, 28, 1502–1505. 1972.
ference on Probability and
-193. 1975.
t particulate matter con-
ssoc., 64, 1172–1182. 1969.
a sets, J.R. Statist. Soc. B,
atistical Modeling, 2, 281–
he North Atlantic ocean,
ardo, J. Berger, A. Sawid,
ord, U.K., 761–768. 1999.
r, Test, 7, 217–285 (with
nce structure, Biometrika,
onary spatial covariance
orthogonal functions for
986.
onstationary covariance
sis functions in the spec-
New York, 1965.
a, Environ. Eco. Stat., 12,
ıssive data product from
inger-Verlag, New York.
a, J. Royal Statist. Soc. Ser.
ling bird survey data on
ster Modelling, Chapman
man filtering, Biometrika,
ical Bayesian modeling:
mic models, J. Stat. Plan.

9

Constructions for Nonstationary Spatial Processes

Paul D. Sampson

CONTENTS

9.1 Overview

Modeling of the spatial dependence structure of environmental processes is fundamental to almost all statistical analyses of data that are sampled spatially. The classical geostatistical model for a *spatial* process $\{Y(\mathbf{s}) : \mathbf{s} \in D\}$ defined over the spatial domain $D \subset \mathbb{R}^d$, specifies a decomposition into mean (or trend) and residual fields, $Y(\mathbf{s}) = \mu(\mathbf{s}) + e(\mathbf{s})$. The process is commonly assumed to be second order stationary, meaning that the spatial covariance function can be written $C(\mathbf{s}, \mathbf{s}+\mathbf{h}) = \text{Cov}(Y(\mathbf{s}), Y(\mathbf{s}+\mathbf{h})) = \text{Cov}(e(\mathbf{s}), e(\mathbf{s}+\mathbf{h})) = C(\mathbf{h})$, so that the covariance between any two locations depends only on the spatial lag vector connecting them. There is a long history of modeling the spatial covariance under an assumption of "intrinsic stationarity" in terms of the *semivariogram*, $\gamma(\mathbf{h}) = \frac{1}{2}\text{var}(Y(\mathbf{s}+\mathbf{h}) - Y(\mathbf{s}))$. However, it is now widely recognized that most, if not all, environmental processes manifest spatially nonstationary or heterogeneous covariance structure when considered over sufficiently large spatial scales.

A fundamental notion underlying most of the current modeling approaches is that the spatial correlation structure of environmental processes can be considered to be approximately stationary over relatively small or "local" spatial regions. This local structure is typically anisotropic. The methods can then be considered to describe spatially varying, locally stationary, anisotropic covariance structure. The models should reflect the effects of known explanatory environmental processes (wind/transport, topography, point sources, etc.). Ideally we would like to model these effects directly, but there have been only a few recent approaches aiming at such explicit modeling (see Calder, 2008).

We distinguish our focus on nonstationarity in spatial covariance from nonstationarity in the mean or trend, as commonly addressed by variants of universal kriging, and from nonstationary processes modeled by intrinsic functions of order k (IRF-F) and characterized by generalized covariance functions, including the one-dimensional special cases of fractional and integrated Brownian motions. Filtered versions of these processes, or "spatial increments of order k," are stationary. In some cases, appropriately identified universal kriging and intrinsic random function kriging are essentially equivalent (Christensen, 1990). See also Stein (2001) and Buttafuoco and Castrignanò (2005).

The "early" literature (reaching back only to the 1980s) on modeling of nonstationary spatial covariance structure was primarily in the context of models for space–time random fields. Prior to 1990, the only apparent approach to this feature of environmental monitoring data (outside of local analyses in subregions where the process might be more nearly stationary) derived from an empirical orthogonal function decomposition of the space–time data matrix, a technique common in the atmospheric science literature. Reference to this approach in the statistical literature dates at least back to Cohen and Jones (1969) and Buell (1972, 1978), although perhaps the most useful elaboration of the method for spatial analysis appears in Obled and Creutin (1986). A number of new computational approaches were introduced in the late 1980s and early 1990s, beginning with Guttorp and Sampson's spatial deformation approach, first mentioned in print in a 1989 comment in a paper by Haslett and Raftery (1989). Shortly following was Haas' "moving window" spatial estimation (Haas, 1990a, 1990b, 1995), although this approach estimates covariance structure locally without providing a (global) model; Sampson and Guttorp's elaboration of their first approach to the spatial deformation model based on multidimensional scaling (1992); an empirical Bayes shrinkage approach of Loader and Switzer (1992); and Oehlert's kernel smoothing approach (1993). Guttorp and Sampson (1994) reviewed this literature on methods for estimating heterogeneous spatial covariance functions with comments on further extensions of the spatial deformation method. In this chapter we focus on the developments from the late 1990s to the present, updating the review of methods provided by Sampson (2001). There has been considerable development and application of kernel and process convolution models, beginning with the work of Higdon (1998) and Fuentes (2001). But despite a substantial growth in the literature of methods on nonstationary modeling, there is almost no conveniently available software at this point in time for the various methods reviewed here. This chapter presents no illustrative case studies and we refer the reader to the original sources for applications.

We review the current literature under the headings of: smoothing and kernel methods, basis function models, process convolution models, and spatial deformation models, concluding with brief mention of parametric models and further discussion.

