Statistica Sinica **20** (2010), 675-695

A GENERALIZED CONVOLUTION MODEL FOR MULTIVARIATE NONSTATIONARY SPATIAL PROCESSES

Anandamayee Majumdar, Debashis Paul and Dianne Bautista

Arizona State University, University of California, Davis and Ohio State University

Abstract: We propose a flexible class of nonstationary stochastic models for multivariate spatial data. The method is based on convolutions of spatially varying covariance kernels and produces mathematically valid covariance structures. This method generalizes the convolution approach suggested by Majumdar and Gelfand (2007) to extend multivariate spatial covariance functions to the nonstationary case. A Bayesian method for estimation of the parameters in the covariance model based on a Gibbs sampler is proposed, then applied to simulated data. Model comparison is performed with the coregionalization model of Wackernagel (2003) that uses a stationary bivariate model. Based on posterior prediction results, the performance of our model appears to be considerably better.

Key words and phrases: Convolution, nonstationary process, posterior inference, predictive distribution, spatial statistics, spectral density.

1. Introduction

Spatial modeling with flexible classes of covariance functions has become a central topic of spatial statistics in recent years. One of the traditional approaches to modeling spatial stochastic processes is to consider parametric families of stationary processes, or processes that can be described through parametric classes of semi-variograms (Cressie (1993)). However, in spite of its simplicity, computational tractability, and interpretability, the stationarity assumption is often violated in practice, particularly when the data come from large, heterogeneous, regions. In various fields of applications, like soil science, environmental science, etc., it is often more reasonable to view the data as realizations of processes that only in a small neighborhood of a location behave like stationary processes. Also, it is often necessary to model two or more processes simultaneously and account for the possible correlation among various coordinate processes. For example, Majumdar and Gelfand (2007) consider an atmospheric pollution data consisting of 3 pollutants : CO, NO and NO_2, whose concentrations in the atmosphere are correlated. A key question studied in this paper is modeling this correlation among the various coordinates while allowing for nonstationarity in space for the

multivariate process. We propose a flexible semiparametric model for multivariate nonstationary spatial processes. After reviewing the existing literature on nonstationary spatial modeling.

A considerable amount of work over the last decade or so has focussed on modeling *locally stationary* processes (Fuentes (2002), Fuentes, Chen, Davis and Lackmann (2005), Gelfand, Schmidt, Banerjee and Sirmans (2004), Higdon (1997), Paciorek and Schervish (2006) and Nychka, Wikle, and Royle (2002)). Dahlhaus (1996, 1997) gives a more formal treatment of locally stationary processes in the time series context in terms of evolutionary spectra of time series. This research on the modeling of nonstationary processes might be thought of as the semi-parametric modeling of covariance functions. Higdon (2002) and Higdon, Swall, and Kern (1999) model the process as a convolution of a stationary process with a kernel of varying bandwidth. Thus, the observed process $Y(s)$ is of the form $Y(s) = \int K_s(x)Z(x)dx$, where $Z(x)$ is a stationary process, and the kernel K_s depends on the location s. Fuentes (2002) and Fuentes and Smith (2001) consider a convolution model in which the kernel has a fixed bandwidth, while the process has a spatially varying parameter. Thus,

$$Y(s) = \int_D K(s-x)Z_{\theta(x)}(s)dx, \tag{1.1}$$

where $\{Z_{\theta(x)}(\cdot) : x \in D\}$ is a collection of independent stationary processes with covariance function parameterized by the function $\theta(\cdot)$. Nychka, Wikle, and Royle (2002) consider a multiresolution analysis-based approach to model the spatial inhomogeneity that utilizes the smoothness of the process and its effect on the covariances of the basis coefficients, when the process is represented in a suitable wavelet-type basis.

One of the central themes of the various modeling schemes described above is that a process may be represented in the spectral domain *locally* as a superposition of Fourier frequencies with suitable (possibly spatially varying) weight functions. Recent work of Pintore and Holmes (2006) provides a solid mathematical foundation to this approach. Paciorek and Schervish (2006) derive an explicit representation for the covariance function for Higdon's model when the kernel is multivariate Gaussian and use it to define a nonstationary version of the Matérn covariance function by utilizing the *Gaussian scale mixture* representation of positive definite functions. Also, there are works on a different type of nonstationary modeling through spatial deformations (see e.g., Sampson and Guttorp (1992)), but they do not concern us here.

The modeling approaches mentioned so far focus primarily on one-dimensional processes. In this paper, our main focus is on modeling nonstationary, multidimensional spatial processes. Existing approaches to modeling the multivariate

processes include the work by Gelfand, Schmidt, Banerjee and Sirmans (2004) that utilizes the idea of *coregionalization* to model the covariance of $\mathbf{Y}(s)$ (taking values in $\mathbb{R}^N$) as

$$\operatorname{Cov}(\mathbf{Y}(s), \mathbf{Y}(s')) = \sum_{j=1}^{N} \rho_j(s - s')\mathbf{T}_j,$$

where $\rho_j(\cdot)$ are stationary covariance functions, and $\mathbf{T}_j$ are positive semidefinite matrices of rank 1. Christensen and Amemiya (2002) consider a different class of multivariate processes that depend on a latent shifted-factor model structure.

Our work can be viewed as a generalization of the convolution model for correlated Gaussian processes proposed by Majumdar and Gelfand (2007). We extend their model to nonstationary settings. A key motivation is the assertion that when spatial inhomogeneity in the process is well-understood in terms of dependence on geographical locations, it makes sense to use that information directly in the specification of the covariance kernel. For example, soil concentrations of Nitrogen, Carbon, and other nutrients and/or pollutants, that are spatially distributed, are relatively homogenous across similar land-use types (e.g., agricultural, urban, desert, transportation - and so on), but are non-homogeneous across spatial locations with different land-use types. Usually the land-use types and their boundaries are clearly known (typically from satellite imagery). This is then an instance when nonstationary models are clearly advantageous compared to stationary models. Another example concerns land-values and different economic indicators in a spatial area. Usually land-values are higher around (possibly multiple) business centers, and such information may be incorporated in the model as the known centers of the kernels at (3.1). It is also important for modeling multidimensional processes that the degree of correlations among the coordinate processes across different spatial scales is allowed to vary. Keeping these goals in mind, we present a class of models that behave locally like stationary processes, but are *globally* nonstationary. The main contributions of this paper are: (i) specification of the multivariate spatial cross-covariance function in terms of Fourier transforms of spatially varying spectra; (ii) incorporation of correlations among coordinate processes that vary with both frequency and location; (iii) derivation of precise mathematical conditions under which the process is nonsingular; and (iv) the provision for including local information about the process (e.g., smoothness, scale of variability, gradient of spatial correlation along a given direction) directly into the covariance model. The last goal is achieved by expressing the spatially varying coordinate spectra $f_j(s, \omega)$ (as in (2.6)) as a sum of kernel-weighted stationary spectra, where the kernels have known shapes and different (possibly pre-specified) centers, bandwidths and orientations. We also

present a Bayesian estimation procedure based on *Gibbs sampling* for estimating a specific parametric covariance function and study its performance through simulation studies.

The paper is organized as follows. We specify the model and discuss its properties in Section 2. In Section 3, we propose a special parametric subclass that is computationally easier to deal with. Also, we discuss various aspects of the model, such as parameter identifiability and the relation to some existing models, by focussing attention on a special bivariate model. In Section 4, we give an outline of a simulation study that illustrates the characteristics of the various processes generated by our model in the two-dimensional setting. In Section 5, we present a Bayesian estimation procedure and conduct a simulation study to demonstrate its effectiveness. In Section 6, we discuss some related research directions. Some technical details and a detailed outline of the Gibbs sampling procedure for posterior inference are given in the *supplementary material*.

2. Construction of Covariances Through Convolution

We consider a real-valued point-referenced univariate spatial process, $Y(s)$, associated with locations $s \in \mathbb{R}^d$. In this section, we construct a Gaussian spatial process model for an arbitrary finite set of locations in a region $D \subset \mathbb{R}^d$ by generalizing the construction of Majumdar and Gelfand (2007), and then extend it to whole of $\mathbb{R}^d$.

