Advances in Modeling and Inference for Environmental Processes with Nonstationary Spatial Covariance

Paul Sampson Doris Damian Peter Guttorp

NRCSE

Technical Report Series

NRCSE-TRS No. 061

January 18, 2001

The NRCSE was established in 1997 through a cooperative agreement with the United States Environmental Protection Agency which provides the Center's primary funding.

ADVANCES IN MODELING AND INFERENCE FOR ENVIRONMENTAL PROCESSES WITH NONSTATIONARY SPATIAL COVARIANCE

P.D. SAMPSON, D. DAMIAN AND P. GUTTORP
National Research Center for Statistics and the Environment
University of Washington
Box 354323
Seattle, WA 98195-4323
U.S.A.

1. INTRODUCTION

Modeling of the spatial dependence structure of environmental processes is fundamental to almost all statistical analyses of data that are sampled spatially. These analyses address tasks such as spatial estimation (kriging) and monitoring network design, as well as the basic scientific characterization of the second order properties of these processes. Prior to 1990, the lack of general models for the spatial covariance function led to almost exclusive reliance on stationary models of the form $\mathrm{cov}(Z(x), Z(y)) = C(x-y)$ where $\{Z(x) : x \in D\}$ denotes a process defined over a spatial domain $D \subset R^d$. However, it is now widely recognized that most, if not all, spatio-temporal environmental processes (and many spatial processes without a temporal aspect) manifest spatially nonstationary or heterogeneous covariance structure when considered over a sufficiently large spatial range.

There is a rapidly growing body of literature on methods for modeling nonstationary spatial covariance structure. The majority of the literature concerns methods that are semi-parametric: they are nonparametric with respect to the way that spatial variation in covariance structure is described, but the local covariance structure is described by conventional parametric models. Much of this recent literature discusses Bayesian modeling strategies which enable the uncertainty in the estimated spatial covariance structure to be reflected in spatial estimation. Many of the methods discussed here concern models assuming temporally independent replications (possibly after preprocessing of space-time monitoring data). Temporal correlation, and more generally space-time covariance structure, is important in many applications, depending on the magnitude of the time step relevant for analysis. See Gneiting and Schlather (2001) for discussion of various separable and non-separable space-time models, mostly stationary, and Wikle and Cressie (1999) for other approaches to dynamic modeling of spatially varying space-time structure.

Let $Z(x,t)$ represent a space-time random field modeled as

$$Z(x,t) = \mu(x,t) + E(x,t) + \varepsilon(x,t), \tag{1}$$

where $\mu(x,t)$ represents a spatio-temporal mean field or trend, $\varepsilon(x,t)$ represents measurement error and small-scale spatial variability, and $E(x,t)$ is a mean zero space-time process which is L_2-continuous in space and independent of $\varepsilon(x,t)$. For the purposes of this article we will assume that t indexes T temporally independent replications of the spatio-temporal field and our focus is on the spatial covariances, $\mathrm{cov}(Z(x,t), Z(y,t))$. In the case of direct measurements of atmospheric processes, such as temperatures, winds, pressure, and rainfall, or in the case of many atmospherically driven pollutant processes, spatial covariances are affected by spatially varying features of landscape, topography and/or circulation patterns, so that they are cannot be expressed as simple homogeneous functions of the vector separation $x - y$.

A fundamental notion underlying most of the current modeling approaches is that the spatial correlation structure of environmental processes such as these 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 none of the methods reviewed here attempts such explicit modeling.

Prior to 1990 the main (possibly only) approach to modeling or characterizing the nonstationary spatial covariance structure of spatio-temporal environmental monitoring data (outside of local analyses in subregions where the process might be more nearly stationary) was based on 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 back at least to Cohen and Jones (1969) and Buell (1978), although perhaps the most useful elaboration of the method for spatial analysis was provided by Obled and Creutin (1986). A number of new modeling and inference methods were introduced in the late 1980's and early 1990's, beginning with Guttorp and Sampson's spatial deformation approach, first mentioned in print in a 1989 comment on a paper by Haslett and Raftery (1989). Shortly following were: Haas' "moving window" spatial estimation (Haas 1990a,b, 1995), although this approach estimates covariance structure locally without providing a global model; Sampson and Guttorp's (1992) elaboration of their first approach to the spatial deformation model based on multidimensional scaling; an empirical Bayes shrinkage approach of Loader and Switzer (1992); and Oehlert's kernel smoothing approach (1993). Guttorp and Sampson (1994) reviewed this early literature on methods for estimating heterogeneous spatial covariance functions with comments on further extensions of the spatial deformation method. In the current article we focus on the developments of the late 1990s as these largely subsume the earlier proposals. At the time of this writing, a number of the methods reviewed here are not yet published.