9.2 Smoothing and Kernel-Based Methods

Perhaps the simplest approaches to dealing with nonstationary spatial covariance structure begin either from the perspective of locally stationary models, which are empirically smoothed over space, or from the perspective of the smoothing and/or interpolation of empirical covariances estimated among a finite number of monitoring sites. Neither of these perspectives incorporate any other explicit modeling of the spatial heterogeneity in the spatial covariance structure. Haas' approach to spatial estimation for nonstationary processes (Haas 1990a, 1990b, 1995) simply computes local estimates of the spatial covariance structure, but does not integrate these into a global model. Oehlert's (1993) kernel smoothing approach and Loader and Switzer's (1992) empirical Bayesian shrinkage and interpolation both aim to smoothly interpolate empirical covariances.

Papers by Fuentes (2001, 2002a, 2002b) and by Nott and Dunsmuir (2002) propose conceptually related approaches for representing nonstationary spatial covariance structure in terms of spatially weighted combinations of stationary spatial covariance functions assumed to represent the local covariance structure in different regions. First, consider dividing the spatial domain D into k subregions S_i, each with a sufficient number of points to estimate a (stationary) variogram or spatial covariance function locally. Fuentes (2001) represents the spatial process $Y(\mathbf{s})$, as a weighted average of "orthogonal local stationary processes":

$$Y(\mathbf{s}) = \sum_{i=1}^{k} w_i(\mathbf{s})Y_i(\mathbf{s}) \tag{9.1}$$

where $w_i(\mathbf{s})$ is a chosen weight function, such as inverse squared distance between $\mathbf{s}$ and the center of subregion S_i. The nonstationary spatial covariance structure is given by

$$\begin{aligned}\mathrm{Cov}(Y(\mathbf{s}), Y(\mathbf{u})) &= \sum\nolimits_{i=1}^{k} w_i(\mathbf{s})w_i(\mathbf{u})\mathrm{Cov}(Y_i(\mathbf{s}), Y_i(\mathbf{u})) \\ &= \sum\nolimits_{i=1}^{k} w_i(\mathbf{s})w_i(\mathbf{u})C_{\theta_i}(\mathbf{s}-\mathbf{u})\end{aligned} \tag{9.2}$$

where $C_{\theta_i}(\mathbf{s}-\mathbf{u})$ represents a stationary spatial covariance function. Fuentes chooses the number of subgrids, k, using a Bayesian information criterion (BIC). The stationary processes $Y_i(\mathbf{s})$ are actually "local" only in the sense that their corresponding covariance functions, $C_{\theta_i}(\mathbf{s}-\mathbf{u})$, are estimated locally, and they are "orthogonal" by assumption in order to represent the overall nonstationary covariance simply as a weighted sum of covariances. Fuentes estimates the parameters with a Bayesian approach providing predictive distributions accounting for uncertainty in the parameter estimates without resorting to computationally intensive MCMC methods.

Fuentes and Smith (2001) proposed to extend the finite decomposition of $Y(x)$ of Fuentes (2001) to a continuous convolution of local stationary processes:

$$Y(\mathbf{x}) = \int_D w(\mathbf{x}-\mathbf{s})Y_{\theta(\mathbf{s})}(\mathbf{x})d\mathbf{s}. \tag{9.3}$$

Estimation would require that the spatial field of parameter vectors $\theta(s)$, indexing the stationary Gaussian processes, be constrained to vary smoothly. In practice, the integrals of (9.3) and spectral representations of the spatial covariance (Fuentes, 2002a) are approximated with discrete sums involving k independent spatial locations s_i and corresponding processes $Y_{\theta_i}(\mathbf{s})$, as in Equation (9.2) above. (See also Fuentes, 2002b.)

Nott and Dunsmuir's (2002) approach, proposed as a more computationally feasible alternative to something like the spatial deformation model of Sampson and Guttorp (1992), has the stated aim of reproducing an empirical covariance matrix at a set of monitoring sites and then describing the conditional behavior given monitoring site values with a collection of stationary processes. We will use the same notation as that above, although for Nott and Dunsmuir, i will index the monitoring sites rather than a smaller number of subregions, and the $C_{\theta_i}(x-y)$ represent local *residual* covariance structure after conditioning on values at the monitoring sites. These are derived from locally fitted stationary models. In their general case, Nott and Dunsmuir's representation of the spatial covariance structure can be written

$$\mathrm{Cov}(Y(\mathbf{x}), Y(\mathbf{y})) = \Sigma_0(\mathbf{x}, \mathbf{y}) + \sum\nolimits_{i=1}^{k} w_i(\mathbf{x})w_i(\mathbf{y})C_{\theta_i}(\mathbf{x}-\mathbf{y})$$

where $\Sigma_0(x, y)$ is a function of the empirical covariance matrix at the monitoring sites, $\mathbf{C} = [c_{ij}]$, and the local stationary models computed so that $\mathrm{Cov}(Y(\mathbf{x}_i), Y(\mathbf{x}_j)) = c_{ij}$. They further propose to replace the empirical covariance matrix $\mathbf{C}$ by the Loader and Switzer (1992) empirical Bayes shrinkage estimator $\hat{\mathbf{C}} = \gamma\mathbf{C} + (1-\gamma)\mathbf{C}_\theta$, where $\mathbf{C}_\theta$ is a covariance matrix obtained by fitting some parametric covariance function model. In this case, it can be shown that the Nott and Dunsmuir estimate for covariances between monitored and unmonitored sites is the same as that of the proposed extrapolation procedure of Loader and Switzer, but the estimate for covariances among unmonitored sites is different, and in particular, not dependent on the order with which these unmonitored sites are considered, as was the case for Loader and Switzer's proposal.

