

Some Topics in Convolution-Based Spatial Modeling

Catherine A. Calder¹ and Noel Cressie

Department of Statistics
The Ohio State University
1958 Neil Ave.
Columbus, OH 43210 USA

Appears in the *Proceedings of the 56th Session of the International Statistics Institute*. Lisbon, Portugal. August 22-29, 2007.

Over the last decade, convolution-based models for spatial data have increased in popularity as a result of their flexibility in modeling spatial dependence and their ability to accommodate large datasets. The modeling flexibility is due to the framework's moving-average construction that guarantees a valid (*i.e.*, non-negative definite) spatial covariance function. This constructive approach to spatial modeling has been used (1) to provide an alternative to the standard classes of parametric variogram/covariance functions commonly used in geostatistics; (2) to specify Gaussian-process models with nonstationary and anisotropic covariance functions; and (3) to create non-Gaussian classes of models for spatial data. Beyond the flexible nature of convolution-based models, computational challenges associated with modeling large datasets can be alleviated in part through dimension reduction, where the dimension of the convolved process is less than the dimension of the spatial data. In this paper, we review various types of convolution-based models for spatial data and point out directions for future research.

We consider models for continuously indexed spatial (*i.e.*, geostatistical) data, and let $\mathbf{Z} = (Z(\mathbf{s}_1), \dots, Z(\mathbf{s}_n))'$ be an $n \times 1$ vector of observations associated with known locations $\{\mathbf{s}_i : i = 1, \dots, n\} \subset \mathcal{D} \subset \mathbb{R}^d$. Using the observations \mathbf{Z} , we wish to make inference on an underlying process $Y(\cdot)$ that, when perturbed by random noise, yields the observation process $Z(\cdot)$. Here we assume that for all $\mathbf{s} \in \mathcal{D}$, $Z(\mathbf{s}) = Y(\mathbf{s}) + \epsilon(\mathbf{s})$, where the noise process $\epsilon(\cdot)$ has mean zero, variance σ_ϵ^2 , and is independent across space, and the processes $Y(\cdot)$ and $\epsilon(\cdot)$ are independent. To simplify our discussion of convolution-based models for spatial data, we modify the model specification above by taking the process $Y(\cdot)$ to have mean zero and decomposing $Z(\cdot)$ as follows:

$$Z(\mathbf{s}) = \mu(\mathbf{s}) + Y(\mathbf{s}) + \epsilon(\mathbf{s}), \quad \mathbf{s} \in \mathcal{D}.$$

¹E-mail: calder@stat.osu.edu

The goal is then to use the data \mathbf{Z} to infer the unknown parameters in the mean function $\mu(\cdot)$ and the covariance structure of the spatially dependent process $Y(\cdot)$. In this paper, we consider convolution-based models for $Y(\cdot)$.

In the traditional geostatistical setting (*e.g.*, Cressie, 1993), the covariance function of $Y(\cdot)$, $C_Y(\mathbf{s}_1, \mathbf{s}_2) \equiv \text{cov}(Y(\mathbf{s}_1), Y(\mathbf{s}_2))$, defined for all $\mathbf{s}_1, \mathbf{s}_2 \in \mathcal{D}$, is usually assumed to belong to a parametric class of covariance functions $\{C^{(P)}(\cdot, \cdot; \theta) : \theta \in \Theta\}$, where for all $\theta \in \Theta$, $C^{(P)}(\cdot, \cdot; \theta)$ is a valid covariance function (*i.e.*, non-negative definite). For a stationary covariance function, we denote $C_Y(\mathbf{s}_1, \mathbf{s}_2) = C_Y^0(\mathbf{h})$, where $\mathbf{h} = \mathbf{s}_1 - \mathbf{s}_2$. Then Bochner's Theorem states that

$$C_Y^0(\mathbf{h}) = \int_{\mathbb{R}^d} e^{i\boldsymbol{\omega}'\mathbf{h}} F(d\boldsymbol{\omega}),$$

where F is a nonnegative symmetric measure on \mathbb{R}^d . The process $Y(\cdot)$ can also be written as

$$Y(\mathbf{s}) = \int_{\mathbb{R}^d} e^{i\boldsymbol{\omega}'\mathbf{s}} V(d\boldsymbol{\omega}),$$

where $V(\cdot)$ is a process with independent increments and

$$\mathbb{E}[|V(d\boldsymbol{\omega})|^2] = F(d\boldsymbol{\omega}).$$

Now assume that $Y(\cdot)$ is Gaussian, and write

$$Y(\mathbf{s}) = \int_{\mathbb{R}^d} k(\mathbf{s}, \mathbf{u}) W(d\mathbf{u}), \mathbf{s} \in \mathcal{D}, \quad (1)$$