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

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 incorporates any other explicit modeling of the spatial heterogeneity in the spatial covariance structure. As noted above, Haas' (1990a,b, 1995) moving window approach 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. See Guttorp and Sampson (1994) for a brief summary and Nott and Dunsmuir (1998) for further assessment.

2.1. SPATIALLY SMOOTHED LOCAL MODELS

Recent manuscripts by Fuentes (2000) and by Nott and Dunsmuir (1998) propose somewhat 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 K_{θ_i}. Fuentes (2000) represents the spatial process $Z(x)$ as a weighted average of "orthogonal local stationary processes":

$$Z(x) = \sum_{i=1}^{k} w_i(x) Z_i(x) \tag{2}$$

where $w_i(x)$ is a chosen weight function such as inverse squared distance between x and the center of subregion S_i and $Z_i(x)$ denotes a spatial process (defined over the entire region) with covariance function K_{θ_i}. The nonstationary spatial covariance structure is given by

$$\begin{aligned} \operatorname{cov}(Z(x), Z(y)) &= \textstyle\sum_{i=1}^{k} w_i(x) w_i(y) \operatorname{cov}(Z_i(x), Z_i(y)) \\ &= \textstyle\sum_{i=1}^{k} w_i(x) w_i(y) K_{\theta_i}(x-y) \end{aligned} \tag{3}$$

Fuentes proposes to choose the number of subgrids, k, using a BIC criterion. The stationary processes $Z_i(x)$ are actually "local" only in the sense that their corresponding covariance functions, $K_{\theta_i}(x-y)$, 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 in the context of a complete Bayesian spatial estimation with predictive distributions accounting for uncertainty in the parameter estimates.

Work in progress by Fuentes and Smith proposes to extend the finite decomposition of $Z(x)$ of Fuentes (2000) to a continuous convolution of local stationary processes:

$$Z(x) = \int_D w(x-s) Z_{\theta(s)}(x) ds \tag{4}$$

Estimation will require that the spatial field of parameter vectors $\theta(s)$, indexing the stationary Gaussian processes, be constrained to vary smoothly.

Nott and Dunsmuir's (1998) approach, proposed as a more computationally feasible alternative to the spatial deformation model of Sampson and Guttorp, has the stated aim of reproducing an empirical covariance matrix at a set of monitoring sites and 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 $K_{\theta_i}(x-y)$ 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 as

$$\operatorname{cov}(Z(x), Z(y)) = \Sigma_0(x,y) + \textstyle\sum_{i=1}^{k} w_i(x) w_i(y) K_{\theta_i}(x-y) \tag{5}$$

where $\Sigma_0(x, y)$ is a function of the empirical covariance matrix at the monitoring sites, $\mathbf{S} = [s_{ij}]$, and the local stationary models computed so that $\mathrm{cov}(Z(x_i), Z(x_j)) = s_{ij}$. This exact interpolation is relaxed by replacing the empirical covariance matrix $\mathbf{S}$ by the Loader and Switzer (1992) empirical Bayes shrinkage estimator $\hat{\mathbf{S}} = \gamma\mathbf{S} + (1-\gamma)\mathbf{C}$, where $\mathbf{C}$ is a covariance matrix obtained by fitting some parametric covariance function model.

While the models introduced by Fuentes and by Nott and Dunsmuir look similar, the details are substantially different. Nott and Dunsmuir use hypothetical conditional processes and assume an empirical covariance matrix computed from spatio-temporal data. Fuentes' method uses unconditional processes and applies as well to purely spatial data. It involves a complete Bayesian analysis without resort to computationally intensive MCMC methods. Neither of these two methods has involved fitting of locally anisotropic covariance models, but doing so should not greatly complicate the computations.

While convenient for accommodating nonstationary covariance structure, the underlying spatial process models of both of these approaches—the decomposition in terms of orthogonal stationary processes with spatially varying local covariance structure—do not seem (to us) to represent usefully interpretable scientific models. Furthermore, certain key elements of the approach, such as the number of locally stationary component models (for Fuentes), or the size of the neighborhoods for fitting of the local models, and the nature of the weight or kernel must be determined by somewhat ad hoc means.

2.2. KERNEL SMOOTHING OF EMPIRICAL COVARIANCE MATRICES

Guillot et al. (2001) have proposed a different type of kernel estimator, similar to one introduced by Oehlert (1992), although they do not refer to this earlier work. Let D denote the spatial domain, so that the covariance function $S(x, y)$ is defined on $D \times D$, and suppose that an empirical covariance matrix $\mathbf{S} = [s_{ij}]$ is computed for sites $\{x_i, i = 1, \ldots, n\}$. Define a non-negative 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 tesselation). The nonparametric, nonstationary estimator of S, is

$$\hat{S}_\varepsilon(x, y) = \sum_{i,j} s_{ij} \int_{D_i \times D_j} K_\varepsilon(x-u, y-v)\,du\,dv \qquad (6)$$

The authors prove positive-definiteness of the estimator for positive-definite kernels, which include step functions, Gaussian, and exponential forms. They consider various approaches to the selection of the bandwidth parameter ε, including plug-in minimization of mean integrated squared error in estimation of $S(x, y)$ and two forms of cross-validation. The local behavior of the resulting models is not discussed. They demonstrate an application to West Africa rainfall event data where, surprisingly, kriging with the nonstationary covariance model is outperformed by kriging with a fitted stationary model.