2.1. Nonstationary covariance structure on a finite set in $\mathbb{R}^d$

In this subsection, we construct a class of nonstationary multivariate stochastic processes on a finite set of points in $\mathbb{R}^d$. Assume that the points $\{s_l : l = 1, \ldots, k\}$ in $\mathbb{R}^d$ are given. Let $\{C_{jl} : j = 1, \ldots, N;\ l = 1, \ldots, k\}$ be a set of stationary covariance kernels on $\mathbb{R}^d$ with corresponding spectral density functions $\{f_{jl} : j = 1, \ldots, N; l = 1, \ldots, k\}$ defined by

$$f_{jl}(\omega) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{-i\omega^T s} C_{jl}(s) ds, \quad \omega \in \mathbb{R}^d.$$

Consider the $Nk \times Nk$ matrix $\mathbf{C}$, whose (j, j')th entry in the (l, l')th block, for $1 \le j, j' \le N$ and $1 \le l, l' \le k$, is denoted by $c_{jl,j'l'}$, and is expressed as

$$c_{jl,j'l'} \equiv C^{\star}_{jj'}(s_l, s_{l'}) = \int_{\mathbb{R}^d} e^{i\omega^T (s_l - s_{l'})} f_{jl}(\omega) f_{j'l'}(\omega) \rho_{jj'}(\omega) \rho^0_{ll'}(\omega) d\omega, \tag{2.1}$$

where $\rho_{jj'}(\cdot)$ are complex-valued functions satisfying $\rho_{jj'}(\omega) = \overline{\rho_{j'j}}(\omega)$, and $((\rho^0_{ll'}(\omega)))^k_{l,l'=1}$ is a non-negative definite matrix for every $\omega \in \mathbb{R}^d$. Thus, $C^{\star} := ((C^{\star}_{jj'}))^N_{j,j'=1}$ is function from $\mathbb{R}^d \times \mathbb{R}^d$ to $\mathbb{R}^{N \times N}$. We require that $\max\{\max_{j,j'} |\rho_{jj'}(\omega)|, \max_{l,l'} |\rho^0_{ll'}(\omega)|\} \le 1$ for all $\omega \in \mathbb{R}^d$.

We show that under appropriate conditions, the $Nk \times Nk$ matrix $\mathbf{C} = ((c_{jl,j'l'}))$ is a non-negative definite matrix. The (l,l')th block (of size $N \times N$) of the matrix $\mathbf{C}$, for $1 \le l, l' \le k$, is

$$C_{ll'} = \begin{pmatrix} C^\star_{11}(s_l, s_{l'}) & \dots & C^\star_{1N}(s_l, s_{l'}) \\ \vdots & \ddots & \vdots \\ C^\star_{N1}(s_l, s_{l'}) & \dots & C^\star_{NN}(s_l, s_{l'}) \end{pmatrix}. \tag{2.2}$$

For all $\omega \in \mathbb{R}^d$, define $A_{ll'}(\omega)$, for $1 \le l, l' \le k$, as

$$A_{ll'}(\omega) = e^{i\omega^T(s_l - s_{l'})}\rho^0_{ll'}(\omega) \begin{pmatrix} (f_{1l}(\omega))^2\rho_{11}(\omega) & \dots & f_{1l}(\omega)f_{Nl}(\omega)\rho_{1N}(\omega) \\ \vdots & \ddots & \vdots \\ f_{Nl}(\omega)f_{1l}(\omega)\rho_{N1}(\omega) & \dots & (f_{Nl}(\omega))^2\rho_{NN}(\omega) \end{pmatrix}, \tag{2.3}$$

where the $f_{jl}(\omega)$'s are as defined above. Let $e(\omega)$ be the $k \times k$ matrix with (l,l')th entry $e^{i\omega^T(s_l - s_{l'})}$, $1 \le l, l' \le k$, $\mathbf{R}(\omega) = ((\rho_{jj'}(\omega)))_{jj'=1}^N$, and $\mathbf{R}^0(\omega) = ((\rho^0_{ll'}(\omega)))_{ll'=1}^k$. Let

$$\mathbf{F}(\omega) = diag(f_{11}(\omega), \dots, f_{N1}(\omega), \dots, f_{1k}(\omega), \dots, f_{Nk}(\omega)),$$

and define $A(\omega)$ to be the $Nk \times Nk$ matrix with (l,l')th block $A_{ll'}(\omega)$, for $1 \le l, l' \le k$. Then $A(\omega) = \mathbf{F}(\omega)[(e(\omega) \odot \mathbf{R}^0(\omega)) \otimes \mathbf{R}(\omega)]\mathbf{F}(\omega)$, where $\odot$ denotes Schur (or Hadamard) product, i.e., coordinate-wise product of two matrices of same dimension, and $\otimes$ denotes the Kronecker product.

Note that, for an arbitrary $a \in \mathbb{C}^k$, $a^*(e(\omega) \odot \mathbf{R}^0(\omega))a = b^*\mathbf{R}^0(\omega)b$, where $b_l = a_l e^{-i\omega^T s_l}$, $l = 1, \dots, k$. Therefore, if $\mathbf{R}^0(\omega)$ is positive definite, then so is the $k \times k$ matrix $e(\omega) \odot \mathbf{R}^0(\omega)$. Since $\mathbf{F}(\omega)$ is diagonal with non-negative diagonal entries, from (2.3), wherever $\mathbf{F}(\omega)$ is p.d., $A(\omega)$ is p.d. (n.n.d.) if both $\mathbf{R}(\omega)$ and $\mathbf{R}^0(\omega)$ are p.d. (at least one n.n.d. but not p.d.). From (2.1),

$$\mathbf{C} = \int_{\mathbb{R}^d} A(\omega)d\omega, \tag{2.4}$$

where the integral is taken over every element of the matrix $A(\omega)$. By the Cauchy-Schwarz inequality and the fact that $\max\{|\rho_{jj'}(\omega)|, |\rho^0_{ll'}(\omega)|\} \le 1$, a sufficient condition for the integral in (2.4) to be finite is that $\max_{1\le j\le N} \max_{1\le l\le k} \int (f_{jl}(\omega))^2 d\omega < \infty$.

Lemma 1. *Sufficient conditions for* $\mathbf{C}$ *to be positive definite are that* (i) *the* $Nk \times Nk$ *matrix* $A(\omega)$ *is non-negative definite on* $\mathbb{R}^d$*, and positive definite on a set of positive Lebesgue measure in* $\mathbb{R}^d$*; and* (ii) $\int_{\mathbb{R}^d}(f_{jl}(\omega))^2 d\omega < \infty$ *for all* $j = 1, \dots, N$ *and* $l = 1, \dots, k$.

Lemma 2. *Suppose there exists $B \subset \mathbb{R}^d$ with positive Lebesgue measure such that, for all $\omega \in B$, we have $f_{jl}(\omega) > 0$ for each $j = 1, \ldots, N$, $l = 1, \ldots, k$, and that both $\mathbf{R}(\omega)$ and $\mathbf{R}^0(\omega) := ((\rho^0_{ll'}(\omega)))^k_{ll'=1}$ are positive definite matrices. Then $A(\omega)$ is a positive definite matrix on B.*

As an immediate consequence of Lemmas 1 and 2 we have the following.

Theorem 1. *Suppose that C_{jl}, $1 \le j \le N$, $1 \le l \le k$, are positive definite functions, and $\mathbf{R}(\omega) = ((\rho_{jj'}(\omega)))^N_{j,j'=1}$, and $\mathbf{R}^0(\omega) := ((\rho^0_{ll'}(\omega)))^k_{ll'=1}$ are non-negative definite matrices for all $\omega \in \mathbb{R}^d$. If there exists a set $B \subset \mathbb{R}^d$ with nonzero Lebesgue measure such that, for all $\omega \in B$, we have $f_{jl}(\omega) > 0$, $\int_{\mathbb{R}^d}(f_{jl}(\omega))^2 d\omega < \infty$ for each j and l, and both $\mathbf{R}(\omega)$ and $\mathbf{R}^0(\omega)$ are positive definite on B, then the matrix $\mathbf{C}$ at (2.1) is a valid cross-covariance structure of an N-dimensional stochastic process on $D = \{s_1, \ldots, s_k\}$.*

In the above construction, since the C_{jl}'s, $\rho_{jj'}$'s, and $\rho^0_{ll'}$ are arbitrary, a rich framework for modeling spatial processes is achieved if we can generalize this from any arbitrary finite set $\{s_l; l = 1, \ldots, k\}$ to an arbitrary spatial region $D \in \mathbb{R}^d$. This does hold in the stationary case (i.e., when $f_{jl}(\omega) = \widetilde{f}_j(\omega)$ for all $l = 1, \ldots, k$, for all j, and $\rho^0_{ll'}(\omega) \equiv 1$) if the matrix $\mathbf{R}(\omega) = ((\rho_{jj'}(\omega)))^N_{j,j'=1}$ is non-negative definite for all $\omega \in \mathbb{R}^d$.

