BAYESIAN STATISTICS 9,
J. M. Bernardo, M. J. Bayarri, J. O. Berger, A. P. Dawid,
D. Heckerman, A. F. M. Smith and M. West (Eds.)

Modelling multivariate counts varying continuously in space

Alexandra M. Schmidt & Marco A. Rodríguez
U. Federal do Rio de Janeiro, Brazil *U. du Québec à Trois-Rivières, Canada*
alex@im.ufrj.br marco.rodriguez@uqtr.ca

Summary

We discuss models for multivariate counts observed at fixed spatial locations of a region of interest. Our approach is based on a continuous mixture of independent Poisson distributions. The mixing component is able to capture correlation among components of the observed vector and across space through the use of a linear model of coregionalization. We introduce here the use of covariates to allow for possible non-stationarity of the covariance structure of the mixing component. We analyze joint spatial variation of counts of four fish species abundant in Lake Saint Pierre, Quebec, Canada. Models allowing the covariance structure of the spatial random effects to depend on a covariate, geodetic lake depth, showed improved fit relative to stationary models.

Keywords and Phrases: Animal Abundance; Anisotropy; Linear Model of Coregionalization; Non-Stationarity; Poisson Log-Normal Distribution; Random Effects.

1. INTRODUCTION

Count data from fixed spatial locations within a region of interest commonly arise in different areas of science, such as ecology, epidemiology, and economics. Moreover, multiple processes are frequently observed simultaneously at each location. Here, we discuss multivariate models for abundances of four different fish species observed at locations along the shorelines of a lake. Our models are based on a Poisson-multivariate lognormal mixture, as proposed by Aitchison & Ho (1989). We concentrate on the covariance structure of the mixing component, which can capture correlations among species and across locations. In particular, we make

---

A. M. Schmidt (www.dme.ufrj.br/∼alex) is Associate Professor of Statistics at the Federal University of Rio de Janeiro, Brazil. M. A. Rodríguez is Professor of Biology at Université du Québec à Trois-Rivières, Département de chimie-biologie, 3351 boul. des Forges, Trois-Rivières, Québec, G9A 5H7, Canada. We are grateful to CNPq and FAPERJ (A. M. Schmidt) and the Natural Sciences and Engineering Council of Canada (M. A. Rodríguez) for financial support, and to A. E. Gelfand for fruitful discussions.

use of the linear model of coregionalization (LMC) (Wackernagel, 2003; see also Schmidt & Gelfand, 2003; Gelfand et al., 2004) to describe the underlying complex covariance structure. Additionally, following Schmidt et al. (2010a), we investigate the influence of a covariate on the covariance structure of the spatial random effects.

1.1. *Ecological Motivation and Data*

Ecologists generally seek to interpret variations in species abundance in terms of environmental features and interactions with other species. Abundance in the form of counts is often highly variable in space and time and shows overdispersion relative to the Poisson. Species count data are simultaneously influenced by numerous factors, the individual effects of which can be difficult to disentangle in the absence of suitable models. Furthermore, spatial and temporal correlations in counts may complicate interpretation of environmental effects if not accounted for properly.

Data on fish abundances in Lake Saint Pierre, an ecosystem showing strong spatial heterogeneity and temporal variability, serve here to illustrate some of these difficulties, as well as possible solutions. Lake Saint Pierre (46^o 12' N; 72^o 50'W) is a fluvial lake of the Saint Lawrence River (Quebec, Canada). The lake is large (surface area: annual mean = 315 km^2; 469 km^2 during the spring floods) and shallow (mean water depth = 3.17 m). Both lake surface area and water level (range = 1.23 m) varied markedly over the study period (14 June - 22 August 2007). Lake Saint Pierre has distinct water masses along its northern, central, and southern portions. These water masses differ consistently in physical and chemical characteristics because lateral mixing is limited by a deep (>14 m) central navigation channel that has strong current and may act as a barrier to fish movement.

Counts of four fish species (yellow perch, brown bullhead, golden shiner, and pumpkinseed) and measurements of four environmental covariates (water depth, transparency, vegetation, and substrate composition) were obtained for 160 locations equally distributed between the North and South shores of the lake (Figure 1). Fish counts and environmental measurements were made along fishing trajectories approximately 650 m in length and parallel to the lake shoreline. Fish counts are expected to respond locally to habitat at the site of capture, which is characterized by the set of environmental covariates. In Lake Saint Pierre, suitable habitat for the target fish species is concentrated in the shallow littoral zones that border the lake shores. Deeper areas nearer to the central navigation channel are actively avoided by the target species. The sampling trajectories therefore provided extensive coverage of suitable littoral habitat along both shorelines.

Geodetic lake depth, measured as water depth minus the lake level relative to a fixed International Great Lakes low-water datum (IGLD55), was calculated for each location and sampling date. Locations at a given geodetic depth lie along a common isobath, or equal-depth contour along the lake bottom. Geodetic depth is linked to potential determinants of fish abundance. These include patterns of current flow, influence from terrestrial inputs, duration of flood period, as well as behavioural processes such as fish movements along depth contours. Any of these processes may induce spatial correlation in fish abundance, yet their effects may not be adequately captured by the local environmental covariates, particularly in lakes subject to large fluctuations in water level. Therefore, we explore here a correlation function that includes geodetic depth as a covariate allowing for non-stationarity of the covariance structure of the spatial random effects.

Along each shore, a set of 10 approximately evenly spaced sampling sectors was chosen to provide full longitudinal coverage of the shoreline. On each of 38 sampling

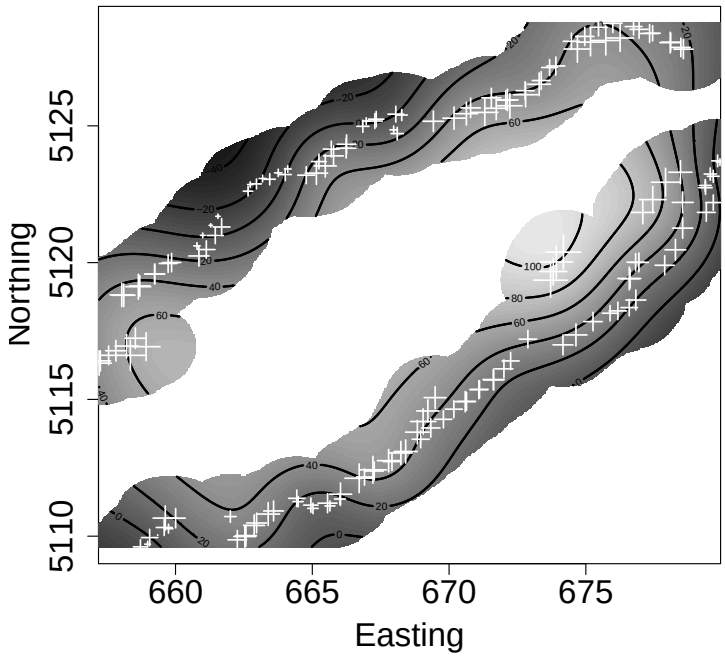

Figure 1: *Study locations along the North and South shores of Lake Saint Pierre (+). Each location represents the centroid of a fishing trajectory approximately 650 m in length. Symbol size is proportional to geodetic lake depth. Contour curves represent interpolated geodetic depth (relative scale)*

dates, measurements were made at each location from a cluster of spatially adjacent locations on one shore, within a sector selected at random among the predetermined set of sectors on that shore. Clusters comprised four locations on 36 sampling dates and eight locations on two sampling dates. Sampling dates were unevenly spaced in time over a period of 70 days, and the North and South shores were visited in alternation on consecutive sampling dates. This sampling design yielded measurements that were clustered both in space and in time, in contrast with the simultaneous sampling of all locations at all occasions characteristic of many spatio-temporal sampling schemes.