Guillot et al. (2001) proposed a kernel estimator similar to the one introduced by Oehlert (1993), although they do not reference this earlier work. Let D denote the spatial domain so that the covariance function $C(\mathbf{x}, \mathbf{y})$ is defined on $D \times D$, and suppose an empirical covariance matrix $\mathbf{C} = [c_{ij}]$ computed for sites $\{x_i, i = 1, \ldots, n\}$. Define a nonnegative kernel K integrating to one on $D \times D$ and let $K_\varepsilon(u, v) = \varepsilon^{-4} K(u/\varepsilon, v/\varepsilon)$ for any real positive ε. Then define a partition $\{D_1, \ldots, D_n\}$ of D (such as the Voronoi partition). The nonparametric, nonstationary estimator of C obtained by regularization of $\mathbf{C}$ is

$$\hat{C}_\varepsilon(\mathbf{x}, \mathbf{y}) = \sum_{i,j} c_{ij} \int_{D_i \times D_j} K_\varepsilon(\mathbf{x} - \mathbf{u}, \mathbf{y} - \mathbf{v}) d\mathbf{u} d\mathbf{v}. \tag{9.4}$$

The authors prove positive definiteness of the estimator for positive definite kernels, discuss selection of the bandwidth parameter ε, and demonstrate an application where, surprisingly, kriging with the nonstationary covariance model is outperformed by kriging with a fitted stationary model.

Finally, we note the nonsmooth, piecewise Gaussian model approach of Kim, Mallick and Holmes (2005), which automatically partitions the spatial domain into disjoint regions using Voronoi tessellations. This model structure, specifying stationary processes within regions (tiles of the tessellation) and independence across regions, is fitted within a Bayesian framework. It is applied to a soil permeability problem where this discrete nonstationary structure seems justified.

9.3 Basis Function Models

The earliest modeling strategy in the literature for nonstationary spatial covariance structure in the context of spatial-temporal applications was based on decompositions of spatial processes in terms of empirical orthogonal functions (EOFs). The original methodology in this field has received renewed attention recently in the work of Nychka and colleagues (Nychka and Saltzman, 1998; Holland et al., 1998; Nychka et al., 2002). Briefly, considering the same spatial-temporal notation as above, the $n \times n$ empirical covariance matrix $\mathbf{C}$ may be written with a spectral decomposition as

$$\mathbf{S} = \mathbf{F}^T \Lambda \mathbf{F} = \sum_{k=1}^{n_T} \lambda_k \mathbf{F}_k \mathbf{F}_k^T \tag{9.5}$$

where $n_T = \min(n, T)$. The extension of this finite decomposition to the continuous spatial case represents the spatial covariance function as

$$C(\mathbf{x}, \mathbf{y}) = \sum_{k=1}^{\infty} \lambda_k F_k(\mathbf{x}) F_k(\mathbf{y}) \tag{9.6}$$

where the eigenfunctions $F_k(\mathbf{x})$ represent solutions to the Fredholm integral equation and correspond to the Karhunen–Loève decomposition of the (mean-centered) field as

$$Y(\mathbf{x}, t) = \sum_{k=1}^{\infty} A_k(t) F_k(\mathbf{x}). \tag{9.7}$$

The modeling and computational task here is in computing a numerical approximation to the Fredholm integral equation, or equivalently, choosing a set of generating functions $e_1(\mathbf{x}), \ldots, e_p(\mathbf{x})$ that are the basis for an extension of the finite eigenvectors $\mathbf{F}_k$ to eigenfunctions $F_k(\mathbf{x})$. (See Guttorp and Sampson (1994), Creutin and Obled (1982), Obled and Creutin (1986), and Preisendorfer (1988, Sec. 2d) for further details.)

In Holland et al. (1998), the spatial covariance function is represented as the sum of a conventional stationary isotropic spatial covariance model and a finite decomposition in terms of empirical orthogonal functions. This corresponds to a decomposition of the spatial process as a sum of a stationary isotropic process and a linear combination of M additional basis functions with random coefficients, the latter sum representing the deviation of the spatial structure from stationarity.

Nychka et al. (2002) introduced a multiresolution wavelet basis function decomposition with a computational focus on large problems with observations discretized to the nodes of a (large) $N \times M$ grid. The example application in this chapter is to air quality model output on a modest 48×48 grid. In the current notation, suppressing the temporal index, they write

$$Y(\mathbf{x}) = \sum_{k=1}^{NM} A_k F_k(\mathbf{x}). \tag{9.8}$$

In the discrete case, they write $\mathbf{F} = [F_{ki}]$, where $F_{ki} = F_k(\mathbf{x}_i)$, $\mathbf{x}_i$ being the ith grid point, so that one can write $\mathbf{Z} = \mathbf{FA}$ and $\mathbf{C} = \mathbf{F}\Sigma_A\mathbf{F}^T$. For the basis functions F_k, they use a "W" wavelet basis with parent forms that are piecewise quadratic splines that are *not* orthogonal or compactly supported. These were chosen because they can approximate the shape of common covariance models, such as the exponential, Gaussian and Matérn, depending on the specification (and off-diagonal sparcity) of the matrix Σ_A. Recent work (Matsuo, Nychka, and Paul, 2008) has extended the methodology to accommodate irregularly spaced monitoring data and a Monte Carlo expectation-maximization (EM) estimation procedure practical for large datasets. They analyze an ozone monitoring network dataset with 397 sites discretized (again) to a 48×48 grid.

