assisted with collection of field data over the years.

Appendix A. Full conditionals

The full conditionals for the parameters for MCMC in the context of TGMRF regression are:

$$\pi(\mu_i|\text{rest}) \propto \pi(y_i|\mu_i) \exp\left(-\frac{1}{2}\boldsymbol{\varepsilon}^\top \mathbf{Q}_\rho \boldsymbol{\varepsilon}\right) \frac{f_{\boldsymbol{\beta},\nu,\mathbf{X}}(\mu_i)}{\phi(\varepsilon_i)}, \quad \text{for } i = 1, \ldots, n,$$

where $\pi(y_i|\mu_i)$ is the appropriate likelihood for the response, $f_{\boldsymbol{\beta},\nu,\mathbf{X}}$ is the density corresponding to $\boldsymbol{F}_{\boldsymbol{\beta},\nu,\mathbf{X}}$ with parameters, ϕ is the density of $N(0,1)$ and $\boldsymbol{\varepsilon} = \left[\Phi^{-1}\left\{F_{\boldsymbol{\beta},\nu,\mathbf{X}}(\mu_1)\right\}, \ldots, \Phi^{-1}\left\{F_{\boldsymbol{\beta},\nu,\mathbf{X}}(\mu_n)\right\}\right]^\top$.

$$\pi(\beta_j|\text{rest}) \propto \prod_{i=1}^{n}\left[\pi(y_i|\mu_i)\frac{f_{\boldsymbol{\beta},\nu,\mathbf{X}}(\mu_i)}{\phi(\varepsilon_i)}\right] \times \exp\left(-\frac{\tau}{2}\beta_j^2 - \frac{1}{2}\boldsymbol{\varepsilon}^\top \mathbf{Q}_\rho \boldsymbol{\varepsilon}\right), \quad \text{for } j = 1, \ldots, q,$$

$$\pi(\rho|\text{rest}) \propto \prod_{i=1}^{n}\pi(y_i|\mu_i) \times |\mathbf{Q}_\rho|^{\frac{1}{2}} \exp\left(-\frac{1}{2}\boldsymbol{\varepsilon}^\top \mathbf{Q}_\rho \boldsymbol{\varepsilon}\right) I(a < \rho < b),$$

and

$$\pi(\nu|\text{rest}) \propto \prod_{i=1}^{n}\left[\pi(y_i|\mu_i)\frac{f_{\boldsymbol{\beta},\nu,\mathbf{X}}(\mu_i)}{\phi(\varepsilon_i)}\right] \times \exp\left((1-\kappa_1)\log(\nu) - \kappa_2\nu - \frac{1}{2}\boldsymbol{\varepsilon}^\top \mathbf{Q}_\rho \boldsymbol{\varepsilon}\right) I(\nu > 0),$$

where τ, κ_1 and κ_2 are the hyperparameters of the Normal and Gamma prior for β_j and ν respectively. The prior for ρ is set as an uniform between (a, b).

Appendix B. Supplementary material

Supplementary material related to this article can be found online at http://dx.doi.org/10.1016/j.spasta.2015.07.004.

References

Banerjee, S., Carlin, P.B., Gelfand, E.A., 2004. Hierarchical Modeling and Analysis for Spatial Data. Chapman & Hall, New York.
Bárdossy, A., 2006. Copula-based geostatistical models for groundwater quality parameters. Water Resour. Res. 42 (11), W11416.
Berrocal, V.J., Raftery, A.E., Neiting, T.G., 2008. Probabilistic quantitative precipitation field forecasting using a two-stage spatial model. Ann. Appl. Stat. 2, 1170–1193.
Besag, J., 1974. Spatial interaction and the statistical analysis of lattice data systems (with discussion). J. R. Stat. Soc. Ser. B 36, 192–225.
Besag, J., York, J., Mollie, A., 1991. Bayesian image restoration with two application in spatial statistics (with discussion). Ann. Inst. Statist. Math. 43, 1–59.
Bloch, C.P., Willig, M.R., 2006. Context-dependence of long-term responses of terrestrial gastropod populations to large-scale disturbance. J. Tropical Ecol. 22, 111–122.
Breslow, N.E., Clayton, D.G., 1993. Approximate inference in generalized linear mixed models. J. Amer. Statist. Assoc. 88, 9–25.