Corollary 1. *Suppose $C_1, \ldots, C_N$ are valid covariance functions on $\mathbb{R}^d$ with spectral densities $\widetilde{f}_1, \ldots, \widetilde{f}_N$, respectively, and the functions $\rho_{jj'}$ are such that $\mathbf{R}(\omega) := ((\rho_{jj'}(\omega)))^N_{jj'=1}$ is non-negative definite a.e. $\omega \in \mathbb{R}^d$. Then there is a mean-zero Gaussian stationary stochastic process $\mathbf{Y}(s) = (Y_1(s), \ldots, Y_N(s))$ on $\mathbf{R}^d$ such that*

$$\mathrm{Cov}\,(Y_j(s), Y_{j'}(t)) = C^\star_{jj'}(s-t) := \int_{\mathbb{R}^d} e^{\omega^T(s-t)} \widetilde{f}_j(\omega)\widetilde{f}_{j'}(\omega)\rho_{jj'}(\omega)d\omega. \qquad (2.5)$$

2.2. Construction of nonstationary covariances on $\mathbb{R}^d$

We now generalize the construction of the nonstationary $N \times N$ covariance function $C^\star$ from the set $\{s_1, \ldots, s_k\}$ to the entire space $\mathbb{R}^d$. Since a Gaussian process is determined entirely by its mean and covariance, given points $s_1, \ldots, s_k \in \mathbb{R}^d$, we can find a zero mean Gaussian random vector $(Y_{jl} : 1 \le j \le N,\ 1 \le l \le k)$ with covariance matrix given by $C^\star$. Moreover, this vector can be viewed as the realization of an N-dimensional random process $\mathbf{Y}(s) = (Y_1(s), \ldots, Y_N(s))$ at the points $s_1, \ldots, s_k$, if we define $Y_{jl} = Y_j(s_l)$. The next theorem states that an extension of the process $\mathbf{Y}(s)$ to arbitrary domains in $\mathbb{R}^d$ is possible.

Theorem 2. *Let $\{f_j(s,\omega)\}_{j=1}^N$, be non-negative functions on $\mathbb{R}^d \times \mathbb{R}^d$, such that $\sup_{s\in\mathbb{R}^d} \int_{\mathbb{R}^d} (f_j(s,\omega))^2 d\omega < \infty$. Let $\rho^0(s,s',\omega)$ be a valid correlation function on $\mathbb{R}^d \times \mathbb{R}^d$ for a.e. $\omega \in \mathbb{R}^d$. Also, let $\mathbf{R}(\omega) = ((\rho_{jj'}(\omega)))_{jj'=1}^N$ be non-negative definite for every $\omega \in \mathbb{R}^d$. If there exist a set $B \in \mathbb{R}^d$ with positive Lebesgue measure so that for every $\omega \in B$, the function $f_j(\cdot,\omega) > 0$, the matrix $\mathbf{R}(\omega)$ is positive definite, and the correlation function $\rho^0(\cdot,\cdot,\omega)$ is positive definite, then there exists an N-dimensional Gaussian spatial process $\mathbf{Y}(s)$ on $\mathbb{R}^d$ with $N \times N$-dimensional covariance kernel $C^\star(s,s')$ whose entries are given by,*

$$C^\star_{jj'}(s,s') = \int_{\mathbb{R}^d} e^{i\omega^T(s-s')} f_j(s,\omega) f_{j'}(s',\omega) \rho^0(s,s',\omega) \rho_{jj'}(\omega) d\omega, \;\; s,s' \in \mathbb{R}^d. \quad (2.6)$$

The function $f_j(s,\omega)$ can be interpreted as a location-dependent spectral density of a *locally stationary* stochastic process. If $f_j(s,\omega) = f_j(\omega)$ for all $j = 1,\ldots,N$, and $\rho^0(s,s',\omega) = 1$, then $C^\star$ as in Theorem 2 becomes a covariance function of an N-dimensional stationary process on $\mathbb{R}^d$.

2.3. Sufficient conditions for positive definiteness

In this subsection, we present a sufficient condition on the Fourier transforms of the cross-correlation functions, namely $\{\rho_{jj'}\}_{j\neq j'}$, that guarantee the positive-definiteness of the covariance function in the convolution model presented in Section 2.2 when the number of variables N is at most 4.

Theorem 3. *When $N \leq 4$, sufficient conditions for positive definiteness of $\mathbf{R}(\omega)$ are the following.*

(i) $1 > |\rho_{jj'}|^2$ *for all* $1 \leq j < j' \leq N$.

(ii) $1 > |\rho_{jl}|^2 + |\rho_{lm}|^2 + |\rho_{mj}|^2 - 2\mathrm{Re}(\rho_{jl}\rho_{lm}\rho_{mj})$, *for all* $1 \leq j < l < m \leq N$.

(iii) *If* $1 \leq j \neq l \neq m \neq n \leq N$, *then*

$$\begin{aligned} &1 - |\rho_{lm}|^2 - |\rho_{mn}|^2 - |\rho_{nl}|^2 + 2\mathrm{Re}(\rho_{lm}\rho_{mn}\rho_{nl}) \\ &\quad > |\rho_{jl}|^2 + |\rho_{jm}|^2 + |\rho_{jn}|^2 - (|\rho_{jl}|^2|\rho_{mn}|^2 + |\rho_{jm}|^2|\rho_{ln}|^2 + |\rho_{jn}|^2|\rho_{lm}|^2) \\ &\qquad + 2\mathrm{Re}(\rho_{jl}\rho_{lm}\rho_{mj}) + 2\mathrm{Re}(\rho_{jm}\rho_{mn}\rho_{nj}) + 2\mathrm{Re}(\rho_{jn}\rho_{nl}\rho_{lj}) \\ &\qquad - 2\mathrm{Re}(\rho_{jl}\rho_{ln}\rho_{nm}\rho_{mj}) - 2\mathrm{Re}(\rho_{jm}\rho_{ml}\rho_{ln}\rho_{nj}) - 2\mathrm{Re}(\rho_{jn}\rho_{nm}\rho_{ml}\rho_{lj}). \end{aligned}$$

Equality in place of any of the inequalities implies singularity of the matrix $\mathbf{R}(\omega)$.

2.4. A general model

A general formulation for the nonstationary covariance kernels comes from introducing some structure to the correlation function $\rho^0(s,t,\omega)$. One proposal

is to consider

$$\rho^0(s,t,\omega) = \sum_{l=1}^{\infty} \widetilde{\rho}_l(s,t)\psi_l(\omega), \tag{2.7}$$

where $\widetilde{\rho}_l$ are correlation functions on $\mathbb{R}^d \times \mathbb{R}^d$, and the non-negative ψ_l are such that $\sum_{l=1}^{\infty} \psi_l(\omega) \leq 1$ a.e.

Recall that by the *spectral representation theory* of stationary stochastic processes (Yaglom (1962) and Schabenberger and Gotway (2005)), there exists an N-dimensional stochastic process $\mathbf{Z}(\omega)$ with independent coordinates defined on $\mathbb{R}^d$, such that the one-dimensional stationary process $X_j(s)$ with covariance function given by $C_{jj}(s-t) = \int_{\mathbb{R}^d} e^{i\omega^T(s-t)}(f_j(\omega))^2 d\omega$, with $\int (f_j(\omega))^2 d\omega < \infty$, can be represented as $X_j(s) = \int_{\mathbb{R}^d} e^{i\omega^T s} f_j(\omega) d\mathbf{Z}_j(\omega)$. Pintore and Holmes (2006) consider processes of the form

$$\widetilde{X}_j(s) = \int_{\mathbb{R}^d} e^{i\omega^T s} f_j(s,\omega) d\mathbf{Z}_j(\omega),$$

where $f_j(s,\omega)$ are of the form $h_j(s)\widetilde{f}_j(\omega;\theta(s))$, where $\widetilde{f}_j(\cdot;\theta)$ is the spectral density function of a stationary stochastic process with parameter θ, and $h_j(\cdot)$ is a non-negative function, for each $j = 1,\ldots,N$. These processes have covariance functions $C_{jj'}(s,t) = \delta_{j-j'} \int_{\mathbb{R}^d} e^{i\omega^T(s-t)} f_j(s,\omega) f_{j'}(t,\omega) d\omega$, where $\delta_0 = 1$ and $\delta_k = 0$ if $k \neq 0$. Our proposal can therefore be viewed as extending their method to the multidimensional case while introducing spatially varying cross-correlation functions.

The setting described by (2.7) can be realized by describing the process $\mathbf{Y}(s)$ as

$$\mathbf{Y}(s) = \sum_{l=1}^{\infty} \xi_l(s) \int_{\mathbb{R}^d} e^{i\omega^T s} \mathbf{F}(s,\omega) \cdot \mathbf{R}^{1/2}(\omega) \sqrt{\psi_l(\omega)} d\mathbf{Z}(\omega), \tag{2.8}$$

where $\mathbf{R}^{1/2}(\omega)$ is a non-negative square-root of the matrix $\mathbf{R}(\omega)$, and $\mathbf{F}(s,\omega)$ is a diagonal matrix with jth diagonal element $f_j(s,\omega)$. Here $\{\xi_l(s)\}_{l=1}^{\infty}$ are uncorrelated (in the Gaussian case, independent) stochastic processes, independent of the process $\mathbf{Z}(\omega)$, with $\mathrm{Cov}\,(\xi_l(s),\xi_l(t)) = \widetilde{\rho}_l(s,t)$. Observe that we have the formal expansion $U(s,\omega) = \sum_{l=1}^{\infty} \xi_l(s)\sqrt{\psi_l(\omega)}$, defining a mean zero, L^2 stochastic process on $\mathbb{R}^d \times \mathbb{R}^d$ with covariance function $\rho_U(s,t,\omega,\omega') = \mathrm{Cov}\,(U(s,\omega),U(t,\omega'))$. Also then, $\rho^0(s,t,\omega) = \rho_U(s,t,\omega,\omega)$. Then we can formally define

$$\mathbf{Y}(s) = \int_{\mathbb{R}^d} e^{i\omega^T s} U(s,\omega) \mathbf{F}(s,\omega) \cdot \mathbf{R}^{1/2}(\omega) d\mathbf{Z}(\omega), \tag{2.9}$$

where the processes $U(s,\omega)$ are $\mathbf{Z}(\omega)$ are assumed to be independent and defined on the same probability space. Note that (2.9) is a formal integral representation,

and we are assuming that all the measurability conditions needed on the processes to define the stochastic integral are satisfied. The most manageable case from a practical point of view though, in our opinion, is when $\widetilde{\rho}_l(s,t) = \widetilde{\rho}(s-t;\theta_l)$ for some parametric correlation function $\widetilde{\rho}(\cdot;\theta)$.