Pintore, Holmes, and colleagues (Pintore and Holmes, 2004; Stephenson et al., 2005) work with both Karhunen–Loève and Fourier expansions. Nonstationarity is introduced by evolving the stationary spectrum over space in terms of a latent spatial power process. The resulting models are valid in terms of the original covariance function, but with local parameters. A Bayesian framework is used with MCMC estimation.

9.4 Process Convolution Models

Higdon (1998) introduced a process convolution approach for accommodating nonstationary spatial covariance structure. (See also Higdon, Swall, and Kern (1999).) The basic idea is to consider the fact that any stationary Gaussian process $Z(s)$ with correlogram $\rho(d) = \int_{\mathbb{R}^2} k(s)k(s-d)ds$ can be expressed as the convolution of a Gaussian white noise process $\zeta(s)$ with kernel $k(s)$

$$Y(\mathbf{s}) = \int_{\mathbb{R}^2} k(\mathbf{s}-\mathbf{u})\zeta(\mathbf{u})d\mathbf{u}. \tag{9.9}$$

A particular case of interest is the choice of bivariate Gaussian density functions with 2×2 covariance matrix Σ for the kernel, which results in processes with stationary anisotropic Gaussian correlation functions with the principal axes of Σ determining the directions of the anisotropic structure.

To account for nonstationarity, Higdon (1998) and Higdon et al. (1999) let the kernel vary smoothly with spatial location. Letting $k_s(\cdot)$ denote a kernel centered at the point s, with a shape depending on s, the correlation between two points $\mathbf{s}$ and $\mathbf{s}'$ is

$$\rho(\mathbf{s}, \mathbf{s}') = \int_{\mathbb{R}^2} k_{\mathbf{s}}(\mathbf{u}) k_{\mathbf{s}'}(\mathbf{u}) d\mathbf{u}. \tag{9.10}$$

Higdon et al. (1999) demonstrate the particular case where the $k_s(\cdot)$ are bivariate Gaussian densities characterized by the shape of ellipses underlying the 2×2 covariance matrices. The kernels are constrained to evolve smoothly in space by estimating the local ellipses under a Bayesian paradigm that specifies a prior distribution on the parameters of the ellipse (the relative location of the foci) as a Gaussian random field with a smooth (in fact, Gaussian) spatial covariance function. It should be noted that the form of the kernel determines the shape of the local spatial correlation function, with a Gaussian kernel corresponding to a Gaussian covariance function. Other choices of kernels can lead to approximations of other common spatial correlation functions.

Paciorek and Schervish (2006) extend this approach and create a class of closed-form nonstationary covariance functions, including a nonstationary Matérn covariance parameterized by spatially varying covariance parameters in terms of an eigen-decomposition of the kernel covariance matrix $k_s(\cdot)$.

Calder and Cressie (2007) discuss a number of topics associated with convolution-based modeling including the computational challenges of large datasets. Calder (2007, 2008) extends the approach to dynamic process convolutions for multivariate space–time monitoring data.

D'Hondt et al. (2007) apply the process convolution model with Gaussian kernels (which they call a nonstationary anisotropic Gaussian kernel (AGK) model) to the nonstationary anisotropic texture in synthetic aperture radar (SAR) images. The Gaussian kernels are estimated locally, in contrast to the Bayesian smoothing methods of Higdon and Paciorek and Schervish.

9.5 Spatial Deformation Models

The spatial deformation approach to modeling nonstationary or nonhomogeneous spatial covariance structures has been considered by a number of authors since the early work represented in Sampson and Guttorp (1992) and Guttorp and Sampson (1994). We first review the modeling approach, as presented by Meiring et al. (1997). We will then review some of the other work on this methodology, focusing on recently introduced Bayesian methods.

Suppose that temporally independent samples $Y_{it} = Y(\mathbf{x}_i, t)$ are available at N sites $\{\mathbf{x}_i, i = 1, \ldots, N$, typically in $R^2\}$ and at T points in time $\{t = 1, \ldots, T\}$. $\mathbf{X} = [\underline{X}_1\ \underline{X}_2]$ represents the matrix of geographic locations. We now write the underlying spatial-temporal process as

$$Y(\mathbf{x}, t) = \mu(\mathbf{x}, t) + \nu(\mathbf{x})^{1/2} E_t(\mathbf{x}) + E_\varepsilon(\mathbf{x}, t), \tag{9.11}$$

where $\mu(\mathbf{x}, t)$ is the mean field, and $E_t(\mathbf{x})$ is a zero mean, variance one, continuous second-order spatial Gaussian process, i.e., $\text{Cov}(E_t(\mathbf{x}), E_t(\mathbf{y})) \rightarrow [\mathbf{x} \geq \mathbf{y}]1$.