Brokaw, N., Crowl, T.A., Lugo, A.E., McDowell, W.H., Scatena, F.N., Waide, R.B., Willig, M.R., 2012. A Caribbean Forest Tapestry: The Multidimensional Nature of Disturbance and Response. Oxford University Press, New York.
Brown, S., Lugo, A.E., Silander, S., Liegel, L., 1983. Research History and Opportunities in the Luquillo Experimental Forest, General Technical Report SO-44. U.S. Dept of Agriculture, Forest Service, Southern Forest Experiment Station, New Orleans, LA.
Chilès, J.P., Delfiner, P., 1999. Geostatistics: Modeling Spatial Uncertainty. Wiley, New York.
Christensen, O.F., Waagepetersen, R., 2002. Bayesian prediction of spatial count data using generalized linear mixed models. Biometrics 58, 280–286.
Gilks, W.R., Best, N.G., Tan, K.K.C., 1995. Adaptive rejection metropolis sampling within Gibbs sampling (Corr: 97V46 P541–542 with R.M. Neal). Appl. Stat. 44, 455–472.
Czado, C., Gneiting, T., Held, L., 2009. Predictive model assessment for count data. Biometrics 65, 1254–1261.
Dey, D.K., Chen, M.H., Chang, H., 1997. Bayesian approach for nonlinear random effects models. Biometrics 53, 1239–1252.
Diggle, P.J., Tawn, J.A., Moyeed, R.A., 1998. Model-based geostatistics. Appl. Stat. 47, 299–350.
Dobra, A., Lenkoski, A., 2011. Copula Gaussian graphical models and their application to modeling functional disability data. Ann. Appl. Stat. 5, 969–993.
Dobra, A., Lenkoski, A., Rodriguez, A., 2011. Bayesian Inference for General Gaussian Graphical Models with Application to Multivariate Lattice Data. J. Amer. Statist. Assoc. 106 (496), 418–1433.
Ewel, J.J., Whitmore, J.L., 1973. The Ecological Life Zones of Puerto Rico and the United States Virgin Islands, Forest Service Research Papers ITF-18. Institute of Tropical Forestry.
Ferrari, P.S., Cribari-Neto, F., 2004. Beta regression for modelling rates and proportions. J. Appl. Stat. 7, 799–815.
Geisser, S., Eddy, W.F., 1979. A predictive approach to model selection (Corr: V75 P765). J. Amer. Statist. Assoc. 74, 153–160.
Gelfand, A.E., Dey, D.K., Chang, H., 1992. Model determination using predictive distributions, with implementation via sampling-based methods (Disc: P160–167). In: Bernardo, J.M., Berger, J.O., Dawid, A.P., Smith, A.F.M. (Eds.), Bayesian Statistics 4. Proceedings of the Fourth Valencia International Meeting. Clarendon Press [Oxford University Press], pp. 147–159.
Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B., 2003. Bayesian Data Analysis, second ed. Chapman and Hall/CRC, ISBN: 9781584883883.
Gelman, A., Rubin, D.B., 1992. Inference from iterative simulation using multiple sequences. Statist. Sci. 7, 457–511.
Genest, C., Neslehova, J., 2007. A primer on copulas for count data. Astin Bull. 37 (2), 475.
Geweke, J., 1992. Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments (Disc: P189–193). In: Bernardo, J.M., Berger, J.O., Dawid, A.P., Smith, A.F.M. (Eds.), Bayesian Statistics 4. Proceedings of the Fourth Valencia International Meeting. Clarendon Press [Oxford University Press], pp. 169–188.
Gneiting, T., Raftery, A.E., 2007. Strictly proper scoring rules, prediction, and estimation. J. Amer. Statist. Assoc. 102, 359–378.
Good, I.J., 1952. Rational decisions. J. R. Stat. Soc. Ser. B 14, 107–114.
Guillot, G., Lebel, T., 1999. Approximation of Sahelian rainfall fields with meta-Gaussian random functions. Stochastic Environ. Res. Risk Assess. 13 (1–2), 113–130.
Kass, E.R., Raftery, E.A., 1995. Bayes factor. J. Amer. Statist. Assoc. 90, 773–795.
Kazianka, H., 2013. spatialCopula: A matlab toolbox for copula-based spatial analysis. Stochastic Environ. Res. Risk Assess. 27 (1), 121–135.
Kazianka, H., Pilz, J., 2010. Copula-based geostatistical modeling of continuous and discrete data including covariates. Stochastic Environ. Res. Risk Assess. 24, 661–673.
Madsen, L., 2009. Maximum likelihood estimation of regression parameters with spatially dependent discrete data. J. Agric. Biol. Environ. Stat. 14, 375–391.
Masarotto, G., Varin, C., 2012. Gaussian copula marginal regression. Electron. J. Stat. 6, 1517–1549.
Mason, C.F., 1970. Snail populations, beech litter production and the role of snails in litter decomposition. Oecologia 5, 215–293.
Nelsen, R., 2006. An Introduction to Copulas, second ed. Springer-Verlag, New York.
Peres-Neto, P.R., Leibold, M.A., Dray, S., 2012. Assessing the effects of spatial contingency and environmental filtering on metacommunity phylogenetics. Ecology 93, S14–S30. http://dx.doi.org/10.1890/11-0494.1.
Pitt, M., Chan, D., Kohn, R., 2006. Efficient Bayesian inference for Gaussian copula regression models. Biometrika 93, 537–554.
Prates, M.O., Aseltine, R.H., Dey, D.K., Yan, J., 2013. Assessing intervention efficacy on high-risk drinkers using generalized linear mixed models with a new class of link functions. Biometrical Journal (ISSN: 1521-4036) 55 (6), 912–924. http://dx.doi.org/10.1002/bimj.201300015.
Reagan, D., Waide, R., 1996. The Food Web of a Tropical Rain Forest. University of Chicago Press, Chichago, Illinois.
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, 1197–1206.
Rue, H., Held, L., 2005. Gaussian Markov random fields: Theory and applications. In: Monographs on Statistics and Applied Probability, vol. 104. Chapman & Hall, London.
Rue, H., Martino, S., Chopin, N., 2009. Approximate Bayesian inference for latent Gaussian models using integrated nested Laplace approximations (with discussion). J. R. Stat. Soc. Ser. B 71, 319–392.
Rue, H., Steinsland, I., Erland, S., 2004. Approximating hidden Gaussian Markov random fields. J. R. Stat. Soc. Ser. B 66, 877–892. http://dx.doi.org/10.1111/j.1467-9868.2004.B5590.x.
Scatena, F.N., Blanco, J.F., Beard, K.H., Waide, R.B., Lugo, A.E., Brokaw, N., Silver, W.L., Haines, B., Zimmerman, J.K., 2012. A Carribean Forest Tapestry, chap. Disturbance regime. Oxford University Press, pp. 164–200.
Schaake, J., Demargne, J., Hartman, R., Mullusky, M., Welles, E., Wu, L., Herr, H., Fan, X., Seo, D.J., 2007. Precipitation and temperature ensemble forecasts from single-value forecasts. Hydrol. Earth Sys. Sci. Discussions 4, 655–717.
Schmid, V., Held, L., 2004. Bayesian extrapolation of space–time trends in cancer registry data. Biometrics 60, 1034–1042.
Song, P.-K., 2000. Multivariate dispersion models Generated from Gaussian Copula. Scand. J. Stat. 27, 305–320.
Spiegelhalter, D.J., Best, N.G., Carlin, B.P., Linde, A., 2002. Bayesian measures of model complexity and fit. J. R. Stat. Soc. Ser. B 64, 583–639.
Wikle, C.K., Berliner, L.M., Cressie, N.A.C., 1998. Hierarchical Bayesian space–time models. Environ. Ecol. Stat. 5, 117–154.
Willig, M.R., Bloch, C.P., Brokaw, N., Higgins, C., Thompson, J., Zimmermann, C.R., 2007. Cross-scale responses of biodiversity to hurricane and anthropogenic disturbance in a tropical forest. Ecosystems 10, 824–838.