3. Specification of the Nonstationary Covariance Model

In this section we give a detailed description of a model that has a natural appeal from the perspective of modeling spatially inhomogeneous multivariate processes, and which renders the problem of estimating the nonstationary covariance kernel computationally quite tractable.

We assume that $\rho^0(s,s',\omega) = \rho_1(s-s')\rho_2(\omega) = \rho_1(s-s')$. Here $\rho_2(\omega) \equiv 1$ for simplicity, and more generally one can assume some parametric form for $\rho_2(\omega)$. We assume a parametric form for the functions $\rho_1(s-s')$ and $\mathbf{R}(\omega)$. Then we model

$$f_j(s,\omega) = \sum_{l=1}^{L} |\Sigma_l|^{-1/2} K_l(\Sigma_l^{-1/2}(s-t_l)) f_j(\omega;\theta_{jl}), \tag{3.1}$$

where $\{t_l : l = 1,\dots,L\}$ is a sequence of points in $\mathbb{R}$; for each l, $K_l(\cdot)$ is a nonnegative kernel with $\int K_l(x)dx = 1$; $\{\Sigma_l : l = 1,\dots,L\}$ is a sequence of $d \times d$ positive definite matrices; and for every fixed $\theta_{jl} \in \Theta_j$, $f_j(\cdot;\theta_{jl})$ is a spectral density function belonging to a parametric family parameterized by θ_{jl}. Also, we assume that $\rho_{jj'}(\omega) = \rho_0(\omega;\nu_{jj'},\kappa)$, for parameters $\{\nu_{jj'}\}_{j,j'=1}^{N}$ and κ; and $\rho_1(s-t) \equiv \rho_1(s-t;\tau)$ for some parameter τ. Under this setting, $C^\star(s,t)$, the covariance kernel of $\mathbf{Y}(t)$, is determined through

$$C^\star_{jj'}(s,t) = \rho_1(s-t) \sum_{l,l'=1}^{L} |\Sigma_l|^{-1/2} K_l(\Sigma_l^{-1/2}(s-t_l)) |\Sigma_{l'}|^{-1/2} K_{l'}(\Sigma_{l'}^{-1/2}(t-t_{l'}))$$
$$\cdot \int_{\mathbb{R}^d} e^{i\omega^T(s-t)} f_j(\omega;\theta_{jl}) f_{j'}(\omega;\theta_{j'l'}) \rho_{jj'}(\omega) d\omega, \quad 1 \le j,j' \le N. \tag{3.2}$$

Thus, defining $G_{jj'}(s;\theta_{jl},\theta_{j'l'},\nu_{jj'},\kappa) = \int_{\mathbb{R}^d} e^{i\omega^T s} f_j(\omega;\theta_{jl}) f_{j'}(\omega;\theta_{j'l'}) \rho_0(\omega;\nu_{jj'},\kappa)\, d\omega$,

$$C^\star_{jj'}(s,t) = \rho_1(s-t;\tau) \sum_{l,l'=1}^{L} |\Sigma_l|^{-1/2} K_l(\Sigma_l^{-1/2}(s-t_l)) |\Sigma_{l'}|^{-1/2} K_{l'}(\Sigma_{l'}^{-1/2}(t-t_{l'}))$$
$$\cdot G_{jj'}(s-t;\theta_{jl},\theta_{j'l'},\nu_{jj'},\kappa). \tag{3.3}$$

Typically, the sequence $\{t_l\}_{l=1}^{L}$ may be assumed given.

3.1. Specification of the parametric spectral density and correlation

We now give a complete description of a model that maintains a balance between flexibility, and computational cost and interpretability. We choose $\rho_1(\cdot;\tau)$ to be an arbitrary parametric stationary correlation function on $\mathbb{R}^d$ with parameter τ. We assume that $f_j(\omega;\theta_{jl})$ is of the form $c_{jl}\gamma(\omega;\widetilde{\theta}_{jl})$ for some scale parameter $c_{jl} > 0$ (note that we express $\theta_{jl} = (c_{jl}, \widetilde{\theta}_{jl})$), and a parametric class of spectral densities $\gamma(\cdot;\widetilde{\theta})$ that is closed under product. The latter means that given any $m \geq 1$, there exists a function $\widetilde{\gamma}(\cdot;\cdot)$ and functions $\mathbf{c}_\gamma(\cdots)$, $\mathbf{d}_\gamma(\cdots)$ of m variables such that, given parameters $\widetilde{\theta}_1, \ldots, \widetilde{\theta}_m$,

$$\prod_{i=1}^{m} \gamma(\cdot;\widetilde{\theta}_i) = \mathbf{d}_\gamma(\widetilde{\theta}_1, \cdots, \widetilde{\theta}_m)\widetilde{\gamma}(\cdot;\mathbf{c}_\gamma(\widetilde{\theta}_1, \cdots, \widetilde{\theta}_m)).$$

In particular, $\gamma(\cdot;\widetilde{\theta}_1) = \mathbf{d}_\gamma(\widetilde{\theta}_1)\widetilde{\gamma}(\cdot;\mathbf{c}_\gamma(\widetilde{\theta}_1))$. For example, the spectral densities of the Matérn family (under some restrictions on the parameters), and the Gaussian family satisfy this property.

For $j \neq j'$, we express $\rho_0(\omega;\nu_{jj'},\kappa)$ as $\nu_{jj'}\alpha(\omega;\kappa)$, where $\alpha(\omega;\kappa) \equiv \alpha(\omega)$ is a real-valued function satisfying $-1/(N-1) \leq \alpha(\omega) \leq 1$. We choose $\{\nu_{jj'}\}_{1\leq j\neq j'\leq N}$ in such a way that the $N \times N$ matrix $\mathbf{N} = ((\nu_{jj'}))_{1\leq j,j'\leq N}$, with $\nu_{jj} \equiv 1$ for all j, is positive definite (in fact, a correlation matrix). Since the $N \times N$ matrix $\mathbf{A}(\omega)$ with diagonal elements 1 and off-diagonal elements $\alpha(\omega)$ is clearly positive definite (under the restriction $\alpha(\omega) \in (-1/(N-1), 1]$), the matrix $\mathbf{R}(\omega)$ thus specified is positive semidefinite for all ω, since the latter is just $\mathbf{N} \odot \mathbf{A}(\omega)$. Taking $\alpha(\omega) = 1$ for all ω corresponds to the situation where the different coordinate processes have the same correlation structure at all frequencies. To add flexibility to the model without making it computationally too cumbersome, we propose the use of

$$\alpha(\omega) = \frac{\gamma(\omega;\alpha_1)}{\gamma(0;\alpha_1)} - \beta\frac{\gamma(\omega;\alpha_2)}{\gamma(0;\alpha_2)}, \tag{3.4}$$

where $\beta \in [0, 1/(N-1))$ and α_1, α_2 are free parameters, and γ belongs to the same family of spectral densities as the one used in specifying f_j's. Thus $\kappa = (\alpha_1, \alpha_2, \beta)$.