where $k(\cdot, \cdot)$ is a square-integrable (*i.e.*, $\int k^2(\mathbf{s}, \mathbf{u}) d\mathbf{u} < M < \infty$) kernel function and $W(\cdot)$ is d -dimensional Brownian motion (*e.g.*, Yaglom, 1987). We can replace $W(\cdot)$ with a general process $V(\cdot)$ whose increments are independent, have mean zero, and have finite variance proportional to the volume of the increment. Due to the independent increments of $V(\cdot)$, it follows that the covariance function of $Y(\cdot)$ is indirectly implied through the choice of the kernel function $k(\cdot, \cdot)$. That is,

$$C_Y(\mathbf{s}_1, \mathbf{s}_2) = \int_{\mathbb{R}^d} k(\mathbf{s}_1, \mathbf{u}) k(\mathbf{s}_2, \mathbf{u}) d\mathbf{u},$$

where, without loss of generality, $E(V(d\mathbf{u})^2) = d\mathbf{u}$. Then, for any finite number k and any real numbers $\{a_i : i = 1, \dots, k\}$,

$$\sum_{i=1}^k \sum_{j=1}^k a_i a_j C_Y(\mathbf{s}_i, \mathbf{s}_j) = \sum_{i=1}^k \sum_{j=1}^k a_i a_j \int_{\mathbb{R}^d} k(\mathbf{s}_i, \mathbf{u}) k(\mathbf{s}_j, \mathbf{u}) d\mathbf{u} = \int_{\mathbb{R}^d} \left(\sum_{i=1}^k a_i k(\mathbf{s}_i, \mathbf{u}) \right)^2 d\mathbf{u} \geq 0.$$

This is precisely the required non-negative-definiteness condition, and hence $C_Y(\cdot, \cdot)$ is a valid covariance function. This constructive approach to specifying a Gaussian process was suggested by Matérn (1986) and Thiébaux and Pedder (1987) and has been used to develop classes of models

for spatial data that are referred to as spatial moving averages (SMA) or process convolutions; we refer to this general class of spatial models as *convolution-based* to emphasize its origins and its generality. In virtually all that is to follow, we make the Gaussian assumption (that is, the statistical models are based on the representation of $Y(\cdot)$ given in (1)), although this assumption is mostly not needed.

Rather than estimating parameters in the covariance function, it is common in geostatistics to work with the variogram function. Letting $\mathbf{h} = \mathbf{s}_1 - \mathbf{s}_2 \in \mathbb{R}^d$, the variogram of $Y(\cdot)$ at lag \mathbf{h} is defined to be

$$2\gamma_Y(\mathbf{h}) \equiv \text{var} [Y(\mathbf{s}_1) - Y(\mathbf{s}_2)] , \quad \mathbf{s}_1, \mathbf{s}_2 \in \mathcal{D} ,$$

provided that the right-hand side is a function of $\mathbf{h} = \mathbf{s}_1 - \mathbf{s}_2$ (which is true if $Y(\cdot)$ is intrinsically stationary). Suppose we consider a special case of (1) where a Gaussian process is constructed as

$$Y(\mathbf{s}) = \int_{\mathbb{R}^d} k^0(\mathbf{s} - \mathbf{u})W(d\mathbf{u}), \mathbf{s} \in \mathcal{D} , \quad (2)$$

where $k^0(\cdot)$ is a kernel function in \mathbb{R}^d . In this situation, the variogram can be written as a function of the kernel $k^0(\cdot)$, as follows:

$$2\gamma(\mathbf{h}) = \int_{\mathbb{R}^d} (k^0(\mathbf{u}) - k^0(\mathbf{u} - \mathbf{h}))^2 d\mathbf{u}. \quad (3)$$

Further, the covariance function of $Y(\cdot)$ is stationary (*i.e.*, $C_Y^0(\mathbf{h}) = C_Y(\mathbf{s}_1, \mathbf{s}_2)$, where $\mathbf{h} = \mathbf{s}_1 - \mathbf{s}_2$), since

$$C_Y^0(\mathbf{h}) = \int_{\mathbb{R}^d} k^0(\mathbf{u})k^0(\mathbf{u} - \mathbf{h}) d\mathbf{u}. \quad (4)$$