Willig, M.R., Bloch, C.P., Covich, A.P., Hall, C.A.S., Lodge, D.J., Lugo, A.E., Silver, W.L., Waide, R.B., Walker, L.R., Zimmerman, J.K., 2012. A Caribbean Forest Tapestry: The Multidimensional Nature of Disturbance and Response, chap. Long-term Research in the Luquillo Mountains: Synthesis and Foundations for the Future. Oxford University Press, New York, pp. 361–441.
Willig, M.R., Camilo, G.R., 1991. The effect of hurricane Hugo on six invertebrate species in the Luquillo experimental forest of Puerto Rico. Biotropica 23, 455–461.
Willig, M.R., McGinley, M.A., 1999. Ecosystems of Disturbed Ground, chap. The Response of Animals to Disturbance and Their Roles in Patch Generation. Elsevier Science, Amsterdam, Netherlands, pp. 633–657.
Willig, M.R., Presley, S.J., Bloch, C.P., Castro-Arellano, I., Cisneros, L.M., Higgins, C.L., Klinbeil, B.T., 2011. Tropical metacommunities along elevational gradients: Effects of forest type and other environmental factors. Oikos 120 (10), 1497–1508. http://dx.doi.org/10.1111/j.1600-0706.2011.19218.x.
Willig, M.R., Secrest, M.F., Cox, S.B., Camilo, G.R., Cary, J.F., Alvarez, J., Gannon, M.R., 1998. Forest Biodiversity in North, Central South America and the Caribbean: research and Monitoring. UNESCO and the Parthenon Publishing Group, chap. Long-term Monitoring of Snails in the Luquillo Experimental Forest of Puerto Rico: Heterogeneity, Scale, Disturbance, and Recovery. The Parthenon Press, Cranforth, Lancashire, UK, pp. 293–322.
Wolpert, R.L., Ickstadt, K., 1998. Poisson/gamma random field models for spatial statistics. Biometrika 85, 251–267.