Our main ecological objectives were to: (1) assess the influence of local habitat (as characterized by the environmental covariates) on the abundance of fish species, (2) determine whether species abundances are correlated across space and among themselves, and (3) understand the spatial distribution of each species. The relationship of fish abundances with local habitat is typically determined by short-term

behavioural responses of individuals. In view of this objective, temporal variation in our study can be viewed as a nuisance that arises because collection of samples in the field was time-consuming. In this context, inclusion of temporal components in the models is primarily a means of adjusting for potential short-term fluctuations in fish abundances (e.g., seasonal declines in counts arising from short-term mortality) that are not in themselves of substantive interest.

The remainder of this paper is organized as follows. Section 2 provides a brief overview of models for multivariate counts and discusses extensions to account for spatial structure in the data. Section 3 introduces a model based on mixtures of the Poisson distribution and continuous variables, and discusses the resultant covariance structures for various mixing distributions. Section 4 presents an analysis of fish abundances based on this model. Section 5 concludes and points to open avenues for research.

2. A BRIEF OVERVIEW OF MULTIVARIATE COUNT DISTRIBUTIONS

A variety of proposals in the literature deal with multivariate distributions for counts. In this section we first sketch a derivation of the multivariate Poisson based on a sum of independent Poisson random variables. We then discuss the construction of alternative multivariate count distributions through the use of continuous mixtures of independent Poisson distributions.

2.1. *Derivation of the Multivariate Poisson as a Sum of Independent Poisson Random Variables*

The multivariate Poisson distribution can be derived as a sum of independent random variables, each following a Poisson distribution (Tsionas, 2001; Karlis & Meligkotsidou, 2005).

Here we follow the notation of Karlis & Meligkotsidou (2005) and Chandina (2007). Assume $\mathbf{W} = (W_1, \cdots, W_q)^T$, where W_i follows independent Poisson distributions, with mean λ_i, $i = 1, \cdots, q$, and $\mathbf{B}$ is a $K \times q$ matrix with zeros and ones, with $K \leq q$. The vector $\mathbf{Y} = (Y_1, \cdots, Y_K)^T$, defined as $\mathbf{Y} = \mathbf{BW}$, follows a K-variate Poisson distribution. The most general form assumes $\mathbf{B}$ to be a $K \times (2^K - 1)$ matrix, such that $\mathbf{B} = (\mathbf{B}_1, \mathbf{B}_2, \cdots, \mathbf{B}_K)$, with $\mathbf{B}_i$ being a submatrix of dimension $K \times \begin{pmatrix} K \\ i \end{pmatrix}$, where each column of $\mathbf{B}_i$ has exactly i ones and $(K - i)$ zeros, and no duplicate column exists (Chandina, 2007). It is simple to see that $E(\mathbf{Y}) = \mathbf{B}\boldsymbol{\lambda}$ and $V(\mathbf{Y}) = \mathbf{B}\boldsymbol{\Sigma}\mathbf{B}^T$, where $\boldsymbol{\Sigma} = \text{diag}\ (\lambda_1, \cdots, \lambda_q)$. For example, when $K = 3$, we have

$$\begin{aligned} Y_1 &= W_1 + W_{12} + W_{13} + W_{123} \\ Y_2 &= W_2 + W_{12} + W_{23} + W_{123} \\ Y_3 &= W_3 + W_{23} + W_{13} + W_{123}, \end{aligned}$$

with $W_i \sim Poi(\lambda_i)$, $W_{ij} \sim Poi(\lambda_{ij})$ $W_{ijl} \sim Poi(\lambda_{ijl})$, $i, j, l = \{1, 2, 3\}$, $i < j < l$. For simplicity, Karlis & Meligkotsidou (2005) do not consider the full structure of the model and drop the term W_{123} above, i.e., they consider only two-way covariances.

Given n observations $\mathbf{Y}_i$, $i = 1, \cdots, n$, the likelihood of this model assumes a complicated form and is time-consuming to evaluate; however, data augmentation can be used to simplify its evaluation (Karlis & Meligkotsidou, 2005). The main drawback of this model is that the mean and variance of Y_i are assumed identical, i.e., the model does not capture overdispersion. Furthermore, the model only accounts

for positive covariance structures. One way to incorporate overdispersion would be to consider random effects in the mean structure of, say, W_i, but we do not pursue this approach here because we anticipate that it would lead to heavy computation and difficulty of interpretation.

2.2. *Multivariate Count Distributions Based on Mixtures*

Mixtures of independent Poisson distributions provide a natural means of accounting for overdispersion in multivariate count data. Consider a random vector $\delta = (\delta_1, \cdots, \delta_K)$ that follows some probability density function with support on $\mathbb{R}^+$, say $g(\delta \mid \theta)$, where θ is the parameter vector of $g(.)$. We define $f(.)$ as the marginal multivariate probability function of the random vector of counts $\mathbf{Y}$. More specifically, let $\mathbf{Y} = (Y_1, \cdots, Y_K)$, with $Y_k \mid \delta_k \sim Poi(\delta_k)$, $k = 1, \cdots, K$, conditionally independent, then

$$f(y \mid \theta) = \int \prod_{k=1}^{K} p(y_k \mid \delta_k) g(\delta \mid \theta) d\delta,$$

where $p(.)$ is the probability function of the univariate Poisson distribution. In the following, we discuss the particular cases when $g(\delta \mid \theta)$ follows either multivariate gamma or multivariate lognormal distributions.

2.2.1. *Multivariate Poisson-Gamma Mixture*

A multivariate gamma distribution for a K-dimensional random vector $\boldsymbol{\delta}$ can be obtained by defining

$$\delta_k = \frac{b_0}{b_k} W_0 + W_k, \ k = 1, \cdots, K,$$

with $W_k \sim Ga(a_k, b_k)$ for $k = 0, 1, 2, \cdots, K$; W_k independent among themselves (Mathai & Moschopoulos, 1991). Similarly, we can define a $K \times (K+1)$ matrix $\mathbf{B}$ such that

$$\boldsymbol{\delta} = \mathbf{B}\,\mathbf{W} = \begin{pmatrix} \frac{b_0}{b_1} & 1 & 0 & \cdots & 0 \\ \frac{b_0}{b_2} & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{b_0}{b_K} & 0 & 0 & \cdots & 1 \end{pmatrix} \begin{pmatrix} W_0 \\ W_1 \\ \vdots \\ W_K \end{pmatrix}. \tag{1}$$

It is easy to show that $\mathrm{Cov}(\delta_k, \delta_l) = \frac{a_0}{b_k\, b_l}$, $k, l = 1, 2, \cdots, K$.

If $\mathbf{Y}$ is a K-dimensional vector, such that $Y_i | \delta_i \sim Poi(\delta_i)$, independently, and $\boldsymbol{\delta}$ is defined as in Equation (1), it can be shown that, marginally,

$$\begin{aligned} E(Y_i) &= E(E(Y_i \mid \delta_i)) = \frac{a_i}{b_i} + \frac{a_0}{b_i} = \mu_i \qquad i, j = 1, \cdots, K \\ V(Y_i) &= E(V(Y_i \mid \delta_i)) + V(E(Y_i \mid \delta_i)) = \mu_i + \frac{a_0}{b_i^2} + \frac{a_i}{b_i} \\ \mathrm{Cov}(Y_i, Y_j) &= Cov[E(Y_i \mid \delta_i), E(Y_j \mid \delta_j)] + E[Cov(Y_i, Y_j \mid \delta_i, \delta_j)] = \frac{a_0}{b_i\, b_j}. \end{aligned} \tag{2}$$

This model accounts for overdispersion in, and covariance among, components of Y. However, both parameters of the gamma distributions must be positive and thus the model can only capture positive covariances. Because this restriction is a severe limitation in many applications, we do not consider this model further.