In this case, we can relate the kernel $k^0(\cdot)$ and the covariance function $C_Y^0(\cdot)$ using the convolution theorem for the Fourier transform, which states that the Fourier transform of the convolution of two functions is proportional to the product of the Fourier transforms of each individual function. Therefore, if we take the square root of the Fourier transform of $C_Y^0(\cdot)$, and then we take its inverse Fourier transform, we obtain a function that is proportional to the kernel $k^0(\cdot)$ (see Kern, 2000, for details on and examples of this result). Thus, depending on the assumptions made about the process $Y(\cdot)$, expressions (1) - (2) can be used to motivate a convolution-based representation of the covariance function of $Y(\cdot)$. This is a powerful approach due to the flexibility in the choice of the kernel $k(\cdot, \cdot)$ or $k^0(\cdot)$.

Frequently, convolution-based models are implemented using discretized versions of (1) - (2). To illustrate this in the most general case (Eq. (1)), we define $\{\omega(\mathbf{u}_i) : i = 1, \dots, m\}$ to be a discrete white-noise process, namely a collection of m independent normal random variables with mean zero and variance λ^2 , associated with fixed regular-lattice locations $\{\mathbf{u}_i : 1, \dots, m\} \subset \mathcal{D}$. Then (1) can be approximated as follows:

$$Y(\mathbf{s}) \approx \sum_{i=1}^m k(\mathbf{s}, \mathbf{u}_i)\omega(\mathbf{u}_i). \quad (5)$$

The right-hand side of (5) can be made to converge to that of (1) by (say) successively increasing the density of the lattice locations by a factor of two in each dimension and decreasing the variance of the $\{\omega(\mathbf{u}_i)\}$ by a factor of 2^d . Besides being used to approximate the integral in the convolution-based representations of $Y(\cdot)$, the discretized convolution framework given by (5) has been extended to allow the $\{\omega(\mathbf{u}_i)\}$ to be dependent processes themselves; examples of such approaches are discussed in Lee et al. (2005) and Cressie and Johannesson (2006).

In the remaining sections of this paper, we review the various ways in which convolution-based representations of Gaussian processes have been exploited to develop flexible and computationally efficient statistical models for spatial data.

Nonparametric Covariance Functions

One of the first examples of the use of convolution-based representations of Gaussian processes in spatial statistics appeared in Barry and Ver Hoef (1996). They proposed a “blackbox” approach to kriging (*i.e.*, best linear unbiased spatial prediction; see, for example, Cressie, 1993, Ch. 3) using a flexible nonparametric specification of the variogram represented by (3). To illustrate their approach, we take $Y(\cdot)$ to be a process defined on a subset of \mathbb{R}^1 (*i.e.*, $d = 1$). In this case, Barry and Ver Hoef’s nonparametric kernels are taken to be piecewise constant functions of the form,

$$k^0(h) = \sum_{j=1}^p a_j \mathcal{I} \left(\frac{(j-1)c}{p} < h \leq \frac{jc}{k} \right),$$

where $\mathcal{I}(\cdot)$ denotes the indicator function, so that $k^0(\cdot)$ has p steps each of range $c > 0$. Using (3) and assuming that the lag h is a multiple m of c/p , a variogram of the form,

$$\begin{aligned} 2\gamma(h) &= \frac{c}{p} \sum_{j=1}^m a_j^2 + \frac{c}{p} \sum_{j=m+1}^p (a_j - a_{j-m})^2 + \frac{c}{p} \sum_{j=p+1}^{p+m} a_{j-m}^2 \\ &= \frac{2c}{p} \sum_{j=1}^p a_j^2 - \frac{2c}{p} \sum_{j=m+1}^p a_j a_{j-m}, \end{aligned}$$

is obtained. Then, Barry and Ver Hoef (1996) show that these evaluations of $2\gamma(h)$ at multiples of c/p can be linearly interpolated to produce a valid variogram defined for all h :

$$2\gamma(h) = (1 - V)2\gamma(q_l c/p) + V2\gamma(q_u c/p),$$

where $q_l = \lfloor hp/c \rfloor$ is the greatest integer less than or equal to hp/c , $q_u = \lceil hp/c \rceil$ is the smallest integer greater than or equal to hp/c , and $V = [h - (q_l c/p)]/(c/p)$ is the fraction that the distance h is from q_l to q_u . Similarly, in two dimensions, a nonparametric variogram can be produced using kernels that are rectangles. In a later paper, Ver Hoef et al. (2004) show how the Fast Fourier Transform (FFT) can be used to increase computational efficiency in calculating such nonparametric variograms. A similar nonparametric convolution-based approach to modeling spatial processes in subsets of \mathbb{R}^2 based on kernels constructed by stacking cylinders of varying heights and decreasing radii was explored in Kern (2000).