The correlation structure of the spatial process is expressed as a function of Euclidean distances between site locations in a bijective transformation of the geographic coordinate system

$$\mathrm{cor}(E_t(\mathbf{x}), E_t(\mathbf{y})) = \rho_\theta(\|f(\mathbf{x}) - f(\mathbf{y})\|), \tag{9.12}$$

where $f(\cdot)$ is a transformation that expresses the spatial nonstationarity and anisotropy, ρ_θ belongs to a parametric family with unknown parameters θ, $\nu(\mathbf{x})$ is a smooth function representing spatial variance, and $E_\varepsilon(\mathbf{x}, t)$ represents measurement error and/or very short scale spatial structure, assumed Gaussian and independent of E_t. For mappings from R^2 to R^2, the geographic coordinate system has been called the "G-plane" and the space representing the images of these coordinates under the mapping is called the "D-plane," Perrin and Meiring (1999) prove that this spatial deformation model is identifiable for mappings from R^k to R^k assuming only differentiability of the isotropic correlation function $\rho_\theta()$. Perrin and Senoussi (2000) derive analytic forms for the mappings $f(\cdot)$ under differentiability assumptions on the correlation structure for both the model considered here, where $\rho_\theta()$ is considered to be a stationary and isotropic correlation function ("stationary and isotropic reducibility"), and for the case where this correlation function is stationary, but not necessarily isotropic ("stationary reducibility").

Mardia and Goodall (1992) were the first to propose likelihood estimation and an extension to modeling of multivariate spatial fields (multiple air quality parameters) assuming a Kronecker structure for the space × species covariance structure. Likelihood estimation and an alternative radial basis function approach to representation of spatial deformations was proposed by Richard Smith in an unpublished report in 1996.

Meiring et al. (1997) fit the spatial deformation model to the empirically observed correlations among a set of monitoring sites by numerical optimization of a weighted least squares criterion constrained by a smoothness penalty on the deformation computed as a thin-plate spline. The problem is formulated so that the optimization is with respect to the parameters, θ, of the isotropic correlation model and the coordinates of the monitoring sites, $\xi_i = f(\mathbf{x}_i)$, in the deformation of the coordinate system. This is a large and often difficult optimization problem. It becomes excessively taxing when uncertainty in the estimated model is assessed by resampling methods or cross-validation. However, it is the approach that is implemented in the most conveniently available software for fitting the deformation model. These are the EnviRo.stat R programs that accompany the text by Le and Zidek (2006) on the analysis of environmental space–time processes (http://enviro.stat.ubc.ca/).

Iovleff and Perrin (2004) implemented a simulated annealing algorithm for fitting the spatial deformation model by optimization, with respect to correlation function parameters θ and D-plane coordinates of the monitoring sites, $\xi_i = f(\mathbf{x}_i)$, of a least squares criterion of goodness-of-fit to an empirical sample covariance matrix. Rather than impose an analytic smoothness constraint on the mapping (such as the thin-plate, spline-based, bending energy penalty of Meiring et al. (1997)), they use a Delaunay triangulation of the monitoring sites to impose constraints on the random perturbations of the D-plane coordinates ξ_i that guarantee that the resulting mapping $f(\mathbf{x}_i)$ is indeed bijective, i.e., it does not "fold." Using any of the other methods discussed here, the achievement of bijective mappings has relied on appropriate tuning of a smoothness penalty or prior probability model for the family of deformations.

Damian et al. (2001, 2003) and Schmidt and O'Hagan (2003) independently proposed similar Bayesian modeling approaches for inference concerning this type of spatial deformation model and for subsequent spatial estimation accounting for uncertainty in the estimation of the spatial deformation model underlying the spatial covariance structure. We present here details of the model of Damian et al. (2001, 2003).

For a Gaussian process with constant mean, $\mu(\mathbf{x}, t) \equiv \mu$, and assuming a flat prior for μ, the marginal likelihood for the covariance matrix Σ has the Wishart form

$$f(\{y_{it}|\Sigma\}) = |2\pi\Sigma|^{-(T-1)/2} \exp\left\{-\frac{T}{2} tr\,\Sigma^{-1}\mathbf{C}\right\} \tag{9.13}$$

where $\mathbf{C}$ is the sample covariance with elements,

$$c_{ij} = \frac{1}{T}\sum_{t=1}^{T}(y_{it} - \bar{y}_i)(y_{jt} - \bar{y}_j), \tag{9.14}$$

and the true covariance matrix is parameterized as $\Sigma = \Sigma(\theta, \nu_i, \xi_i)$, with $\Sigma_{ij} = (\nu_i \nu_j)^{1/2} \rho_\theta(\|\xi_i - \xi_j\|)$, and $\xi_i = f(\mathbf{x}_i)$. The parameters to be estimated are $\{\theta, \nu_i, \xi_i; i = 1, \ldots, N\}$.

The Bayesian approach requires a prior on all of these parameters. The novel and challenging aspect of the problem concerns the prior for the spatial configuration of the ξ_i. Writing the matrix $\Xi = [\xi_1, \ldots, \xi_N]^T = [\,\Xi_1\ \Xi_2\,]$, Damian et al. (2001, 2003) use a prior of the form

$$\pi(\Xi) \propto \exp\left\{-\frac{1}{2\tau^2}\left[\Xi_1^T \mathbf{K} \Xi_1 + \Xi_2^T \mathbf{K} \Xi_2\right]\right\} \tag{9.15}$$

where $\mathbf{K}$ is a function of the geographic coordinates only—the "bending energy matrix" of a thin-plate spline (see Bookstein, 1989)—and τ is a scale parameter penalizing "nonsmoothness" of the transformation f. Mardia, Kent, and Walder (1991) first used a prior of this form in the context of a deformable template problem in image analysis. It should be noted that the bending energy matrix $\mathbf{K}$ is of rank $n - 3$ and the quadratic forms in the exponent of this prior are zero for all affine transformations, so that the prior is flat over the space of all affine deformations and thus is improper.