2.2.2. *Multivariate Poisson-lognormal Mixture*

The lognormal mixture of independent Poisson distributions (Aitchison & Ho, 1989) provides a versatile alternative that allows for both overdispersion and negative covariances when modelling multivariate counts. Assume that $\boldsymbol{\delta} = (\delta_1, \cdots, \delta_K)$ follows a multivariate lognormal distribution whose associated multivariate normal has mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$, with elements σ_{kl}, for $k, l = 1, \cdots, K$. If we now assume that Y_k follows a Poisson distribution with mean δ_k, then, marginally,

$$\begin{aligned} E(Y_k) &= \exp\left(\mu_k + \frac{1}{2}\sigma_{kk}\right) = \alpha_k \\ V(Y_k) &= \alpha_k + \alpha_k^2\{\exp(\sigma_{kk}) - 1\} \\ \mathrm{Cov}(Y_k, Y_l) &= \alpha_k\alpha_l\{\exp(\sigma_{kl}) - 1\},\ k, l = 1, \cdots, K. \end{aligned} \tag{3}$$

Notice that the covariance structure of $\mathbf{Y}$ is determined by the covariance matrix $\boldsymbol{\Sigma}$ of the mixing component $\boldsymbol{\delta}$. This approach is appealing because it allows for modelling of the Poisson means as a function of covariates X and random effects $\boldsymbol{\epsilon} = (\epsilon_1, \cdots, \epsilon_K)$ such that:

$$\log \delta_k = \beta X_k + \epsilon_k,\ k = 1, \cdots, K \tag{4}$$

$$\boldsymbol{\epsilon} \sim N_K(0, \boldsymbol{\Sigma}). \tag{5}$$

Chib & Winkelmann (2001) were the first to provide a full Bayesian treatment of this model; additionally, they extended it by allowing the random effects $\boldsymbol{\epsilon}$ to follow a multivariate t distribution, providing flexibility to handle outliers. Building models that can capture covariances across space as well as among components of the response vector is a challenge. A number of proposals deal with multivariate area-level counts following ideas similar to those first proposed by Aitchison and Ho (1989). For example, Carlin and Banerjee (2003), and Gelfand and Vounatsou (2003) developed multivariate conditionally autoregressive (MCAR) models for hierarchical modeling of diseases. More recently, Jin et al. (2007) proposed an alternative to the MCAR model based on the linear model of coregionalization. Models for multivariate spatial processes are reviewed in Gelfand and Banerjee (2010). An overview of models for multivariate disease analysis, including models for count data, can be found in Lawson [chapter 9] (2009). In the context of point patterns, Møller et al. (1998) propose a multivariate Cox process directed by a log Gaussian intensity process.

3. MULTIVARIATE SPATIAL MODELS FOR SPECIES ABUNDANCE

Let $Y_k(s_{i_j})$ represent the number of individuals (counts) of species k, $k = 1, \cdots, K$, observed at location s_{i_j}, $j = 1, \cdots, n_i$ and sampling day $i = 1, \cdots, I$. Our aim is to propose a joint distribution for the vector $\mathbf{Y}(s) = (Y_1(s), \cdots, Y_K(s))'$.

Following section 2, we pursue a constructive approach based on hierarchical structures to induce the covariance structure among components of $\mathbf{Y}(s)$, and across locations of the region of interest. For each location s and sampling day i, we assume a positive continuous mixture of conditionally independent Poisson distributions, that is,

$$Y_k(s_{i_j}) \mid \theta_k(s_{i_j}), \delta_k(s_{i_j}) \sim Poi(\theta_k(s_{i_j})\delta_k(s_{i_j})).$$

We assume $\log \theta_k(s_{i_j}) = \mathbf{X}_k(s_{i_j})\boldsymbol{\beta}_k$, where $\mathbf{X}_k(s_{i_j})$ represents a p_k-dimensional row vector comprising environmental covariates at location s_{i_j} and a value of 1 associated

with the intercept, and $\boldsymbol{\beta}_k$ is the respective p_k dimensional vector of coefficients. We allow each component of $\boldsymbol{\theta}(s_{i_j}) = (\theta_1(s_{i_j}), \cdots, \theta_K(s_{i_j}))'$ to have its own set of coefficients and covariates. The parameter vector $\boldsymbol{\delta}(s_{i_j}) = (\delta_1(s_{i_j}), \cdots, \delta_K(s_{i_j}))'$ plays the role of the mixing component.

3.1. *Specification of the Mixing Components*

We now discuss the prior distribution of $\delta_k(s_{i_j})$. We assume an additive structure for time and space, that is

$$\log \delta_k(s_{i_j}) = \gamma_k(s_i) + \nu_k(s_{i_j}), \tag{6}$$

where each component of $\boldsymbol{\gamma}(s_i) = (\gamma_1(s_i) \cdots \gamma_K(s_i))'$ represents a species-specific, spatially homogeneous temporal effect for the set of locations $s_i = (s_{i_1}, \cdots, s_{i_{n_i}})$, observed at sampling day i. The elements of $\boldsymbol{\nu}(s_{i_j}) = (\nu_1(s_i) \cdots \nu_K(s_i))'$ capture local structure left in the data after adjusting the mean of the Poisson to account for covariate and temporal effects.

We formulate $\gamma_k(s_i)$ and $\nu_k(s_{i_j})$ as independent linear models of coregionalization. In the geostatistical literature, the LMC has been widely applied to capture correlation across space and among components of the response vector of interest (Foley and Fuentes, (2008); Berrocal et al., (2010)).

3.1.1. *Modelling the Temporal Effects*

We explore continuous correlation structures across time to account for the uneven spacing in time of the observations. Following the LMC approach, one can assume $\boldsymbol{\gamma}(s_i) = \mathbf{B}\boldsymbol{\upsilon}(i)$, where $\mathbf{B}$ is a lower triangular matrix, and the $\boldsymbol{\upsilon}(i) = (\upsilon_1(i), \cdots, \upsilon_K(i))'$ are independent components, each following a zero-mean Gaussian process with unit variance and correlation function $\rho(i, i'; \phi_{\gamma_k}) = \exp(-\phi_{\gamma_k}|t(i) - t(i')|)$, where $t(i)$ is the Julian day associated with the i^{th} sampling day. As shown in Gelfand et al. (2004) this leads to a non-separable covariance structure for the $\boldsymbol{\gamma}(.)$ random effects. We initially explored various model specifications under different priors for ϕ_{γ_k} and found little evidence for time-dependence of $\upsilon_k(i)$. Therefore, we consider here only a particular case of the coregionalization model in which each component of $\boldsymbol{\upsilon}(.)$ follows a standard normal distribution. This is similar to assuming $\boldsymbol{\gamma}(s_i) \sim N_K(0, \boldsymbol{\Omega})$, $\forall\, i = 1, \cdots, I$, where $\boldsymbol{\Omega} = \mathbf{B}\mathbf{B}^T$ captures the covariance among species at each time t. Considering the joint distribution of the KI dimensional vector, $\boldsymbol{\gamma} = (\boldsymbol{\gamma}(s_1), \cdots, \boldsymbol{\gamma}(s_I))'$, we have that $\boldsymbol{\gamma} \mid \boldsymbol{\Omega} \sim N(0, I_I \otimes \boldsymbol{\Omega})$, where I_I is the I-dimensional identity matrix.