Dimension Reduction

Models based on the discrete form of the convolution-based representation of $Y(\cdot)$ (Eq. (5)) with a sparse support set for the latent process $\omega(\cdot)$ have been proposed to alleviate the computation burden associated with fitting standard geostatistical models (*e.g.*, kriging) to large datasets. In his model for ocean temperatures in the North Atlantic, Higdon (1998) proposed a space-time model based on discrete convolutions: a three-dimensional kernel with a separable (in space and time) structure is used to convolve a latent three-dimensional discrete white-noise process on a spatially sparse grid covering the study region. Both the coarse resolution of the latent process and the separable form of the kernel result in dramatic reductions in computational burden over what would have been required to fit a standard Gaussian process model with a nonseparable covariance function to the data directly.

As a general framework for dimension reduction in spatial modeling, Higdon (2002) considers models for a $n \times 1$ -dimensional data vector \mathbf{Z} of the form

$$\mathbf{Z} = \mu \mathbf{1}_n + \mathbf{K} \boldsymbol{\omega} + \boldsymbol{\epsilon},$$

where $\mathbf{1}_n$ is an n -dimensional vector of 1s, \mathbf{K} is a $n \times m$ matrix with entries $k(\mathbf{s}_i, \mathbf{u}_j)$, $\boldsymbol{\omega}$ is an m -dimensional vector that is assumed to be $N(\mathbf{0}, \tau^2 \mathbf{I}_m)$, $\boldsymbol{\epsilon}$ is an n -dimensional vector that is assumed to be $N(\mathbf{0}, \sigma_\epsilon^2 \mathbf{I}_n)$, and $\boldsymbol{\omega}$ and $\boldsymbol{\epsilon}$ are independent. Written in this manner, it is clear that the discrete convolution-based model can be thought of as a linear mixed-effects model (*e.g.*, Pinheiro and Bates, 2000). Thus, model-fitting can be carried out using functions in standard statistical software packages.

This discrete convolution-based approximation to a Gaussian process used in Higdon (1998) and Higdon (2002) relies on using a kernel function that has a Gaussian form. It has been demonstrated through simulation studies in \mathbb{R}^2 that the distance between the lattice points should be no more than the parameter τ of the Gaussian kernel. For more peaked (*e.g.*, exponential) kernels, a finer lattice for the latent process is required and the computational savings over traditional kriging are not as substantial.

Nonstationary and Anisotropic Models

One of the primary uses of a convolution-based representation of a Gaussian process is to build models with nonstationary covariance structures. Two distinct approaches have been developed in the literature, due to Higdon (Higdon, 1998; Higdon et al., 1998) and Fuentes (Fuentes, 2002a,b). The Higdon approach is to rewrite (4) as

$$Y(\mathbf{s}) = \int_{\mathcal{D}} k_{\mathbf{s}}^0(\mathbf{s} - \mathbf{u}) W(d\mathbf{u}), \quad (6)$$

where the *parameters* of the kernel $k(\mathbf{s}, \mathbf{u}) \equiv k_{\mathbf{s}}^0(\mathbf{s} - \mathbf{u})$ are allowed to vary spatially. Using a discretized version of (6), Higdon (1998) defines the spatially referenced kernel $k_{\mathbf{s}}^0(\cdot)$ to be a mixture of a finite set of Gaussian kernels, where the mixing weights vary smoothly with $\mathbf{s} \in \mathcal{D}$.

This approach was extended in Higdon et al. (1998) to a more general class of nonstationary and anisotropic convolution-based models. The kernels were chosen to have a Gaussian form:

$$k_{\mathbf{s}}^0(\mathbf{h}) = (2\pi)^{-1} |\Sigma(\mathbf{s})|^{-1/2} \exp(-\mathbf{h}'\Sigma(\mathbf{s})^{-1}\mathbf{h}/2), \mathbf{h} \in \mathbb{R}^2. \quad (7)$$