The parameter space is highly multidimensional and the posterior distributions are not of closed form, therefore, a Metropolis–Hastings algorithm was implemented to sample from the posterior. (See Damian et al. (2001) for details of the MCMC estimation scheme.) Once estimates for the new locations have been obtained, the transformation is extrapolated to the whole area of interest using a pair of thin-plate splines.

Schmidt and O'Hagan (2003) work with the same Gaussian likelihood, but utilize a general Gaussian process prior for the deformation. When considered in terms of the coordinates ξ_i, the effect of this on the form of the prior $\pi(\Xi)$ is to center the coordinate vectors Ξ_j, $j = 1, 2$, at their geographic locations and to replace $\mathbf{K}$ with a full rank covariance matrix of a form to be specified. Utilizing the known interpretation of thin-plate splines as kriging for an intrinsic random function with a particular form of (generalized) covariance matrix, we see that the Damian et al. (2001) approach may be considered similarly to correspond to a prior for the deformation considered as an intrinsic random function. Schmidt and O'Hagan (2003) also differ from Damian et al. (2001) in their choice of parametric isotropic correlation models and in many of the details of the MCMC estimation scheme, but they are otherwise similarly designed methods.

The atmospheric science literature includes a number of papers with deformation models motivated or determined explicitly by physical processes. (See, for example, Riishojgaard (1998) and Fu et al. (2004).) Xiong et al. (2007) implement a nonlinear mapping model for nonstationary covariance-based kriging in a high-dimensional ($p = 19$) metamodeling problem using computer simulation data.

Anderes and Stein (2008) are the first authors to address the application of the deformation model to the case of a single realization of a spatial process obtained as the deformation of an isotropic Gaussian random field. They present a complete mathematical analysis and methodology for observations from a dense network with approximate likelihood computations derived from partitioning the observations into neighborhoods and assuming independence of the process across partitions.

9.6 Discussion

There is clearly much active work on the development and application of models for nonstationary spatial processes in an expanding range of fields beyond the atmospheric science and environmental applications that motivated most of the early work in this field. We have seen novel applications in image analysis (D'Hondt et al., 2007) and "metamodeling in engineering design" (Xiong et al., 2007). It appears unlikely that there will prove to be one "best" approach for all applications from among the major classes reviewed here: kernel smoothing, process convolution models, spectral and basis functions models, and deformation models.

Although this chapter covers substantial literature, the recent methodologies are still not mature in a number of respects. First, most of the approaches reviewed here are not easily applied as the developers of these methods have, for the most part, not made software available for use by other investigators. A number of questions of practical importance remain to be addressed adequately through analysis and application. Most of the literature reviewed above addresses the application of the fitted spatial covariance models to problems of spatial estimation, as in kriging. The Bayesian methods, all propose to account for the uncertainty in the estimation of the spatial covariance structure, but the practical effects of this uncertainty have not yet been demonstrated. There remains a need for further development of diagnostic methods and experience in diagnosing the fit of these alternative models. In particular, the nature of the nonstationarity, or equivalently, the specification or estimation of the appropriate degree of spatial smoothness in these models expressed in prior distributions or regularization parameters, needs further work. For the Bayesian methods, this translates into a need for further understanding and/or calibration of prior distributions.

This chapter has focused on nonparametric approaches to the modeling of nonstationary spatial covariance structure for univariate spatial processes. In some cases one may wish to formally test the hypothesis of nonstationarity (Fuentes, 2005; Corstanje et al., 2008). Mardia and Goodall (1992), Gelfand et al. (2004), and Calder (2007, 2008) address multivariate problems that are addressed in further detail in Chapter 28. Some parametric models have also been introduced. These include parametric approaches to the spatial deformation model, including Perrin and Monestiez' (1998) parametric radial basis function approach to the representation of two-dimensional deformations. Parametric models appropriate for the characterization of certain point source effects have been introduced by Hughes-Oliver et al. (1998, 1999, 2009).

References

Anderes, E.B. and Stein, M.L. (2008). Estimating deformations of isotropic Gaussian random fields on the plane. *Annals of Statistics*, **36**, 719–741.

Bookstein, F.L. (1989). Principal warps – Thin-plate splines and the decomposition of deformations. *IEEE Transactions on Pattern Analysis*, **11**, 567–585.

Buell, E.C. (1972). Integral equation representation for factor analysis. *Journal of Atmospheric Science*, **28**, 1502–1505.

Buell, E.C. (1978). The number of significant proper functions of two-dimensional fields. *Journal of Applied Meteorology*, **17**, 717–722.

Buttafuoco, G. and Castirgnanò, A. (2005). Study of the spatio-temporal variation of soil moisture under forest using intrinsic random functions of order k. *Geoderma*, **128**, 208–220.

Calder, C.A. (2007). Dynamic factor process convolution models for multivariate space-time data with application to air quality assessment. *Environmental and Ecological Statistics*, **14**, 229–247.

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

Calder, C.A. and Cressie, N. (2007). Some topics in convolution-based spatial modeling. In: *Proceedings of the 56th Session of the International Statistical Institute*, Lisboa. International Statistical Institute. Voorburg, the Netherlands.