3.1.2. *Modelling the Spatial Random Effects*

The $\boldsymbol{\nu}(s_{i_j}) = (\nu_1(s_{i_j}), \cdots, \nu_K(s_{i_j}))'$ can be viewed as a residual component left after adjusting for covariates and temporal effects. We allow this component to reflect covariances between species and across locations. Below we discuss different covariance structures for the random vector $\boldsymbol{\nu}(s_{i_j})$. Following the LMC, the general structure we assume is

$$\boldsymbol{\nu}(s_{i_j}) = \mathbf{A}\boldsymbol{\omega}(s_{i_j}), \tag{7}$$

where $\mathbf{A}$ is a K-dimensional lower triangular matrix. Each component of $\boldsymbol{\omega}(s_{i_j}) = (\omega_1(s_{i_j}) \cdots \omega_K(s_{i_j}))'$ is assumed to follow a zero-mean Gaussian process with unit variance and a specific correlation function, and $\omega_k(.)$ is assumed independent of $\omega_j(.)$, for $j \neq k = 1, 2, \cdots, K$.

Separable Covariance Structure for the Spatial Effects

We assume that all $\omega_j(.)$, $j = 1, \cdots, K$ follow a Gaussian process with zero mean, unit variance, and common correlation function $\rho(s - s'; \boldsymbol{\vartheta})$. Consequently, the nK dimensional vector obtained by stacking the $\boldsymbol{\nu}(s_{i_j})$, that is, $\boldsymbol{\nu} = (\boldsymbol{\nu}(s_{1_1}), \cdots, \boldsymbol{\nu}(s_{1_{n_1}}), \cdots, \boldsymbol{\nu}(s_{I_1}), \cdots, \boldsymbol{\nu}(s_{I_{n_I}}))'$, follows a zero-mean normal distribution with a separable covariance matrix given by $\mathbf{R} \otimes \mathbf{M}$, where $\mathbf{R}$ is the common correlation matrix of the spatial processes, and $\mathbf{M} = \mathbf{A}\mathbf{A}^T$ is the covariance among species. We assume an exponential correlation function for the spatial processes, such that $\rho(s - s'; \phi) = \exp(-\phi||s - s'||)$. This assumption of separability implies that after adjusting for the covariates and temporal effects there is a single (common among species) spatial structure left in the mean of the Poisson distribution.

Geometric Anisotropy in the Correlation Structure of the Spatial Effects

Schmidt et al. (2010b) examined the spatial distribution of counts for the most abundant fish species (yellow perch) in Lake Saint Pierre. They fitted models based on a wide range of assumptions about the spatial correlation structure, including spatial independence, geometrical anisotropy (Diggle & Ribeiro, 2007), and covariates in the correlation structure (Schmidt et al., 2010b), and found that stationary spatial effects did not provide adequate fits for this species. We consider here models based on linear transformations of the coordinate system. More specifically, the correlation function of each process $\omega_k(.)$ is given by $\rho(s, s'; \boldsymbol{\vartheta})$, with $d(s) = s\,D$, and $D = \begin{bmatrix} \cos\psi_A & -\sin\psi_A \\ \sin\psi_A & \cos\psi_A \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \psi_R^{-1} \end{bmatrix}$, where ψ_A is the *anisotropy angle* and $\psi_R > 1$ is the *anisotropy ratio* (Diggle & Ribeiro, 2007). The spatial correlation is given by $\rho(s, s'; \boldsymbol{\vartheta}) = \exp(-\phi||d(s) - d(s')||)$, with parameters $\boldsymbol{\vartheta} = (\phi, \psi_A, \psi_R)$.

Considering Covariates in the Correlation Structure of $\omega_j(.)$

A limitation of the geometrical anisotropic structure is that it can only capture anisotropy along a single direction that is constant across the study region. An alternative approach, which retains model simplicity while affording additional flexibility, is to allow for inclusion of covariates in the correlation structure of Gaussian processes (Schmidt et al., 2010a). We followed this approach to model the correlation structure of the spatial process as $\rho(s, s'; z(s), z(s'); \boldsymbol{\vartheta}) = \exp(-\phi_1||s - s'|| - \phi_2|z(s) - z(s')|)$, where $z(s)$ is a covariate that potentially influences the correlation structure of the latent spatial process, and $\boldsymbol{\vartheta} = (\phi_1, \phi_2)$ are parameters to be estimated. This correlation function yields a nonstationary structure in $\mathbb{R}^2$. The resulting *projection model* (Schmidt et al., 2010a) can be viewed as a particular case of the deformation approach introduced by Sampson & Guttorp (1992).

Let $\boldsymbol{\delta}$ be the nK-dimensional vector containing the elements of the mixing component, such that $\boldsymbol{\delta} = (\boldsymbol{\delta}(s_{1_1}), \cdots, \boldsymbol{\delta}(s_{I_{n_I}}))'$. Recalling that $\boldsymbol{\gamma} = (\gamma(s_1), \cdots, \gamma(s_I))'$ is the KI-dimensional vector of random effects shared by all locations visited on sampling day i, the resultant distribution of the logarithm of the mixing component in equation (6), $\log \boldsymbol{\delta}$, can be derived as follows. Let $\mathbf{C}$ be a $n \times I$ matrix, where $n = n_1 + \cdots + n_I$ is the total number of visited locations, and the rows are given by n_i replications of the I-dimensional row vector, $\mathbf{e}_i$, which has i^{th} element equal to 1 and all others equal to 0. Under this assumption of separability of the spatial effects, the distribution of the mixing component, conditioned on $\boldsymbol{\gamma}$, follows a normal distribution with mean $(I_K \otimes C)\boldsymbol{\gamma}$ and covariance structure $\mathbf{R} \otimes \mathbf{M}$. Integrating with respect to $\boldsymbol{\gamma}$, $\log \boldsymbol{\delta}$ follows a zero-mean normal distribution with covariance matrix $\boldsymbol{\Sigma} = (I_K \otimes \mathbf{C})(I_I \otimes \boldsymbol{\Omega})(I_K \otimes \mathbf{C})^T + (\mathbf{R} \otimes \mathbf{M})$.

Nonseparable Covariance Structure for the Spatial Effects

A more general, nonseparable, covariance structure is obtained by assuming that the $\omega_j(.)$ follow zero-mean, unit variance Gaussian processes, now with different correlation functions $\rho(s-s';\boldsymbol{\vartheta}_j)$. The covariance structure of $\boldsymbol{\omega}$ is then given by $\sum_{j=1}^{K}(\mathbf{R}_j \otimes \mathbf{M}_j)$ (Gelfand et al., 2004). If $\rho(s-s';\boldsymbol{\vartheta}_j)$ is stationary, the resultant correlation structure of $\nu_j(.)$ is a linear combination of stationary processes with different spatial ranges.

Different directional effects can also be incorporated to each of the latent spatial processes $\omega_j(.)$. In this case, $\boldsymbol{\vartheta}_j = (\phi_j, \psi_{A_j}, \psi_{R_j})$, $j = 1, \cdots, K$, such that $\rho(s,s';\boldsymbol{\vartheta}_j) = \exp(-\phi_j||d_j(s) - d_j(s')||)$, and each component of $\boldsymbol{\nu}(.)$ is a linear combination of processes with different directional components. Different decay parameters for the covariate can be used as well, providing the correlation structure $\rho(s,s';z(s),z(s');\boldsymbol{\vartheta}_j) = \exp(-\phi_{1j}||s-s'|| - \phi_{2j}|z(s)-z(s')|)$ for each $\omega_j(.)$; in this case, $\boldsymbol{\vartheta}_j = (\phi_{1j}, \phi_{2j})$, $j = 1, \cdots, K$. The last two models result in a nonseparable and non-stationary covariance structure for $\nu_j(.)$.