The covariance matrix of the Gaussian kernel associated with location \mathbf{s} , $\Sigma(\mathbf{s})$, is then parameterized by the lengths of the major and minor axes of its one-standard-deviation ellipse, namely $a_{\mathbf{s}}$ and $b_{\mathbf{s}}$, respectively, and the rotation angle $\theta_{\mathbf{s}}$:

$$\Sigma(\mathbf{s}) = \begin{pmatrix} a_{\mathbf{s}}^2 \cos^2(\theta_{\mathbf{s}} - \pi/2) + b_{\mathbf{s}}^2 \sin^2(\theta_{\mathbf{s}} - \pi/2) & \sin(\theta_{\mathbf{s}} - \pi/2) \cos(\theta_{\mathbf{s}} - \pi/2)(b_{\mathbf{s}}^2 - a_{\mathbf{s}}^2) \\ \sin(\theta_{\mathbf{s}} - \pi/2) \cos(\theta_{\mathbf{s}} - \pi/2)(b_{\mathbf{s}}^2 - a_{\mathbf{s}}^2) & a_{\mathbf{s}}^2 \sin^2(\theta_{\mathbf{s}} - \pi/2) + b_{\mathbf{s}}^2 \cos^2(\theta_{\mathbf{s}} - \pi/2) \end{pmatrix}.$$

Figure 1 illustrates a nonstationary and anisotropic Gaussian process constructed using the Higdon approach. The plot on the left shows the one-standard-deviation ellipses (shrunk by a factor of five) of the Gaussian kernels associated with locations on a 10×10 grid (*i.e.*, the dimension of the latent process $\omega(\cdot)$, m , is 100). The plot on the right shows a realization of the resulting Gaussian process on a finer resolution 21×21 grid. As expected, the process appears smoother in the top right-hand corner where the ellipses are larger, and it exhibits varying amounts of directional-dependence structure corresponding to regions with noncircular ellipses.

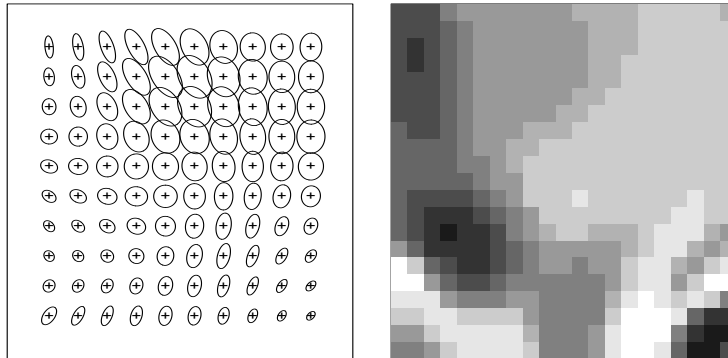


Figure 1: One standard deviation ellipses (shrunk by a factor of five) of the Gaussian kernels (left) and a realization of the resulting nonstationary and anisotropic Gaussian process on a 21×21 grid (right).

Paciorek and Schervish (2006) extend these nonstationary process-convolution models proposed by Higdon et al. (1998). First, they consider a more easily generalizable eigen-decomposition specification for the spatially varying kernel covariance matrices $\Sigma(\mathbf{s})$. In addition, they develop a general class of nonstationary covariance functions that includes a nonstationary Matérn covariance function. They advocate the use of this model since, like the corresponding stationary version of the Matérn covariance function, it has a parameter that controls the smoothness/differentiability of the spatial process.

The Fuentes approach to modeling nonstationary processes involves convolving stationary processes, as opposed to white-noise processes (Fuentes, 2002a,b). The spatial domain \mathcal{D} is divided

into m small subregions centered at nodes $\mathbf{u}_1, \dots, \mathbf{u}_m$. Defined on each subregion is a Gaussian stationary spatial process $Y_{\theta(\mathbf{u}_i)}$ with covariance function $C_{Y_{\theta(\mathbf{u}_i)}}$, where $\theta(\mathbf{u}_i)$ is an unknown region-specific parameter defining the covariance of the region's process $Y_{\theta(\mathbf{u}_i)}$. Then, a nonstationary process on \mathcal{D} is defined by taking

$$Y(\mathbf{s}) = \sum_{i=1}^m k^0(\mathbf{s} - \mathbf{u}_i) Y_{\theta(\mathbf{u}_i)}(\mathbf{s}),$$

where $k^0(\cdot)$ is a stationary kernel (*i.e.*, it does not vary across space). Therefore, unlike the Higdon approach that was based on allowing the kernels to vary spatially, nonstationarity in the Fuentes approach results from the spatially varying processes $\{Y_{\theta(\mathbf{u}_i)}\}$. This approach has been extended in Fuentes and Smith (2001) by allowing the parameter $\theta(\cdot)$ to vary smoothly across space (rather than at discrete nodes) and used to develop formal hypothesis tests for nonstationarity (Fuentes, 2002b, 2005). Related approaches include Nychka et al. (2002)'s model for nonstationary spatial data based on linear combinations of wavelet basis functions and Banerjee et al. (2004)'s spatial model for house prices based on normalized distance-weighted sums of stationary processes.