An obvious advantage of this restriction is that one has a closed form expression for $G_{jj'}(s;\theta_{jl},\theta_{j'l'},\nu_{jj'},\kappa)$ in terms of the inverse Fourier transform of the function $\widetilde{\gamma}$: for $1 \leq j \neq j' \leq N$,

$$G_{jj'}(s;\theta_{jl},\theta_{j'l'},\nu_{jj'},\kappa) = c_{jl}c_{j'l'}\nu_{jj'} \cdot \left[\frac{\mathbf{d}_\gamma(\widetilde{\theta}_{jl},\widetilde{\theta}_{j'l'},\alpha_1)}{\gamma(0;\alpha_1)}(\mathcal{F}^{-1}\widetilde{\gamma})(s;\mathbf{c}_\gamma(\widetilde{\theta}_{jl},\widetilde{\theta}_{j'l'},\alpha_1)) - \frac{\beta\mathbf{d}_\gamma(\widetilde{\theta}_{jl},\widetilde{\theta}_{j'l'},\alpha_2)}{\gamma(0;\alpha_2)}(\mathcal{F}^{-1}\widetilde{\gamma})(s;\mathbf{c}_\gamma(\widetilde{\theta}_{jl},\widetilde{\theta}_{j'l'},\alpha_2))\right], \tag{3.5}$$

where $\mathcal{F}^{-1}\widetilde{\gamma}$ denotes the inverse Fourier transform of $\widetilde{\gamma}$, i.e., the covariance function whose spectral density is $\widetilde{\gamma}$. Also, for $j = j'$,

$$G_{jj'}(s;\theta_{jl},\theta_{j'l'},\nu_{jj'},\kappa) = c_{jl}c_{jl'}\mathbf{d}_\gamma(\widetilde{\theta}_{jl},\widetilde{\theta}_{jl'})(\mathcal{F}^{-1}\widetilde{\gamma})(s;\mathbf{c}_\gamma(\widetilde{\theta}_{jl},\widetilde{\theta}_{jl'})). \quad (3.6)$$

3.2. A bivariate process

In many practical problems we are interested in studying the joint behavior of two processes at a time, e.g., soil salinity and soil moisture content, temperature and pressure fields, etc. In the bivariate case, the general formulation for our model simplifies considerably. As stated in Theorem 3, in order that (2.6) be a valid covariance kernel, it is sufficient that $|\rho_{12}(\omega)|^2 \leq 1$. Also, the condition $Im(\rho_{12}(\omega)) \neq 0$ for $\omega \in B$ for some B with positive Lebesgue measure, is necessary to ensure that $\mathrm{Cov}\,(Y_1(s), Y_2(s')) \neq \mathrm{Cov}\,(Y_2(s), Y_1(s'))$ (asymmetric cross-covariance). Our model bridges the two extremes: $\rho_{12}(\omega) \equiv 0$ yields zero cross-correlation across all spatial locations, and $\rho_{12}(\omega) \equiv 1$ specifies the singular cross-convolution model outlined in Majumdar and Gelfand (2007).

For the purpose of illustrating some of the main features of our model, we focus on a bivariate non-stationary Gaussian process with the parametric covariance model presented in Section 3.1. We assume $\gamma(\cdot;\widetilde{\theta})$ to be a Gaussian spectral density with scale parameter $\widetilde{\theta}$, so that $\gamma(\omega;\widetilde{\theta}) = (1/2)(\pi\widetilde{\theta})^{-1/2}e^{-\omega^2/4\widetilde{\theta}}$. Hence, the parameters describing the product of m such densities with parameters $\widetilde{\theta}_1,\ldots,\widetilde{\theta}_m$ are

$$\mathbf{c}_\gamma(\widetilde{\theta}_1,\cdots,\widetilde{\theta}_m) = \frac{1}{\sum_{i=1}^m 1/\widetilde{\theta}_i} \quad \text{and} \quad \mathbf{d}_\gamma(\widetilde{\theta}_1,\cdots,\widetilde{\theta}_m) = \frac{1}{2^m\pi^{m/2}}\frac{1}{\prod_{i=1}^m \widetilde{\theta}_i^{1/2}}$$

with $\widetilde{\gamma}(\omega;\widetilde{\theta}) = e^{-\omega^2/4\widetilde{\theta}}$. Moreover, in this case, we have $(\mathcal{F}^{-1}\gamma)(s;\widetilde{\theta}) = e^{-\widetilde{\theta}\|s\|^2}$. Thus, the expression for $G_{jj'}(s;\theta_{jl},\theta_{j'l'},\nu_{jj'},\kappa)$ can be simplified by noting that $\gamma(0;\widetilde{\theta}) = \mathbf{d}_\gamma(\widetilde{\theta}) = (\pi\widetilde{\theta})^{-1/2}/2$, and $(\mathcal{F}^{-1}\widetilde{\gamma})(s;\widetilde{\theta}) = 2(\pi\widetilde{\theta})^{1/2}e^{-\widetilde{\theta}\|s\|^2}$. Next, since the Cholesky decomposition of a positive definite matrix can be chosen to be lower triangular, we write

$$\Sigma_l{}^{-\frac{1}{2}} = \begin{pmatrix} \sigma_{11l} & 0 \\ \sigma_{21l} & \sigma_{22l} \end{pmatrix}. \quad (3.7)$$

Using (3.5) and (3.6) we obtain a simplified form of the covariance function $C^\star_{jj'}(s,t)$. For ease of expressions, we consider the case when $\widetilde{\theta}_{jl} = \widetilde{\theta}_l$ for $j = 1,2$. Then,

$$C^\star_{jj'}(s,t) = e^{-\tau(s-t)}\frac{1}{2\pi}\sum_{l,l'=1}^{L}\sigma_{11l}\sigma_{22l}\sigma_{11l'}\sigma_{22l'}\exp(-\frac{1}{2}\parallel\Sigma_l^{-1/2}(s-t_l)\parallel^2$$
$$-\frac{1}{2}\parallel\Sigma_{l'}^{-1/2}(t-t_{l'})\parallel^2)\cdot c_{jl}c_{j'l'}\nu_{jj'}\Gamma_{jj'}(s-t;\widetilde{\theta}_l,\widetilde{\theta}_{l'},\kappa), \quad (3.8)$$

where $\kappa = (\alpha_1, \alpha_2, \beta)$, $\Sigma_l^{-1/2}$ has the form (3.7), and the $\Gamma_{jj'}(s; \widetilde{\theta}_l, \widetilde{\theta}_{l'}, \kappa)$ are given in the *supplementary material*.

3.3. Comparison with other nonstationary models

Here we compare our model with other well-known models for nonstationary spatial covariances. For brevity, we focus on the univariate process specified by (1.1). Assuming that the set D is finite, say $D = \{x_1, \ldots, x_M\}$, the covariance kernel for the process $Y(\cdot)$ is

$$C_Y(s,t) = \sum_{m=1}^{M} K(s - x_m) K(t - x_m) C_{\theta(x_m)}(s - t), \tag{3.9}$$

where $C_\theta(\cdot)$ is the stationary covariance kernel of the process $Z_\theta(\cdot)$. The expression for C_Y bears a similarity with the expression (3.3) when $j = j'$. Indeed, in the latter case, $C^\star_{jj'}$ reduces to

$$C^\star_{jj}(s,t) = \sum_{l,l'=1}^{L} \widetilde{K}_l(s - t_l) \widetilde{K}_{l'}(t - t_{l'}) \widetilde{\rho}_1(s - t; \tau) G_{jj}(s - t; \theta_{jl}, \theta_{jl'}) \tag{3.10}$$

for appropriate kernels $\widetilde{K}_l(\cdot)$. Observe that if the kernels $\{\widetilde{K}_l\}_{l=1}^L$ and the centers $\{t_l\}_{l=1}^L$ are such that, for all $l \neq l'$, $\widetilde{K}_l(\cdot - t_l)\widetilde{K}_l(\cdot - t_{l'}) = 0$, and if $U(s,\omega) \equiv 1$ (so that $\widetilde{\rho}_1(\cdot) \equiv 1$), then (3.10) can be expressed in the form (3.9).

Adopting a different viewpoint, we consider the representation of the process $Y(\cdot)$ described by (1.1), and the process $Y_j(\cdot)$ with covariance kernel described by (3.10) in the spectral domain. The former is

$$Y(s) = \int_{\mathbb{R}^d} e^{i\omega^T s} \sum_{m=1}^{M} K(s - x_m) \widetilde{f}(\omega; \theta(x_m)) d\widetilde{Z}_m(\omega), \tag{3.11}$$

where $\widetilde{Z}_m(\cdot)$, $m = 1, \ldots, M$, are i.i.d. zero mean Brownian processes, and $\widetilde{f}(\cdot; \theta)$ is the spectral density function of $C_\theta(\cdot)$. Whereas, from (2.9) and (3.1), $Y_j(\cdot)$ can be written

$$Y_j(s) = \int_{\mathbb{R}^d} e^{i\omega^T s} U(s,\omega) \sum_{l=1}^{L} \widetilde{K}_l(s - t_l) f_j(\omega; \theta_{jl}) d\mathbf{Z}_j(\omega), \tag{3.12}$$

where $\mathbf{Z}_j(\cdot)$ is the jth coordinate of $\mathbf{Z}(\cdot)$. The two processes thus differ mainly by the fact that (3.11) is a representation of $Y(\cdot)$ in terms of weighted sum of *independent* spectral processes $\{\widetilde{Z}_m(\cdot)\}$ and $Y_j(\cdot)$ is represented in terms of one spectral process $\mathbf{Z}_j(\cdot)$.

Table 1. Parameter specification of the 8 different models.