Now, following equation (6) and the previous definition of $\boldsymbol{\delta}$, the resultant marginal joint distribution of $\log \boldsymbol{\delta}$ is a zero-mean multivariate normal distribution, with covariance matrix $\boldsymbol{\Sigma} = (I_K \otimes \mathbf{C})(I_I \otimes \boldsymbol{\Omega})(I_K \otimes \mathbf{C})^T + \sum_{j=1}^{K}(\mathbf{R}_j \otimes \mathbf{M}_j)$.

3.2. *Inference Procedure*

Let $\mathbf{y} = (\mathbf{y}(s_{1_1}), \mathbf{y}(s_{1_2}), \cdots, \mathbf{y}(s_{1_{n_1}}), \cdots, \mathbf{y}(s_{I_1}), \cdots, \mathbf{y}(s_{T_{n_I}}))'$ be the observed counts over the sampling period at each location s_{i_j}, $i = 1, \cdots, I$, $j = 1, \cdots, n_i$. Conditional on $\boldsymbol{\theta}(s_{i_j})$ and $\boldsymbol{\delta}(s_{i_j})$, each observation is an independent realization from a Poisson distribution; therefore, the likelihood function is

$$l(\mathbf{y} \mid \boldsymbol{\theta}, \boldsymbol{\delta}) \propto \prod_{i=1}^{I}\prod_{j=1}^{n_i}\prod_{k=1}^{K} \exp\left\{-\theta_k(s_{i_j})\,\delta_k(s_{i_j})\right\}\left[\theta_k(s_{i_j})\,\delta_k(s_{i_j})\right]^{y_k(s_{i_j})}. \tag{8}$$

Prior Distribution of the Hyperparameters

We now specify the joint prior distribution of the hyperparameters in the model. We assume they are all independent *a priori*. For $\boldsymbol{\Omega}$, the among-species covariance matrix associated with the temporal random effect, we assign an inverse Wishart prior distribution with $K+2$ degrees of freedom and a diagonal scale matrix. The scale elements are given by estimates of the residual standard error from independent log-linear fits for each species. The prior distribution for the coregionalized matrix $\mathbf{A}$ should not be too vague. There are two options when assigning this prior distribution. As $\mathbf{M} = \mathbf{A}\mathbf{A}^T$ is a covariance matrix, one possibility is to assign an inverse Wishart prior distribution to $\mathbf{M}$ and make use of the relationship between $\mathbf{M}$ and $\mathbf{A}$ to obtain the prior distribution for $\mathbf{A}$. This approach is pursued by Gelfand et al. (2004) and Jin et al. (2007). The alternative we choose here assigns independent, zero-mean normal distributions to each of the off-diagonal elements of $\mathbf{A}$, and a lognormal distribution to the diagonal elements of $\mathbf{A}$. The variances of these normal prior distributions were fixed at 5, yielding reasonably vague prior distributions for $\mathbf{M}$.

For the geometrical anisotropic structure, we assign a uniform prior for ψ_A in the interval $(0, \pi)$. Given that $\psi_{R_k} > 1$, we assign a Pareto prior distribution with parameters $\psi_{R_M} = 1$ and $\alpha = 2$, where $p(\psi_{R_k}|\alpha, \psi_M) = \frac{\alpha\psi_{R_M}}{\psi_{R_k}^{\alpha+1}}$ for $\psi_{R_k} > 1$. The

mean of the Pareto is only defined for $\alpha > 1$ and the variance for $\alpha > 2$; the choice of $\alpha = 2$ provides a fat-tailed distribution with infinite variance. For ϕ_{ij}, $i = 1, 2$, $j = 1, \cdots, K$, the decay parameter of the exponential correlation function associated with the j^{th} Gaussian process, we assign an inverse gamma prior distribution with parameters a and b. We fix $a = 2$ providing a distribution with infinite variance, and set the mean based on the idea of practical range. The prior mean of the ϕ_{ij}, b_i, is fixed such that the practical range (correlation = 0.05) is reached at half of the maximum distance between observations, i.e., $-\log(0.05) = b_1 \ \frac{max(||s-s'||)}{2}$ and $-\log(0.05) = b_2 \ \frac{max(|z-z'|)}{2}$.

Following Bayes theorem, the posterior distribution is proportional to the prior distribution times the likelihood function. The posterior distribution resulting from the likelihood in equation (8) and the prior discussed above does not have known closed form. We use Markov chain Monte Carlo (MCMC) methods, specifically, the Gibbs sampler with some Metropolis-Hastings (M-H) steps (Gamerman & Lopes, 2006), to obtain samples from the target posterior distribution.

MCMC Sampling Scheme

To increase the efficiency of the MCMC sampling scheme, we reparametrize the model proposed above. Let $\log \varphi_k(s_{i_j}) = X_k^*(s_{i_j})\boldsymbol{\beta}_k^* + W_k(s_{i_j})$ and $W_k(s_{i_j}) = \beta_{1k} + \gamma_k(s_i) + \nu_k(s_{i_j})$, where $X_k^*(.)$ does not have a column of ones, and $\boldsymbol{\beta}_k^* = (\beta_{2k}, \cdots, \beta_{p_k k})^T$. The posterior full conditional distributions of the covariate coefficients and of $W_k(s_{i_j})$ are not known. We use the M-H algorithm proposed by Gamerman (1997) to sample from these distributions. Let $\mathbf{W}$ and $\boldsymbol{\nu}$ be the vectors obtained by stacking the $\mathbf{W}(s_{i_j})$ and the $\boldsymbol{\nu}(s_{i_j})$. We can write $\mathbf{W} = (I_K \otimes \mathbf{1}_n)\boldsymbol{\beta}_{1\cdot} + (I_K \otimes \mathbf{C})\boldsymbol{\gamma} + \boldsymbol{\nu}$, which follows a Kn-dimensional multivariate normal distribution, with mean $(I_K \otimes \mathbf{1}_n)\boldsymbol{\beta}_{1\cdot} + (I_K \otimes \mathbf{C})\boldsymbol{\gamma}$ and covariance matrix $\sum_{j=1}^{K}(\mathbf{R}_j \otimes \mathbf{M}_j)$. Under this reparametrization the posterior full conditional distribution of $\boldsymbol{\beta}_{1\cdot} = (\beta_{11}, \cdots, \beta_{1K})'$ follows a multivariate normal distribution, and so does the posterior full conditional distribution of $\boldsymbol{\gamma}$. The parameters in the spatial correlation function result in unknown posterior full conditionals and we use M-H steps to sample from them. To sample the decay parameters ϕ_1 and ϕ_2 we use a log-normal proposal based on the current value of the chain, with suitably tuned variance. Parameters ψ_A and ψ_R in the elliptical anisotropy model are truncated. For the former, we apply a transformation to the real line that allows for use of normal proposal distributions; for the latter, we sample from a truncated normal distribution based on the current value of the chain. The full conditional posterior distribution for $\boldsymbol{\Omega}$ follows a inverse Wishart distribution. The elements of $\mathbf{A}$ do not follow a known posterior full conditional; we therefore use M-H steps to sample from them. The MCMC algorithms were implemented in Ox (Doornik and Ooms, 2007), a programming language that deals very efficiently with matrices and vectors.