Christensen, R. (1990). The equivalence of predictions from universal kriging and intrinsic random function kriging. *Mathematical Geology*, **22**, 655–664.

Cohen, A., and Jones, R.H. (1969). Regression on a random field. *Journal of the American Statistical Association*, **64**, 1172–1182.

Corstanje, R., Grunwald, S., and Lark, R.M. (2008). Inferences from fluctuations in the local variogram about the assumption of stationarity in the variance. *Geoderma*, **143**, 123–132.

Creutin, J.D. and Obled, C. (1982). Objective analyses and mapping techniques for rainfall fields: An objective comparison. *Water Resources Research*, **18**, 413–431.

Damian, D., Sampson, P.D., and Guttorp, P. (2001). Bayesian estimation of semi-parametric non-stationary spatial covariance structures. *Environmetrics*, **12**, 161–178.

Damian, D., Sampson, P.S., and Guttorp, P. (2003). Variance modeling for nonstationary spatial processes with temporal replicates. *Journal of Geophysical Research-Atmospheres*, **108**, D24, 8778.

D'Hondt, O., Lopez-Martiez, C., Ferro-Famil, L., and Pottier, E. (2007). Spatially nonstationary anisotropic texture analysis in SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, **45**, 3905–3918.

Fu, W.W., Zhou, G.Q., and Wang, H.J. (2004). Ocean data assimilation with background error covariance derived from OGCM outputs. *Advances in Atmospheric Sciences*, **21**, 181–192.

Fuentes, M. (2001). A new high frequency kriging approach for nonstationary environmental processes. *Environmetrics*, **12**, 469–483.

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

Fuentes, M. (2002b). Interpolation of nonstationary air pollution processes: A spatial spectral approach. *Statistical Modelling*, **2**, 281–298.

Fuentes, M. (2005). A formal test for nonstationarity of spatial stochastic processes. *Journal of Multivariate Analysis*, **96**, 30–54.

Fuentes, M. and Smith, R.L. (2001). A new class of nonstationary spatial models. Tech report, North Carolina State Univ. Institute of Statistics Mimeo Series #2534, Raleigh.

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

Guillot, G., Senoussi, R., and Monestiez, P. (2001). A positive definite estimator of the non stationary covariance of random fields. In: *GeoENV 2000: Third European Conference on Geostatistics for Environmental Applications*, P. Monestiez, D. Allard, and R. Froidevaux, Eds., Kluwer, Dordrecht, the Netherlands, 333–344.

Guttorp, P., and Sampson, P.D. (1994). Methods for estimating heterogeneous spatial covariance functions with environmental applications. In: *Handbook of Statistics*, vol. 12, G.P. Patil and C.R. Rao, Eds., Elsevier Science, New York, pp. 661–689.

Haas, T.C. (1990a). Kriging and automated variogram modeling within a moving window. *Atmospheric Environment*, **24A**, 1759–1769.

Haas, T.C. (1990b). Lognormal and moving window methods of estimating acid deposition. *Journal of the American Statistical Association*, **85**, 950–963.

Haas, T.C. (1995). Local prediction of a spatio-temporal process with an application to wet sulfate deposition. *Journal of the American Statistical Association*, **90**, 1189–1199.

Haslett, J. and Raftery, A.E. (1989). Space-time modelling with long-memory dependence: Assessing Ireland's wind resource (with discussion). *Journal of the Royal Statistical Society, Ser. C*, **38**, 1–50.

Higdon, D. (1998). A process-convolution approach to modeling temperatures in the North Atlantic Ocean, *Journal of Environmental and Ecological Statistics*, **5**, 173–190.

Higdon, D.M., Swall, J., and Kern, J. (1999). Non-stationary spatial modeling. In *Bayesian Statistics 6*, J.M. Bernardo, J.O. Berger, A.P. David, and A.F.M. Smith, Eds., Oxford University Press, Oxford, U.K., pp. 761–768.

Holland, D., Saltzman, N., Cox, L., and Nychka, D. (1998). Spatial prediction of dulfur dioxide in the eastern United States. In: *GeoENV II: Geostatistics for Environmental Applications*. J. Gomez-Hernandez, A. Soares, and R. Froidevaux, Eds. Kluwer, Dordrecht, the Netherlands, pp. 65–76.

Hughes-Oliver, J.M. and Gonzalez-Farías, G. (1999). Parametric covariance models for shock-induced stochastic processes. *Journal of Statistical Planning and Inference*, **77**, 51–72.

Hughes-Oliver, J.M., Gonzalez-Farías, G., Lu, J.C., and Chen, D. (1998). Parametric nonstationary spatial correlation models. *Statistics and Probability Letters*, **40**, 267–278.

Hughes-Oliver, J.M., Heo, T.Y., and Ghosh, S.K. (2009). An auto regressive point source model for spatial processes. *Environmetrics*, **20**, 575–594.

Iovleff, S. and Perrin, O. (2004). Estimating a non-stationary spatial structure using simulated annealing. *Journal of Computational and Graphical Statistics*, **13**, 90–105.

Kim, H.M., Mallick, B.K., and Holmes, C.C. (2005). Analyzing nonstationary spatial data using piecewise Gaussian processes. *Journal of the American Statistical Association*, **100**, 653–658.