Multivariate Spatial Models

Convolution-based models have also been developed for the analysis of multivariate spatial data. Ver Hoef and Barry (1998) and Ver Hoef et al. (2004) consider cross-covariance functions defined as

$$C_{Y_p, Y_q}^0(\mathbf{h}; \rho_p, \rho_q, \theta_p, \theta_q, \Delta_p, \Delta_q) = \rho_p \rho_q \int_D k^0(\mathbf{u}; \theta_p) k^0(\mathbf{u} - \mathbf{h} + \Delta_p - \Delta_q; \theta_q) d\mathbf{u},$$

where the kernel k^0 with parameters θ_p and θ_q describe the spatial dependencies in the marginal processes $Y_p(\cdot)$ and $Y_q(\cdot)$, ρ_p and ρ_q capture the cross-dependencies, Δ_p and Δ_q describe the shift-asymmetry, and the kernel functions are specified in a nonparametric manner as in Barry and Ver Hoef (1996); see also the section, Nonparametric Covariance Functions, above.

Other convolution-based multivariate spatial models include the multivariate space-time models in Calder (2007a,b) and the convolutions of covariance functions developed in Majumdar and Gelfand (2007).

Spatio-temporal Models

Convolution-based models for space-time data can be developed in a straightforward manner by simply convolving a three-dimensional kernel with a three-dimensional white-noise process (*e.g.*, Higdon, 1998). However, capturing dynamic features such as seasonality or directional space-time dependence using kernel functions may be problematic. Additionally, the temporal domain of many space-time processes is discrete; this feature is difficult to make explicit in standard convolution-based modeling. An alternative approach proposed by Calder et al. (2002) and Higdon (2002) is to convolve dynamic processes using two-dimensional spatial kernels. Since it is based on convolutions of dependent processes, this approach has a similar flavor to the Fuentes approach to modeling nonstationary spatial processes (Fuentes, 2002a,b).

As an illustration of these dynamic convolution-based models, Calder et al. (2002) proposed a model for ozone levels across the Eastern U.S. based on latent space-time processes of the form

$$Y(\mathbf{s}, t) = \sum_{i=1}^m k^0(\mathbf{s} - \mathbf{u}_i) Y_{\mathbf{u}_i}(t).$$

The processes $\{Y_{\mathbf{u}_i}(\cdot)\}$ evolve as

$$Y_{\mathbf{u}_i}(t) = f(Y_{\mathbf{u}_1}(t-1), \dots, Y_{\mathbf{u}_m}(t-1); \beta) + \nu_{\mathbf{u}_i}(t),$$

where $f(\cdot; \beta)$ is a parametric function of the values of the latent processes $\{Y_{\mathbf{u}_i}\}$ at the previous time step and $\{\nu_{\mathbf{u}_i}\}$ are independent Gaussian-process innovations. A related approach involving convolutions of latent autoregressive processes was proposed by Sanso et al. (2005).

An alternative convolution-based space-time modeling framework was proposed by Wikle and Cressie (1999), Wikle (2002), and Xu et al. (2005). Here, the evolution of the latent space-time process was specified as a convolution by letting

$$Y(\mathbf{s}, t) = \int k^0(\mathbf{s} - \mathbf{u}, t) Y(\mathbf{u}, t-1) d\mathbf{u} + \eta(\mathbf{s}, t),$$

where $\eta(\cdot, \cdot)$ is a spatially colored error process independent in time. Wikle (2002) and Xu et al. (2005) chose kernel functions corresponding to partial differential equations governing $Y(\cdot)$. When there is no readily available science to drive the dynamics, one might choose the orthogonal-based-function approach used by Wikle and Cressie (1999), or the dynamic linear modeling approach used by Calder et al. (2002) and extended in Calder (2007a,b).

Directions for Future Research

We conclude our review of convolution-based models for spatial processes by proposing some areas for future research.

Implications of Kernel Choice and Discretization

A kernel corresponding to a particular stationary covariance function $C_Y^0(\mathbf{h})$ can be obtained by Fourier inversion of $C_Y^0(\mathbf{h})$ (see above). However, in many situations, for a given kernel/covariance function there is not a closed-form expression for the covariance function/kernel. In addition, the functional form for a kernel corresponding to a particular stationary covariance function may not be the same for all dimensions d , and the existence of a closed-form expression for a covariance function corresponding to a kernel function depends on d (see Kern, 2000, and Cressie and Pavlicová, 2002). Thus, while the convolution-based framework permits a great deal of modeling flexibility through the choice of kernel function, it is important to understand the implications of the choice of kernel on the properties of the resulting process.