	$\sigma_{11l}=1$			$\theta_{jl}=(c_{jl},\widetilde{\theta}_{jl})$	$\alpha_2=0.2$
Model	σ_{21l}	σ_{22l}	c_{jl}	$\widetilde{\theta}_{jl}$	α_1
(1)	3.0	3	$2\sqrt{j/l}$	$\sqrt{1/l}$	0.1
(2)	8.0	3	$2\sqrt{j/l}$	$\sqrt{1/l}$	0.1
(3)	3.0	3	$10jl$	$\sqrt{1/l}$	0.1
(4)	8.0	3	$10jl$	$\sqrt{1/l}$	0.1
(5)	3.0	5	$2\sqrt{j/l}$	$\sqrt{1/l}$	0.1
(6)	8.0	5	$2\sqrt{j/l}$	$\sqrt{1/l}$	0.1
(7)	3.0	3	$2\sqrt{j/l}$	$\sqrt{1/l}$	0.4
(8)	8.0	3	$2\sqrt{j/l}$	$j+l$	0.1

These comparisons also indicate that, as long as the kernels $\{\widetilde{K}_l\}_{l=1}^L$ and the centers $\{t_l\}_{l=1}^L$ are such that for all $l \neq l'$ the products $\widetilde{K}_l(\cdot - t_l)\widetilde{K}_l(\cdot - t_{l'})$ have comparatively small values, the parameters in the model are identifiable. They can be identified essentially from the data on different spatial locations. Indeed this is the case if the centers $\{t_l\}_{l=1}^L$ are well-separated, and the scale parameters Σ_l for the kernel $\widetilde{K}_l$ are comparatively small in magnitude. In practice, we expect to have reasonable *apriori* information about the possible spatial inhomogeneity, so that the specification of fairly accurate prior for the kernel centers $\{t_l\}_{l=1}^L$ is possible.

4. Simulation Results

To understand the dependency of the model on various parameters, we performed a small simulation study for the bivariate ($N = 2$) case, in which we specified $L = 4$; $\sigma_{11l} = 1$, for all $l = 1, \ldots, 4$; $\beta = 0.5$, $\nu_{12} = \nu_{21} = 0.5$, $\alpha_2 = 0.2$, $\tau = 0.5$. We generated 100 realizations (on the unit square $[0,1]\times[0,1]$) of a bivariate spatial process with centers of the four kernels $t_1 = (0.1, 0.7)$, $t_2 = (0.6, 0.1)$, $t_3 = (0.9, 0.6)$, $t_4 = (0.6, 0.9)$. Changing the values of σ_{21l}, σ_{22l}, c_{jl}, $\widetilde{\theta}_{jl}$ and α_1, we generated data from eight different models as given in Table 1. If we generalize this to $N(\geq 2)$ processes and L kernels, then the number of parameters is $[N(N+1)/2 + 3N]L + N(N-1)/2 + 4$. Note that the first term within bracket $N(N+1)/2$ corresponds to Σ_l; the second term within brackets corresponds to $\{c_{jl}\}_{j=1}^N$, $\{\widetilde{\theta}_{jl}\}_{j=1}^N$ and t_l; and the term $N(N-1)/2$ corresponds to $\{\nu_{jj'}\}_{1\leq j<j'\leq N}$.

Qualitative features of the nonstationarity are illustrated through the contour plot of $Var(Y(s))$ against s (Figure 1). A sample realization of each process is plotted in Figure 4 (in the *supplementary material*). From the figures, we

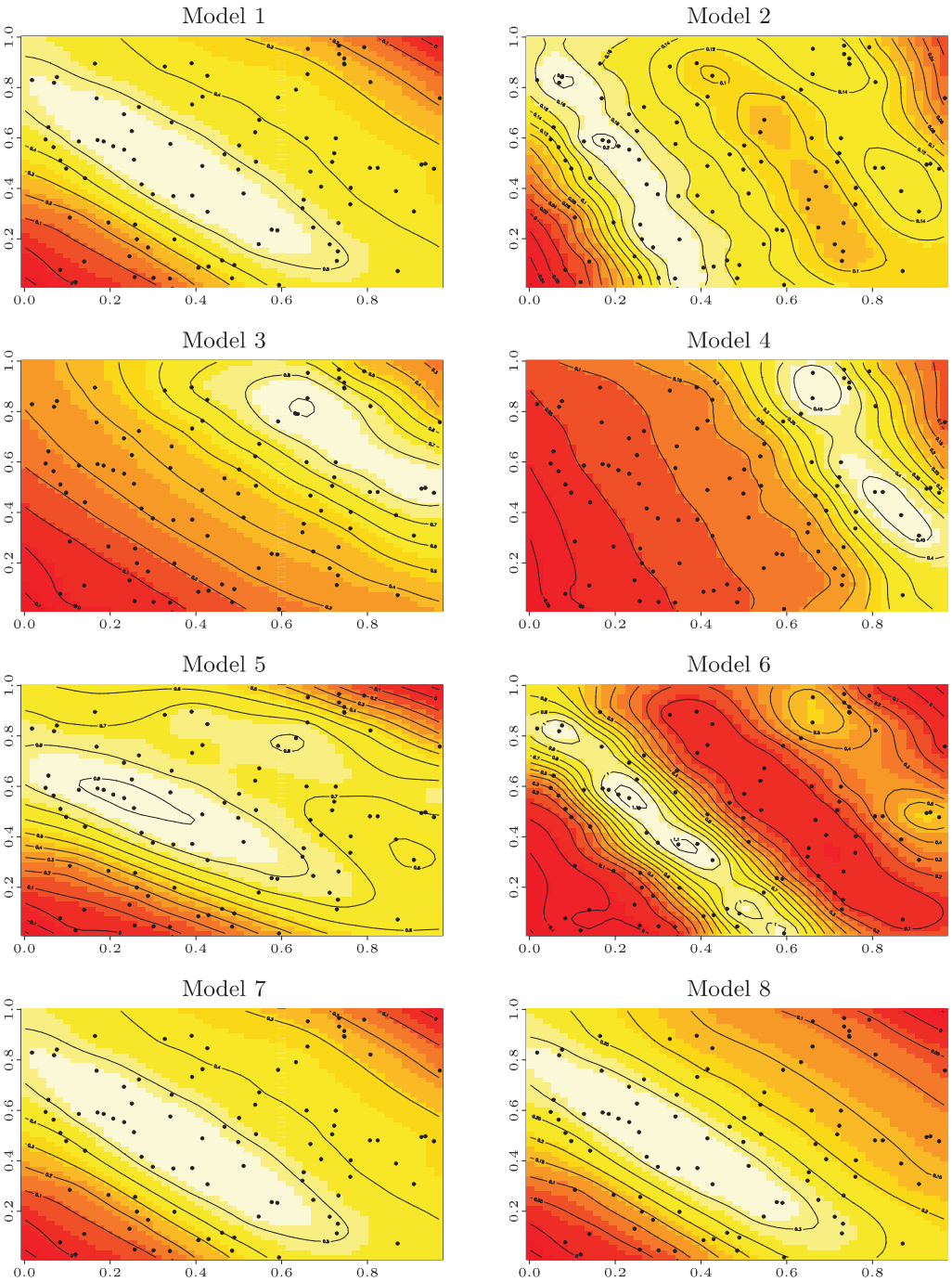

Figure 1. $\mathrm{Var}(Y_1(s))$ vs. s under the 8 different models.

clearly note distinct pattern variations among the variance profiles as well as sample realizations of the eight processes. All parameters seem to have considerable effects on the processes, and with the flexibility of these local and global parameters we can generate a wide class of non-stationary multivariate spatial models.

5. Bayesian Modeling and Inference

We give an outline of a Bayesian approach to estimating the parameters of the general N-dimensional model specified in Section 3.1. We assume an *exponential correlation* structure (Stein (1999)) for $\rho_1(\cdot, \tau)$, and a Gaussian spectral density for $\gamma(\cdot; \theta)$, where $\tau > 0$ is a global decay parameter. We assign a $Gamma(a_\tau, b_\tau)$ prior for τ, with $a_\tau, b_\tau > 0$. We take $c_{jl} \overset{i.i.d.}{\sim} Gamma(a_{c_{jl}}, b_{c_{jl}})$, and $\widetilde{\theta}_{jl} \overset{i.i.d.}{\sim} Gamma(a_{\widetilde{\theta}_{jl}}, b_{\widetilde{\theta}_{jl}})$. We specify i.i.d. $InvWishart(\Psi, 2)$ priors for Σ_l, with mean matrix Ψ. In order to specify a prior for the parameters $\{\nu_{jj'}\}$, we consider an $N \times N$ positive definite matrix $\nu^\star$, and put $\mathbf{N} := ((\nu_{jj'}))_{j,j'=1}^N = diag(\nu^\star)^{-1/2} \nu^\star diag(\nu^\star)^{-1/2}$. We take $\nu^\star \sim InvWishart(\tilde{\nu}, d)$, where $\tilde{\nu}$ is an $N \times N$ positive definite correlation matrix (to avoid over-parametrization) whose structure represents our prior belief about the strength and directionality of association of the different coordinate processes. In a geophysical context this may mean knowledge about the states of the physical process. Note that this ensures the positive-definiteness of the matrix $\mathbf{N}$, and additionally sets the diagonals ν_{jj} to 1, as required. When $N = 2$, there is only one unknown parameter in $\mathbf{N}$ and we specify the prior $\nu_{12} \sim Unif(-1, 1)$ which guarantees that $\mathbf{N}$ is p.d. Since the permissible range of β is $[0, 1/(N-1)]$, we assume the prior of β to be $\beta \sim Unif(0, 1/(N-1))$. We assume independent Gamma priors $Gamma(a_{\alpha_k}, b_{\alpha_k})$, $k = 1, 2$, for the positive parameters α_1, α_2, respectively.