4. DATA ANALYSIS

Ecological interpretation and understanding can be enhanced by focusing on general rules that are largely invariant in space and time. An aim of particular interest is to extract those components of species-environment relationships that hold across different settings, e.g., sites or years, and can be described parsimoniously. Accordingly, we assume that coefficients of the environmental covariates, which describe local species responses to their immediate environment, do not vary across locations in the lake. Temporal random effects were assumed to be independent across time

and common to both shores. However, we allow for different spatial random effects for the North and South shores because previous studies point to marked differences in the spatial structure of fish growth (Glémet & Rodríguez, 2007) and abundance (Schmidt et al., 2010b) between the two shores.

4.1. *Fitted Models*

We fit six models that aim to explain joint variation in abundances of the four fish species described in subsection 1.1. The models differ primarily in the specification of the covariance structure of each component of $\boldsymbol{\omega}(.)$ in equation (7). We consider both separable and non-separable covariance structures for stationary and non-stationary cases:

M1 : Separable isotropic covariance structure;

M2 : Separable elliptical anisotropy covariance structure;

M3 : Separable covariate-dependent (z =geodetic depth) covariance structure;

M4 : Non-separable isotropic covariance structure;

M5 : Non-separable elliptical anisotropy covariance structure;

M6 : Non-separable, covariate-dependent (z =geodetic depth) covariance structure.

For each model we ran two chains for $L = 60,000$ iterations, starting from suitably dispersed initial values for the parameters. We discarded the first $10,000$ sampled values (burn in), and subsequently kept every 50^{th} iteration to reduce autocorrelation of the sampled values. Convergence of the chains was checked by examining the trace plots.

4.2. *Model Comparisons*

Model comparison was performed using two different criteria: (i) the deviance information criterion (DIC) (Spiegelhalter et al., 2002), and (ii) the expected predictive deviance (EPD), a measure of posterior predictive loss (Gelfand & Ghosh, 1998). Disaggregated values for both criteria were computed for individual fish species.

DIC is a generalization of the AIC based on the posterior distribution of the deviance, $D(\boldsymbol{\theta}) = -2 \log l(\mathbf{y} \mid \boldsymbol{\theta}, \boldsymbol{\delta})$. More formally, the DIC is defined as

$$DIC = \overline{D} + p_D = 2\overline{D} - D(\overline{\boldsymbol{\theta}}),$$

where $\overline{D}$ defines the posterior expectation of the deviance, $\overline{D} = E_{\boldsymbol{\theta}|\mathbf{y}}(D)$, p_D is the effective number of parameters, $p_D = \overline{D} - D(\overline{\boldsymbol{\theta}})$, and $\overline{\boldsymbol{\theta}}$ is the posterior mean of the parameters. $\overline{D}$ can be viewed as a measure of goodness of fit, whereas p_D reflects the complexity of the model. Smaller values of DIC indicate better fitting models.

EPD is based on replicates of the observed data, $Y_{i,rep}$, $i = 1, \cdots, n$, where i stands for the i^{th} sampling unit. The selected models are those that perform best under a loss function such as the deviance, a familiar discrepancy-of-fit measurement

Yellow perch

Model	p_D	DIC	EPD
M1	144.8	1039.0	2019063.4
M2	145.2	1038.6	2019061.1
M3	137.1	1031.3	2019023.1
M4	145.1	1040.0	2019030.6
M5	144.0	1036.4	2019067.2
M6	136.5	1032.1	2018988.0

Brown bullhead

Model	p_D	DIC	EPD
M1	109.4	626.2	133777.8
M2	110.1	628.3	133778.2
M3	101.1	621.3	133755.5
M4	111.9	633.7	133768.1
M5	112.6	635.1	133770.4
M6	101.1	622.9	133754.1

Golden shiner

Model	p_D	DIC	EPD
M1	106.0	590.1	246747.8
M2	106.0	591.7	246729.9
M3	98.6	585.4	246714.3
M4	103.2	592.1	246734.8
M5	105.5	596.2	246741.1
M6	99.6	590.6	246699.6

Pumpkinseed

Model	p_D	DIC	EPD
M1	81.9	432.4	54044.9
M2	81.4	429.8	54056.2
M3	75.6	428.9	54037.7
M4	82.9	434.0	54050.2
M5	82.8	434.7	54046.9
M6	75.7	432.7	54038.1

Table 1: Values of the effective number of parameters, p_D, DIC, and EPD for the six fitted models, by fish species.

for generalized linear models (Gelfand & Ghosh, 1998). Under this loss function the EPD for model m results in

$$D_l^{(m)} = 2\sum_{i=1}^{n} w_i\{t_i^{(m)} + l\, t(y_{i,obs})\} + 2(l+1)\sum_{i=1}^{n} w_i \left\{ \frac{t(\mu_i^{(m)}) + l\, t(y_{i,obs})}{l+1} - t\left(\frac{\mu_i^{(m)} + l y_{i,obs}}{l+1}\right)\right\},$$

where, in the Poisson case, $w_i = 1$, $t(y) = y\,\log(y) - y$, $t_i^{(m)} = E[t(y_{i,rep}) \mid y_{i,obs}, m]$. EPD is not sensitive to the choice of the arbitrary constant l (Gelfand & Ghosh, 1998), which was fixed here to $l = 100$. In the EPD equation above, $\mu_i = E(Y_{i,rep} \mid \mathbf{y})$ is the mean of the predictive distribution of $Y_{i,rep}$ given the observed data $\mathbf{y}$. At each iteration of the MCMC we obtain replicates of the observations given the sampled values of the parameters and then compute $D_l^{(m)}$ using Monte Carlo integration techniques. The smallest value of $D_l^{(m)}$ indicates the best fitted model.

Model comparisons based on DIC and EPD (Table 1) indicated that models including geodetic depth as a covariate in the correlation structure of the spatial effects generally fit better than those assuming isotropy or geometrical anisotropy. In particular, EPD generally pointed to M6 as the best model among those fitted.

Posterior predictive distributions for counts of the four species under M6 showed agreement with observed counts (Figure 2).

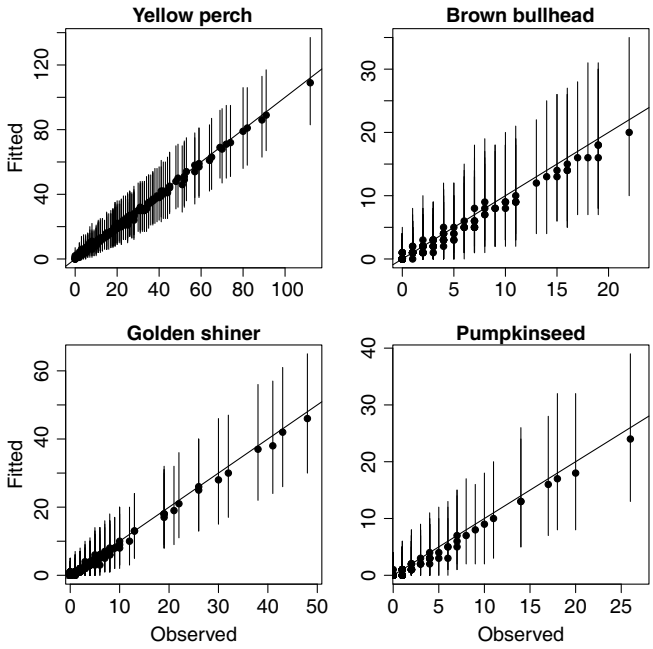

Figure 2: *Summary of the fitted (predictive distribution under model M6) versus observed counts, by species. The posterior mean of the predictive distribution (circles) and 95% posterior predictive interval (vertical lines) are shown.*

The regression coefficients for environmental covariates did not appear to depend strongly on the prior structure assumed for the spatial random effects (Figure 3). M6 indicated that transparency had positive influence on the abundance of three of the four fish species, whereas water depth (negative effect on yellow perch) and vegetation (positive effect on brown bullhead) each influenced the abundance of one fish species. Substrate composition had no apparent effect on fish abundances. The results for water transparency are consistent with previous work showing that mortality risk for fish in Lake Saint Pierre generally declines as transparency increases (Laplante-Albert et al., 2010).