We believe that there is a need for careful consideration of the properties of Gaussian processes constructed via convolution-based models, along the lines of Cressie and Pavlicová (2002) and

Paciorek (2003). A related concern is the impact of the discretization on the properties of discrete convolution-based Gaussian processes, as well as the overall fit of the model. For the nonstationary convolution-based models, this issue is addressed somewhat in Fuentes and Smith (2001) and Paciorek (2003). However, there is still a need to look at the relationship between the sampling design of the data and the level of discretization, as well as how these features impact the ability to detect nonstationary and isotropic features of the data.

Non-Gaussian Convolution-Based Models

As we have pointed out earlier, the specification for the process $Y(\cdot)$ given in (1) can be readily extended by substituting for $W(\cdot)$ other stochastic processes with independent increments (e.g., Lévy processes). This approach has been considered by Wolpert and Ickstadt (1998), who used convolutions of gamma processes to develop a model for spatial point processes. Convolutions of Lévy processes have also been considered by Tu (2006) for modeling spatio-temporal processes.

Processes on Non-Euclidean Spaces

Any space upon which Brownian motion can be defined will in principle support a convolution-based Gaussian model, through a generalization of (1). For example, convolution-based models can be defined on the sphere. More generally, if there is a spectral theory, then the spectral representation hints at how a convolution-based model could be defined; for example, the spectral analysis for groups given in Diaconis (1988) might be explored for abstract convolution-based models.

A non-Euclidean model of immediate applicability on river networks is the class of convolution-based models on directed trees. Ver Hoef et al. (2006) and Cressie et al. (2006) define “spatial” models on river networks whose covariance between the process at two locations on the network is a function of river distance (rather than Euclidean distance) between the two locations. The convolution-based models use kernels $k(\mathbf{s}, \mathbf{u})$ that are *asymmetric*, reflecting the physical nature of a river that “downstream” should be treated differently from “upstream”. Indeed, it is possible that one location is neither upstream nor downstream from another location (e.g., they are on different tributaries), in which case their process values are independent if the kernel is one-sided.

References

- Banerjee, S., Gelfand, A. E., Knight, J. R., and Sirmans, C. F. (2004). “Spatial modeling of house prices using normalized distance-weighted sums of stationary processes.” *Journal of Business and Economic Statistics*, 22, 2, 206–213.
- Barry, R. and Ver Hoef, J. (1996). “Blackbox kriging: Spatial prediction without specifying variogram models.” *Journal of Agricultural, Biological, and Environmental Statistics*, 1, 297–322.
- Calder, C. A. (2007a). “A Bayesian dynamic process convolution approach to modeling PM2.5 and PM10 concentration levels.” *Environmetrics*. Forthcoming.

- (2007b). “Dynamic factor process convolution models for multivariate space-time data with application to air quality assessment.” *Environmental and Ecological Statistics*. Forthcoming.
- Calder, C. A., Holloman, C., and Higdon, D. (2002). “Exploring space-time structure in ozone concentration using a dynamic process convolution model.” In *Case Studies in Bayesian Statistics 6*, eds. C. Gatsonis, R. E. Kass, A. Carriquiry, A. Gelman, D. Higdon, D. Pauler, and I. Verdinelli, 165–176. New York: Springer-Verlag.
- Cressie, N. (1993). *Statistics for Spatial Data, Revised edn.*. New York: John Wiley.
- Cressie, N., Frey, J., Harch, B., and Smith, M. (2006). “Spatial prediction on a river network.” *Journal of Agricultural, Biological, and Environmental Statistics*, 11, 127–150.
- Cressie, N. and Johannesson, G. (2006). “Fixed rank kriging for large spatial datasets.” Tech. Rep. 780, Department of Statistics, The Ohio State University, Columbus, OH.
- Cressie, N. and Pavlicová, M. (2002). “Calibrated spatial moving average simulations.” *Statistical Modelling*, 2, 267–279.
- Diaconis, P. (1988). *Group Representations in Probability and Statistics*. IMS Lecture Notes - Monograph Series, vol. 11, Hayward, CA: Institute of Mathematical Statistics.
- Fuentes, M. (2002a). “Interpolation of nonstationary air pollution processes: a spatial spectral approach.” *Statistical Modelling*, 2, 281–298.
- (2002b). “Spectral methods for nonstationary spatial processes.” *Biometrika*, 89, 197–210.
- (2005). “A formal test for nonstationarity of spatial stochastic processes.” *Journal of Multivariate Analysis*, 96, 30–54.
- Fuentes, M. and Smith, R. (2001). “A new class of nonstationary spatial models.” Tech. Rep. 2534, Department of Statistics, North Carolina State University, Raleigh, NC.
- Higdon, D. (1998). “A process-convolution approach to modeling temperatures in the North Atlantic Ocean.” *Journal of Environmental and Ecological Statistics*, 5, 173–190.
- (2002). “Space and space-time modeling using process convolutions.” In *Quantitative Methods for Current Environmental Issues*, eds. C. Anderson, V. Barnett, P. C. Chatwin, and A. H. El-Shaarawi, 37–56. New York: Springer-Verlag.
- Higdon, D., Swall, J., and Kern, J. (1998). “Non-stationary spatial modeling.” In *Bayesian Statistics 6*, eds. J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith. Oxford: Oxford University Press, 761–768.
- Kern, J. (2000). “Bayesian process-convolution approaches to specifying spatial dependence structure.” Ph.D. thesis, Institute of Statistics and Decision Sciences, Duke University, Durham, NC.