5.1. Results in a special bivariate case

We discuss some simulation results for the special case of the bivariate model specified in Section 3.2. We fixed $\sigma_{11l} = \sigma_{11}$, $\sigma_{22l} = \sigma_{22}$, $\sigma_{21l} = \sigma_{21}$, $c_{jl} = c$, and $\widetilde{\theta}_{jl} = \theta$ for all $l = 1, \ldots, L$; and $\alpha_1 = \alpha_2 = \alpha$. Since β and ν_{12} are not identifiable together, we set $\beta = 0$ in the model. Further, we chose $\sigma_{11} = \sigma_{22} = 1$, $\tau = 0.1$, $c = 2$, $\theta = 0.1$, $\alpha = 0.1$, $\sigma_{21} = 1$, and $\nu_{21} = 0.8$. We generated bivariate Gaussian data with mean $\mathbf{0}$. For estimation, we treated β, σ_{11}, σ_{22} and τ as *known*, and the other five parameters as *unknown* and estimated them from the data using the Gibbs sampling procedure.

From equations (3.8), (S1.1) and (S1.2), c^2 is a scale parameter, we used an $InvGamma(2, 1)$ prior for c^2. For the (positive) parameter α, we assumed a $Gamma(0.01, 10)$ prior. For θ, we assumed a $Gamma(0.1, 10)$ prior. Since

Table 2. Posterior mean, standard deviation, median and 95% credible intervals of parameters

Parameter value	Posterior values	$n = 15$	$n = 25$	$n = 50$
	Mean, s.d.	2.13, 0.43	1.96, 0.45	2.08, 0.49
$c = 2$	Median	2.07	1.91	2.08
	95% credible interval	(1.32, 3.12)	(1.21, 2.89)	(1.12, 3.05)
	Mean, s.d.	0.54, 0.28	0.59, 0.27	0.69, 0.21
$\nu_{21} = 0.8$	Median	0.54	0.65	0.72
	95% credible interval	(0.03, 0.98)	(0.10, 0.98)	(0.22, 0.98)
	Mean, s.d.	0.93, 0.40	0.85, 0.37	1.27,0.36
$\sigma_{21} = 1$	Median	0.96	0.87	1.27
	95% credible interval	(0.03, 1.66)	(0.06, 1.54)	(0.50, 1.92)
	Mean, s.d.	0.10, 0.10	0.12, 0.10	0.12, 0.10
$\alpha = 0.1$	Median	0.07	0.09	0.09
	95% credible interval	(0.003, 0.37)	(0.005, 0.30)	(0.007, 0.35)
	Mean, s.d.	0.30, 0.25	0.27, 0.33	0.13, 0.11
$\theta = 0.1$	Median	0.24	0.15	0.10
	95% credible interval:	(0.02, 0.97)	(0.02, 1.18)	(0.01, 0.42)

ν_{21} is restricted to the interval $(-1, 1)$, and is a measure of global association between processes, we assumed a positive association through a Uniform$(0, 1)$ prior. Finally, we chose a N$(0, 10)$ prior for σ_{21}.

The posterior distribution of c^2 is an Inverse Gamma. The posterior distributions of the rest of the parameters do not have closed forms. Hence we employed Gibbs sampling within a Metropolis Hastings algorithm to obtain posterior samples of the parameters. Burn-in was obtained with $2,000$ iterations, and we thinned the samples by 20 iterations to obtain $1,000$ uncorrelated samples from the joint posterior distribution of $(c, \theta, \alpha, \sigma_{21}, \nu_{21})$ given the data. Sensitivity analysis of the priors was carried out by varying the means and variances. The priors prove to be fairly robust with respect to the posterior inference results. For data simulated using $n = 15, 25$ and 50 locations, we present the results of the posterior inference in Table 2. This table displays the posterior mean, standard deviation (s.d.), median, and 95% credible intervals for each of the five parameters treated as random in the model. Table 2 shows that, for each n, the 95% credible intervals contained the actual values of the parameters. The lengths of the intervals for ν_{21} and α were relatively large.

In order to gain further insight into the performance as sample size increases, we carried out a separate simulation study: we simulated 100 independent samples of sizes $n = 15, 25$ and 50, respectively, from the bivariate model for the specified values of the parameters, and ran the MCMC each time on each of

these samples. Boxplots of the mean and median of the posterior squared errors (SE) of the covariance terms at three different spatial locations for the three sample sizes are displayed in Figure 2. The reported values are generically of the form Mean/Median $(SE(\mathrm{Cov}\,(Y_k(s_i), Y_{k'}(s_{i'})))/\{\mathrm{Cov}\,(Y_k(s_i), Y_{k'}(s_{i'}))\}^2)$, i.e, they represent the mean and the median of the standardized forms of the posterior SE From Figure 2, we observe that the means and medians of the posterior standardized SE were rather small, and these values decreased with larger sample sizes, as was to be expected.

Further, to compare the prediction performance of our model with that of a known bivariate stationary spatial model, we used the coregionalization model (stationary) of Wackernagel (2003) as implemented by the `spBayes` package (Finley, Banerjee and Carlin (2007)) in `R`, and compared predictive distributions of terms such as $(\theta^{pred} - \theta)^2$, where θ is the variance or covariance of the data generated using the true model (i.e, the generalized convolution model) at specific locations, and θ^{pred} are the posterior predictive sample estimates of θ. We compared the two fitted models for $n = 25$ and $n = 50$, using medians of $(\theta^{pred} - \theta)^2/\theta^2$ for standardizing the results. The `spBayes` package uses priors with large variances (infinite variance for scale or variance parameters) for all but one of the parameters used in the model, and that is also the case in our model. One parameter, namely the decay parameter, was assigned a mean of 0.18 and variance of 0.54 in the model implemented by `spbayes`. The generalized convolution model, on the other hand, assigned a prior to the parameter α with mean 0.1 and variance 1. No hyper-prior was used in either model. Figure 3 clearly shows the stark contrast in performance of the stationary coregionalization versus nonstationary generalized convolution model. For our model, the median of posterior predictive squared errors was close to 0 for all values, whereas many of the values of this performance measure corresponding to the stationary model were extremely large. This seems to indicate that for cases where nonstationarity prevails in the underlying multivariate spatial processes, our model may be a better choice than the regular coregionalization model of Wackernagel (2003), as implemented by the `spBayes` package.

6. Discussion

As a remark on an alternative method for fitting models of this type using a frequentist approach, we note that, for one dimensional time series, several aspects of estimating locally stationary processes have been explored (Dahlhaus (1996, 1997)), Dahlhaus and Neumann (2001). One approach due to Dahlhaus (1997) is to consider *tapered local periodogram* estimates of the *evolutionary spectrum* of the time series, then to minimize an asymptotic Kullback-Leibler divergence functional with respect to the parameters. Fuentes (2007) gives a formulation similar to that of Dahlhaus (1996, 1997) involving an approximation of