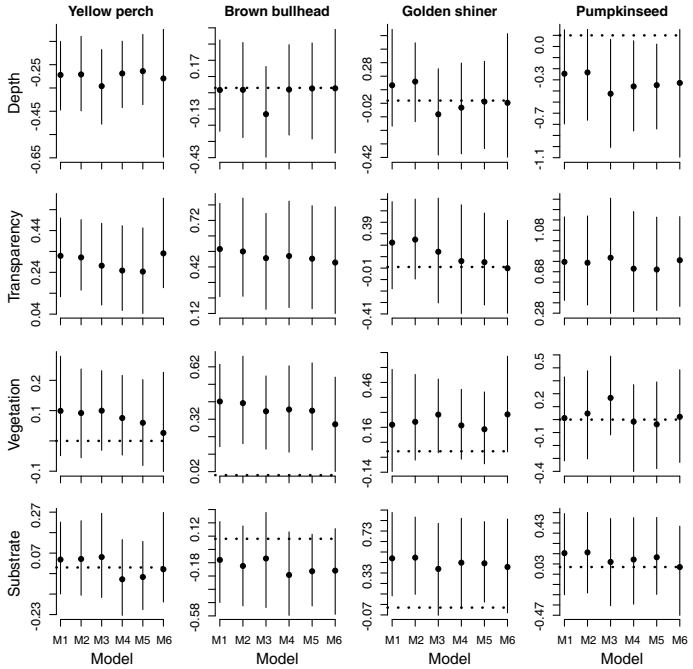

Figure 3: *Posterior summary (median and 95% credible interval) of regression coefficients for the four environmental covariates (rows) under each fitted model, by fish species (columns). Dotted lines indicating a value of zero for the regression coefficients are provided for reference.*

The random effects $\gamma_k(.)$ (not presented here) showed no seasonal temporal trend for any of the species. The off-diagonal elements of covariance matrix $\mathbf{\Omega}$, which provide the covariances between components of $\boldsymbol{\gamma}(.)$, did not significantly differ from zero. The posterior sample of $\mathbf{M}$ was used to obtain both the variances of $\nu_k(.)$ and the correlations between species at a specific location (Figure 4). Variances tended to be greater in the North shore than in the South shore. Correlations between species also tended to be stronger in the North shore. The strongest correlations were positive, between pumpkinseed and yellow perch, and between pumpkinseed and golden shiner, both in the North shore. Although the models allow for negative

correlations, which would be expected, say, in the presence of strong competition between species, none was apparent among any of the species pairs.

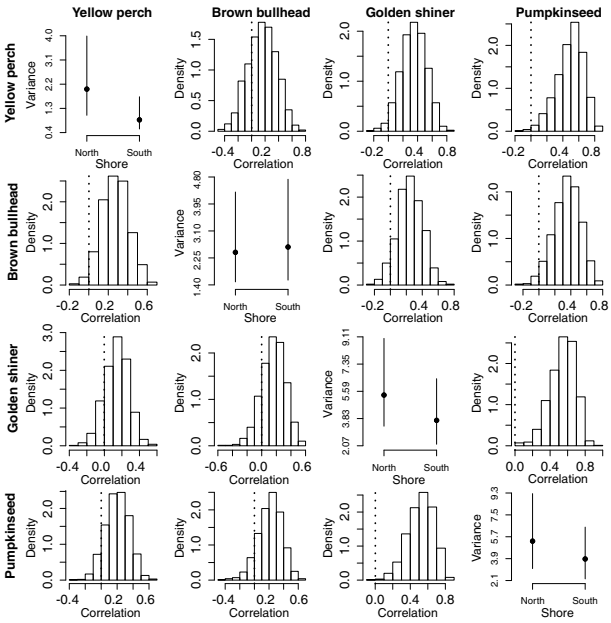

Figure 4: *Estimates of species variance and between-species correlation under M6. Posterior summary (median and 95% credible intervals) of variances of* $\nu_k(.)$*, by shore (main diagonal), and posterior distribution of correlations between species for the North (above main diagonal) and South (below main diagonal) shores. Dotted lines indicating a value of zero for the correlations are provided for reference.*

Inspection of decay parameters of exponential correlation functions for the fitted models (not shown here) showed rapid decay of spatial correlation structures for the South shore locations. We therefore focus on spatial pattern for North shore locations. The spatial correlation between two arbitrary points can be obtained following Section 8 in Gelfand et al. (2004). Maps of the spatial correlation between a fixed central location and all other locations (Figure 5) show contrasting patterns of decay among models and between two selected species, yellow perch (slow decay) and brown bullhead (rapid decay). As expected, decay does not depend on direction under isotropic model M4. Decay is slower in the SE - NE direction under the geometric anisotropic M5. Decay under model M6, which includes geodetic depth as a covariate in the covariance structure of the spatial process, has directionality generally similar to that of M5, but is conspicuously modified by geodetic lake depth. Such modifications are particularly evident in groups of locations having similar distance from the central point but different geodetic depth (cf. Figure 5 and Figure 1). These results are consistent with the biological intuition that observations at similar geodetic depths can be influenced by common environmental or behavioural processes that induce spatial correlation (see *Ecological Motivation and Data* above).

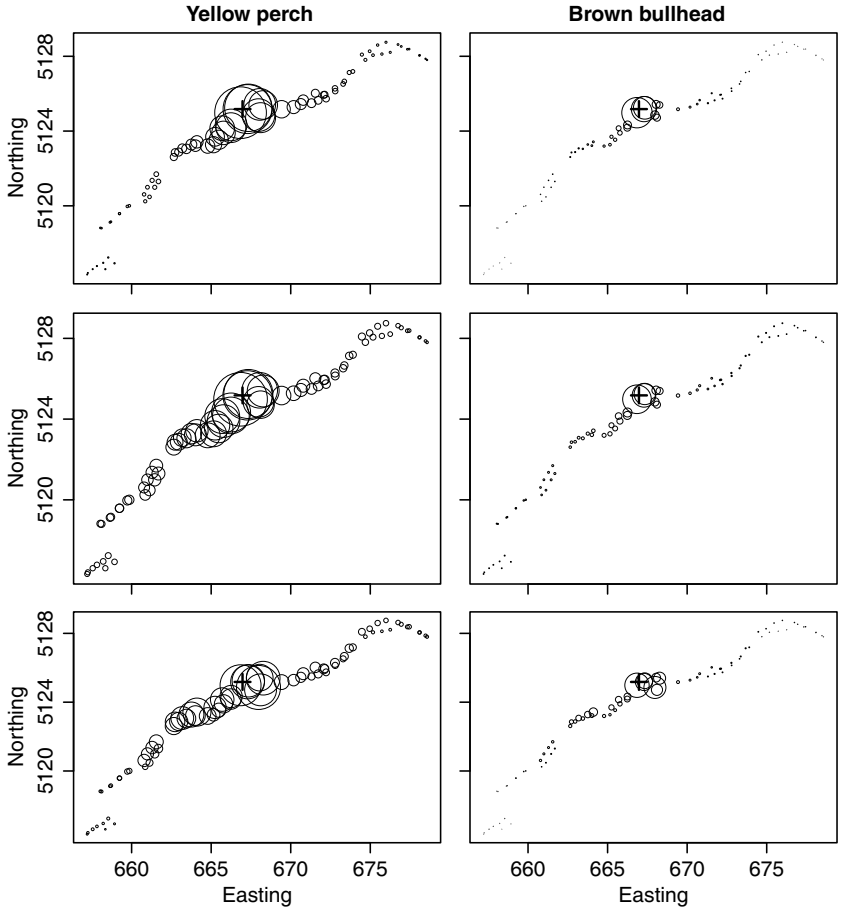

Figure 5: *Posterior means of the spatial correlation of $\nu_k(.)$ between a central location (+) and all other locations along the North shore of the lake (symbol sizes proportional to correlation). Results are presented separately for models M4, M5, and M6 (rows) and two fish species, yellow perch and brown bullhead (columns).*