- Lee, H. K. H., Higdon, D. M., Calder, C. A., and Holloman, C. H. (2005). “Efficient models for correlated data via convolutions of intrinsic processes.” *Statistical Modelling*, 5, 53–74.
- Majumdar, A. and Gelfand, A. (2007). “Multivariate spatial process modeling using convolved covariance functions.” *Mathematical Geology*. Forthcoming.
- Matérn, B. (1986). *Spatial Variation, Second edn.* New York: Springer-Verlag.
- Nychka, D., Wikle, C. K., and Royle, J. K. (2002). “Multiresolution models for nonstationary spatial covariance functions.” *Statistical Modelling*, 2, 315–331.
- Paciorek, C. J. (2003). “Nonstationary Gaussian processes for regression and spatial modelling.” Ph.D. thesis, Department of Statistics, Carnegie Mellon University, Pittsburgh, PA.
- Paciorek, C. J. and Schervish, M. J. (2006). “Spatial modelling using a new class of nonstationary covariance functions.” *Environmetrics*, 17, 483–506.
- Pinheiro, J. and Bates, D. (2000). *Mixed Effects Models in S and S-Plus.* New York: Springer-Verlag.
- Sanso, B., Schmidt, A., and Nobre, A. A. (2005). “Bayesian spatio-temporal models based on discrete convolutions.” Tech. rep., Departamento de Métodos Estatísticos, U Universidade Federal do Rio de Janeiro, Brazil.
- Thiébaux, H. J. and Pedder, M. A. (1987). *Spatial Objective Analysis with Applications in Atmospheric Science.* London: Academic Press.
- Tu, C. (2006). “Bayesian nonparametric modeling using Lévy process priors with applications for function estimation, times series modeling, and spatio-temporal modeling.” Ph.D. thesis, Institute of Statistics and Decision Sciences, Duke University, Durham, NC.
- Ver Hoef, J. and Barry, R. (1998). “Constructing and fitting models for cokriging and multivariable spatial prediction.” *Journal of Statistical Planning and Inference*, 69, 275–294.
- Ver Hoef, J., Cressie, N., and Barry, R. (2004). “Flexible spatial models for kriging and cokriging using moving averages and the Fast Fourier Transform (FFT).” *Journal of Computational and Graphical Statistics*, 13, 265–282.
- Ver Hoef, J. M., Peterson, E., and Theobald, D. (2006). “Spatial statistical models that use flow and stream distance.” *Environmental and Ecological Statistics*, 13, 449–464.
- Wikle, C. K. (2002). “A kernel-based spectral model for non-Gaussian spatio-temporal processes.” *Statistical Modelling*, 2, 299–314.
- Wikle, C. K. and Cressie, N. (1999). “A dimension-reduced approach to space-time Kalman filtering.” *Biometrika*, 86, 4, 815–829.

- Wolpert, R. L. and Ickstadt, K. (1998). “Poisson/gamma random field models for spatial statistics.” *Biometrika*, 85, 251–267.
- Xu, K., Wikle, C., and Fox, N. (2005). “A kernel-based spatio-temporal dynamical model for nowcasting radar precipitation.” *Journal of the American Statistical Association*, 100, 1133–1144.
- Yaglom, A. M. (1987). *Correlation Theory of Stationary and Related Random Functions*. New York: Springer-Verlag.