Le, N.D. and Zidek, J.V. (2006). *Statistical Analysis of Environmental Space-Time Processes*. Springer, New York.

Loader, C., and Switzer, P. (1992). Spatial covariance estimation for monitoring data. In: *Statistics in Environmental and Earth Sciences*, A. Walden and P. Guttorp, Eds., Edward Arnold, London, pp. 52–70.

Mardia, K.V., and Goodall, C.R. (1992). Spatial-temporal analysis of multivariate environmental monitoring data. In: *Multivariate Environmental Statistics 6*. N.K. Bose, G.P. Patil, and C.R. Rao, Eds., North Holland, New York, pp. 347–385.

Mardia, K.V., Kent, J.T., and Walder, A.N. (1991). Statistical shape models in image analysis. In: *Computing Science and Statistics: Proceedings of the 23rd Symposium on the Interface*. E.M. Keramidas, ed., Interface Foundation of America, Fairfax, VA., pp. 550–557.

Matsuo, T., Nychka, D., and Paul, D. (2008). Nonstationary covariance modeling for incomplete data: Monte Carlo EM approach. Forthcoming. See http://www.image.ucar.edu/~nycka/man.html

Meiring, W., Monestiez, P., Sampson, P.D., and Guttorp, P. (1997). Developments in the modelling of nonstationary spatial covariance structure from space-time monitoring data. In: *Geostatistics Wallongong '96*, E.Y. Baafi and N. Schofield, eds., Kluwer, Dordrecht, the Netherland, pp. 162–173.

Nott, D.J. and Dunsmuir, W.T.M. (2002). Estimation of nonstationary spatial covariance structure. *Biometrika*, **89**, 819–829.

Nychka, D. and Saltzman, N. (1998). Design of air quality networks. In: *Case Studies in Environmental Statistics*. D. Nychka, W. Piegorsch, and L. Cox, eds. Springer-Verlag, New York, pp. 51–76.

Nychka, D., Wikle, C., and Royle, J.A. (2002). Multiresolution models for nonstationary spatial covariance functions. *Statistical Modelling*, **2**, 315–331.

Obled, Ch. and Creutin, J.D. (1986). Some developments in the use of empirical orthogonal functions for mapping meteorological fields. *Journal of Applied Meteorology*, **25**, 1189–1204.

Oehlert, G.W. (1993). Regional trends in sulfate wet deposition. *Journal of the American Statistical Association*, **88**, 390–399.

Paciorek, C.J. and Schervish, M.J. (2006). Spatial modelling using a new class of nonstationary covariance functions. *Environmetrics*, **17**, 483–506.

Perrin, O. and Meiring, W. (1999). Identifiability for non-stationary spatial structure. *Journal of Applied Probability*, **36**, 1244–1250.

Perrin, O. and Monestiez, P. (1998). Modeling of non-stationary spatial covariance structure by parametric radial basis deformations. In: *GeoENV II: Geostatistics for Environmental Applications*. J. Gomez-Hernandez, A. Soares, and R. Froidevaux, eds. Kluwer, Dordrecht, the Netherland, pp. 175–186.

Perrin, O. and Senoussi, R. (2000). Reducing non-stationary random fields to stationarity and isotropy using a space deformation. *Statistics and Probability Letters*, **48**, 23–32.

Pintore, A. and Holmes, C. (2004). Spatially adaptive non-stationary covariance functions via spatially adaptive spectra. Forthcoming. (Submitted manuscript available from http://www.stats.ox.ac.uk/~cholmes/).

Preisendorfer, R.W. (1988). *Principal Component Analysis in Meteorology and Oceanography*. Elsevier, Amsterdam.

Riishojgaard, L.P. (1998). A direct way of specifying flow-dependent background error correlations for meteorological analysis systems. *Tellus Series A-Dynamic Meteorology and Oceanography*, **50**, 42–57.

Sampson, P.D. (2001). Spatial covariance. In El-Shaarawi, A.H., Pierorcsh W.W., Eds. *Encyclopedia of Environmetrics*. New York: John Wiley & Sons, 2002, vol. 4, pp. 2059–2067.

Sampson, P.D., and Guttorp, P. (1992). Nonparametric estimation of nonstationary spatial covariance structure. *Journal of the American Statistical Association*, **87**, 108–119.

Schmidt, A. and O'Hagan, A. (2003). Bayesian inference for nonstationary spatial covariance structure via spatial deformations. *Journal of the Royal Statistical Society*, Series B, **65**, 743–758.

Smith, R.H. (1996). Estimating nonstationary spatial correlations. Preprint (http://www.stat.unc.edu/faculty/rs/papers/RLS_Papers.html).

Stein, M.L. (2001). Local stationarity and simulation of self-affine intrinsic random functions. *IEEE Transactions on Information Theory*, **47**, 1385–1390.

Stephenson, J., Holmes, C., Gallagher, K. and Pintore, A. (2005). A statistical technique for modelling non-stationary spatial processes. In: *Geostatistics Banff 2004*, O. Leuangthong and C.V. Deutsch, Eds., pp. 125–134.

Wikle, C.K., and Cressie, N. (1999). A dimension-reduced approach to space-time Kalman filtering. *Biometrika*, **86**, 815–829.

Xiong, Y., Chen, W., Apley, D., and Ding, X.R. (2007). A non-stationary covariance-based kriging method for metamodelling in engineering design. *International Journal for Numerical Methods in Engineering*, **71**, 733–756.