5. CONCLUSIONS

We discussed models for multivariate counts observed at fixed spatial locations within a region. The use of continuous mixtures of Poisson distributions allowed for overdispersion in the components of the response vector and both positive and negative covariance among components. The linear model of coregionalization provided flexible covariance structures for the mixing component. We explored spatial covariance structures that allow for non-stationarity of the latent spatial process while retaining model simplicity. In particular, including information on geodetic lake depth in the spatial covariance structure of the spatial process provided a flexible means of capturing anisotropy along the shorelines of a lake. Our results suggest that further exploration of flexible covariance structures for the spatial components may prove fruitful.

Many challenges in modelling multivariate counts remain open. For example, in some data sets, specific features, such as an excess of observed zeros, and spatial or temporal dependence may be observed in some components of the response vector but not in others. Future research could explore models that allow for such differences yet still capture correlations between components of the response vector. In our study, some species showed very rapid decay in the spatial correlation of the spatial component, indicating that independent structures could be considered for this particular effect. We are currently investigating how to adapt the LMC to cope with this kind of situation.

REFERENCES

Aitchison, J. and Ho, C. H. (1989). The multivariate Poisson-log normal distribution. *Biometrika* **76**, 643–653.

Berrocal, V. J., Gelfand, A. E. and Holland, D. M. (2010) A spatio-temporal downscaler for output from numerical models. *J. Agricul. Biol. Environm. Stat.*, DOI No. 10.1007/s13253-009-0004-z.

Carlin, B. P. and Banerjee, S. (2003). Hierarchical multivariate CAR models for spatio-temporally correlated survival data. *Bayesian Statistics 7* (J. M. Bernardo, M. J. Bayarri, J. O. Berger, A. P. Dawid, D. Heckerman, A. F. M. Smith and M. West, eds.) Oxford: University Press, 45–63 (with discussion).

Chib, S. and Winkelmann, R. (2001). Markov chain Monte Carlo analysis of correlated count data. *J. Business & Economic Statistics*, **19**, 428–435.

Diggle, P. J. and Ribeiro Jr., P. J. (2007). *Model-based Geostatistics* New York: Springer.

Doornik, J. A. and Ooms, M. (2007). *Introduction to Ox: An Object-Oriented Matrix Language*, London: Timberlake Consultants Press.

Foley, K. M. and Fuentes, M. (2008) A statistical framework to combine multivariate spatial data and physical models for hurricane surface wind prediction. *J. Agricul. Biol. Environm. Stat.*, **13**, 37–59.

Gamerman, D. (1997). Sampling from the posterior distribution in generalized linear mixed models. *Statist. Computing* **7**, 57–68.

Gamerman, D. and Lopes, H. F. (2006). *Markov Chain Monte Carlo - Stochastic Simulation for Bayesian Inference.* 2nd Edition, London: Chapman and Hall.

Gelfand, A. E. and Ghosh, S. K. (1998). Model choice: a minimum posterior predictive loss approach. *Biometrika* **85**, 1–11.

Gelfand, A. E. and Vounatsou, P. (2003). Proper multivariate conditional autoregressive models for spatial data analysis. *Biostatistics* **4**, 11–25.

Gelfand, A. E. and Banerjee, S. (2010). Multivariate Spatial Process Models. *Handbook of Spatial Statistics* (A. E. Gelfand, P. J. Diggle, M. Fuentes and P. Guttorp, eds.) Boca Raton, FL: Chapman & Hall/CRC, 495–515.

Gelfand, A. E., Schmidt, A. M., Banerjee, S. and Sirmans, C. F. (2004). Nonstationary multivariate process modeling through spatially varying coregionalization (with discussion). *Test* **13**, 263–312.

Glémet, H. and Rodríguez, M. A. (2007). Short-term growth (RNA/DNA ratio) of yellow perch (*Perca flavescens*) in relation to environmental influences and spatio-temporal variation in a shallow fluvial lake. *Can. J. Fish. Aquat. Sci.*, **64**, 1646–1655.

Jin, X. Banerjee, S. and Carlin, B. P. (2007). Order-free coregionalized areal data models with application to multiple disease mapping. *J. Roy. Statist. Soc. B* **69**, 817–838.

Karlis, D. and Meligkotsidou, L. (2005). Multivariate Poisson regression with covariance structure. *Statist. Computing* **15**, 255–265.

Karunanayake, C. (2007). Multivariate Poisson hidden Markov Models for analysis of spatial counts. Ph.D. Thesis, Department of Mathematics and Statistics, University of Saskatchewan, Saskatoon.

Laplante-Albert, K.A., Rodríguez, M. A., and Magnan, P. (2010). Quantifying habitat-dependent mortality risk in lacustrine fishes by means of tethering trials and survival analyses. *Env. Biol. Fish.*, **87**, 263–273.

Lawson, A. B. (2009). *Bayesian Disease Mapping - Hierarchical Modeling in Spatial Epidemiology.* London: Chapman and Hall

Mathai, A. M. and Moschopoulos, P. G. (1991). On a multivariate gamma. *J. Multivariate Analysis* **39**, 135–153.

Møller, J., Syversveen, A. R. and Waagepetersen, R. P. (1998). Log Gaussian processes. *Scandinavian J. Statist.* **25**,451–482.

Sampson, P. and Guttorp, P. (1992). Nonparametric estimation of nonstationary spatial covariance structure. *J. Amer. Statist. Assoc.* **87**, 108–119.

Schmidt, A. M. and Gelfand, A. E. (2003). A Bayesian coregionalization model for multivariate pollutant data. *J. Geophys. Res - Atmospheres*, **108(D24)**, 8783, doi:10.1029/2002JD002905.

Schmidt, A. M., Guttorp, P. and O'Hagan, A. (2010a). Considering covariates in the covariance structure of spatial processes. *Tech. Rep.*, Departamento de Métodos Estatísticos, IM-UFRJ, Brazil.

Schmidt, A. M., Rodríguez, M. A. and Capistrano, E. S. (2010b). Accounting for latent spatio-temporal structure in animal abundance models. *Tech. Rep.*, Departamento de Métodos Estatísticos, IM-UFRJ, Brazil.

Spiegelhalter, D., Best, N., Carlin, B. and Linde, A. (2002). Bayesian measures of model complexity and fit (with discussion). *J. Roy. Statist. Soc. B* **64**, 583–639.

Tsionas, E. (2001). Bayesian multivariate Poisson regression. *Communications in Statistics - Theory and Methods*, **30**, 243–255.

Tsionas, E. (2004). Bayesian inference for multivariate gamma distributions. *Statist. Computing* **14**, 223–233.

Wackernagel, H. (2003). *Multivariate Geostatistics - An Introduction with Applications*, 3rd Edition. New York: Springer.