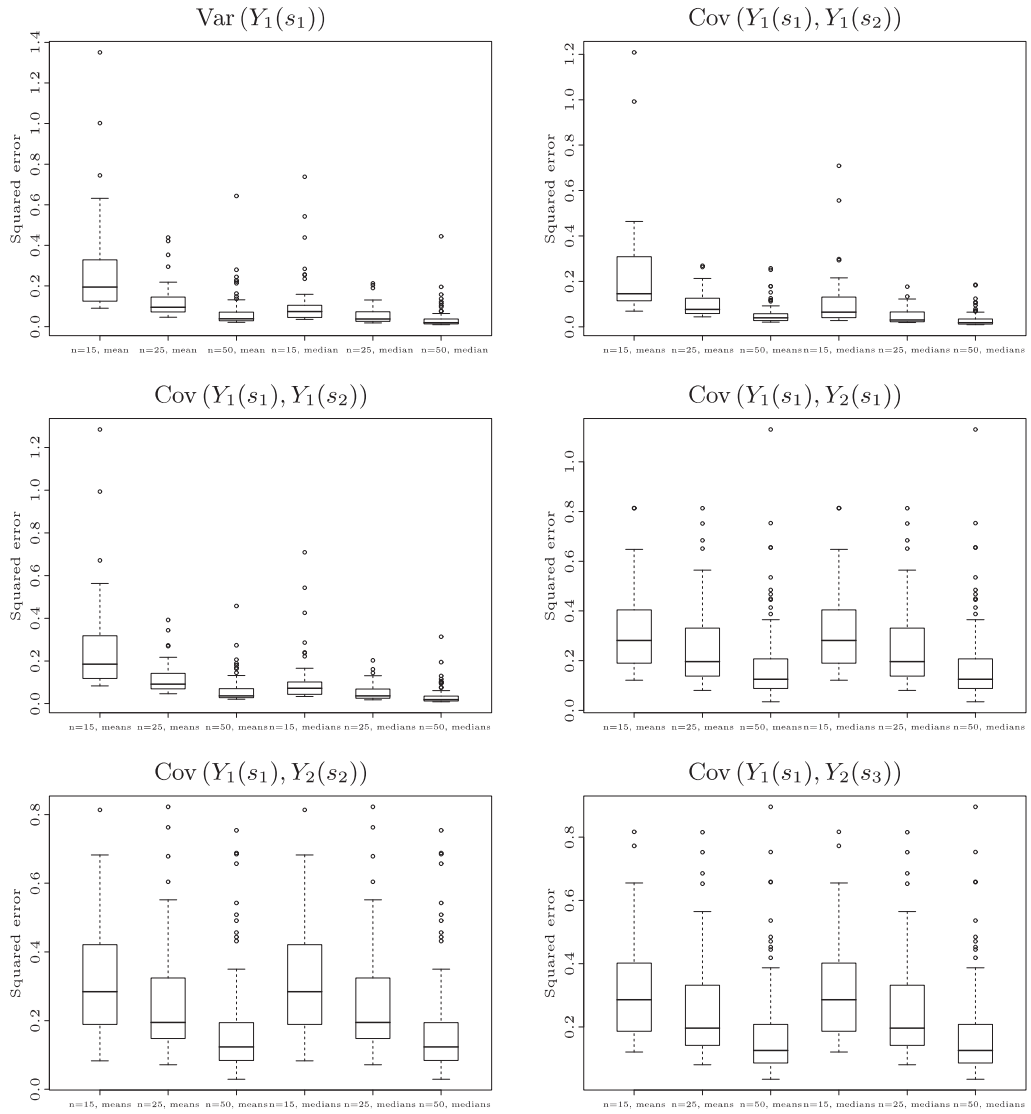

Figure 2. Mean and median posterior standard deviations of "standardized" values of variance and covariance values for $n = 15, 25$ and 50 (based on 100 simulations using the gen. conv. model)

the asymptotic likelihood, and discusses its implications for the estimation of the spectral density of the process in the case of data observed on an incomplete lattice in $\mathbb{R}^d$. The properties of estimators based on this type of likelihood approximation for correlated, multi-dimensional processes are currently under investigation.

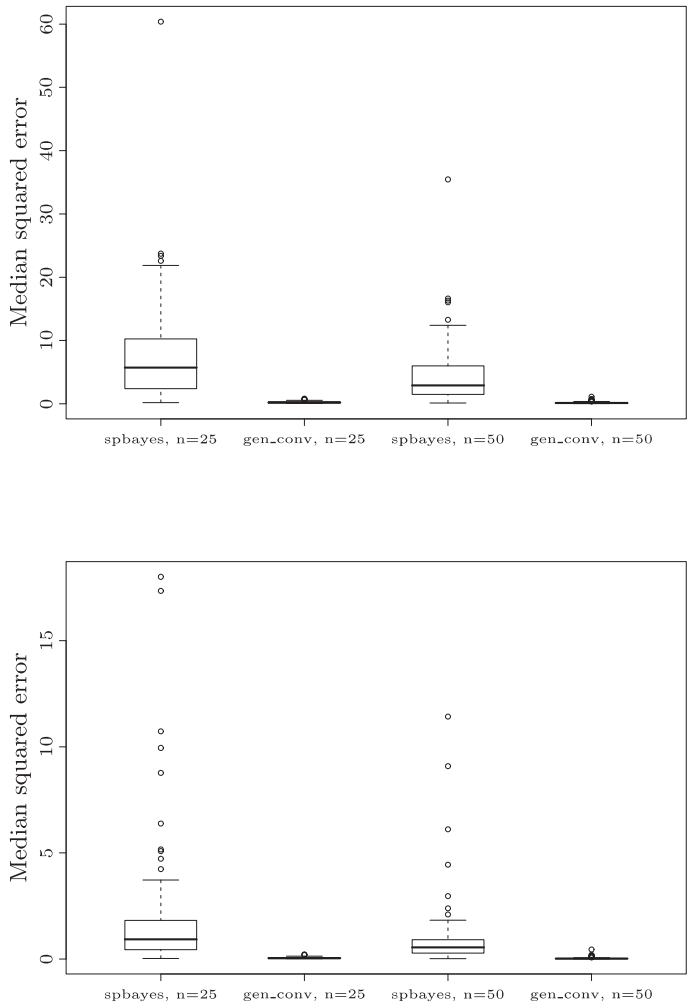

Figure 3. Median posterior predictive "standardized" squared error with sample size $n = 25$ and $n = 50$ for $Var(Y_1(s_1))$ (upper panel) and $\mathrm{Cov}\,(Y_1(s_1), Y_2(s_1))$ (bottom panel) comparing the `spbayes` and the gen. conv. model

Acknowledgement

The authors would like to thank the reviewers for their valuable comments and suggestions. Majumdar's research was supported by grants from NSF.

References

Billingsley, P. (1999). *Convergence of Probability Measures*. Wiley, New York.

Christensen, W. F. and Amemiya, Y. (2002). Latent variable analysis of multivariate spatial data. *J. Amer. Statist. Assoc.* **97**, 302-317.

Cressie, N. (1993). *Statistics for Spatial Data*. Wiley-Interscience.

Dahlhaus, R. (1996). On the Kullback-Leibler information divergence of locally stationary processes. *Stochastic Process. Appl.* **62**, 139-168.

Dahlhaus, R. (1997). Fitting time series models to nonstationary processes. *Ann. Statist.* **25**, 1-37.

Dahlhaus, R. and Neumann, M. H. (2001). Locally adaptive fitting of semiparametric models to nonstationary time series. *Stochastic Process. Appl.* **91**, 277-308.

Finley, A. O., Banerjee, S. and Carlin, B. P (2007). `spBayes`: an R package for univariate and multivariate hierarchical point-referenced spatial models. *J. Statist. Software* **19**, 4.

Fuentes, M. and Smith, R. (2001). A new class of nonstationary models. Technical Report # 2534, North Carolina State University, Institute of Statistics.

Fuentes, M. (2002). Spectral methods for nonstationary spatial processes. *Biometrika* **89**, 197-210.

Fuentes, M. (2007). Approximate likelihood for large irregularly spaced spatial data. *J. Amer. Statist. Assoc.* **102**, 321-331.

Fuentes, M., Chen, L., Davis, J. and Lackmann, G. (2005). Modeling and predicting complex space-time structures and patterns of coastal wind fields. *Environmetrics* **16**, 449-464.

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

Higdon, D. M. (1997). A process-convolution approach for spatial modeling, computer science and statistics, In *Proceedings of the 29th Symposium Interface*. (Edited by D. Scott).

Higdon, D. M. (2002). Space and space-time modeling using process convolutions, *Quantitative Methods for Current Environmental Issues* (C. Anderson, et al. eds), 3754. Springer, London.

Higdon, D. M., Swall, J. and Kern., J. (1999). Non-stationary spatial modeling. In *Bayesian Statistics* **6**. (Edited by J. M. Bernardo et al.), 761-768. Oxford University Press.

Majumdar, A. and Gelfand, A. E. (2007). Multivariate spatial modeling for geostatistical data using convolved covariance functions. *Mathematical Geology* **39**(2), 225-245.

Nychka, D., Wikle, C. and Royle, A. (2002). Multiresolution models for nonstationary spatial covariance functions. *Stat. Model.* **2**, 299-314.

Paciorek, C. J. and Schervish, M. (2006). Spatial modeling using a new class of nonstationary covariance function. *Enviornmentrics* **17**, 483-506.

Pintore, A. and Holmes, C. (2006). Spatially adaptive non-stationary covariance functions via spatially adaptive spectra. *J. Amer. Statist. Assoc.* (to appear).

Robert, C. and Casella, G. (2004). *Monte Carlo Statistical Methods*, 2nd edition. Springer.

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

Schabenberger, O. and Gotway, C. A. (2005). *Statistical Methods for Spatial Data Analysis*. Chapman & Hall/CRC.

Stein, M. L. (1999). *Interpolation of Spatial data: Some Theory for Kriging*. Springer, New York.

Wackernagel, H. (2003). *Multivariate Geostatistics: An Introduction with Applications*. 3rd edition. Springer, Berlin.

Yaglom, A. M. (1962). *An Introduction to the Theory of Stationary Random Functions*. Prentice-Hall.

Department of Mathematics and Statistics, Arizona State University, Tempe, Tempe, Arizona 85287-1804, U.S.A.

E-mail: ananda@math.asu.edu

Department of Statistics, University of California, Davis, California 95616, U.S.A.

E-mail: debashis@wald.ucdavis.edu

1958 Neil Avenue, Department of Statistics, Ohio State University, Columbus, OH 43210-1247, U.S.A.

E-mail: bautista.30@osu.edu

(Received July 2007; accepted January 2009)