3. BASIS-FUNCTION MODELS

As noted in our introduction, the earliest modeling strategy for nonstationary spatial covariance structure was based on decompositions of spatial processes in terms of empirical orthogonal functions (EOF's). The original methodology in this field, as reviewed by Guttorp and Sampson (1994), has received renewed attention recently in the work of Nychka and colleagues (Nychka and Saltzman 1998, Holland et al. 1998, Nychka et al. 2000). Briefly, using the same spatio-temporal notation as above, the $n \times n$ empirical covariance matrix $\mathbf{S}$ may be written with a spectral decomposition as

$$\mathbf{S} = \mathbf{F}'\mathbf{\Lambda}\mathbf{F} = \sum_{k=1}^{n_T} \lambda_k \mathbf{F}_k \mathbf{F}'_k, \tag{7}$$

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

$$S(x, y) = \sum_{k=1}^{\infty} \lambda_k F_k(x) F_k(y), \tag{8}$$

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

$$Z(x, t) = \sum_{k=1}^{\infty} A_k(t) F_k(x). \tag{9}$$

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 $e_1(x), \cdots, e_p(x)$ that are the basis for an extension of the finite eigenvectors $\mathbf{F}_k$ to eigenfunctions $F_k(x)$. See Guttorp and Sampson (1994), Obled and Creutin (1986), and Preisendorfer (1998, sect 2d) for further details.

In Nychka and Saltzman (1998) and 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. The authors demonstrate that model-based calculations of kriging predictive standard errors for an application to seasonally-adjusted SO_2 data in the eastern U.S. were substantially smaller for this nonstationary model than for a fitted stationary model and a fitted model with stationary spatial correlation but spatially varying variance field.

Nychka, et al. (2000) introduced an alternative wavelet basis function decomposition with a computational focus on large problems with observations discretized to the nodes of a (large) $N \times M$ grid. Their example is an analysis of monthly precipitation fields over a region of the Midwest and Rocky Mountains with observations at about 1600 stations discretized to a 128×128 grid. They use a "W" wavelet basis with parent forms that are piecewise quadratic splines, which are not orthogonal or compactly supported. These were chosen because they can approximate the shape of common covariance models such as the exponential and Gaussian, depending on the sequence of variances of the basis functions in the decomposition. Issues concerning the class of covariances spanned by these multiresolution bases remain to be studied.

EOF decompositions are also being used effectively to provide low-dimensional representations of the space-time structure in dynamic linear models; see Wikle and Cressie (1999) for details.

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 of the form $\rho(d) = \int_{R^2} k(s)k(s-d)ds$ can be expressed as the convolution of a Gaussian white noise process $\zeta(s)$ with kernel $k(s)$

$$Z(s) = \int_{R^2} k(s-u)\zeta(u)du. \tag{10}$$

A particular kernel of interest is the bivariate Gaussian density function with 2×2 covariance matrix $\mathbf{\Sigma}$, which results in processes with stationary anisotropic Gaussian correlation function with the principal axes of $\mathbf{\Sigma}$ 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 s and s' is

$$\rho(s, s') = \int_2 k_s(u) k_{s'}(u) du. \tag{11}$$

Higdon et al. (1999) demonstrate the particular case where the $k_s(\cdot)$ are bivariate Gaussian densities characterized by the shapes 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 to other common spatial correlation functions.

Figure 1, based on the presentation in Higdon et al. (1999), illustrates the nature of a fitted model for an analysis of the spatial distribution of dioxin concentrations in the Piazza Road pilot study area, which is part of an U.S. EPA Superfund site in Missouri. In this purely spatial example, dioxin was transported through a small stream channel which follows a curving path generally following the path of greatest concentration from top to bottom as indicated in the figure. The solid ellipses indicate the shape of the Gaussian kernels at sampling sites as given by the posterior distribution of the Bayesian analysis; the long axis of the ellipse indicates the direction of greater spatial correlation, which roughly parallels the direction of the stream channel. The dotted ellipses represent the spatially varying estimates of these local kernels on a regular grid, in accordance with the Gaussian random field prior for their parameters.

5. SPATIAL DEFORMATION MODELS

The spatial deformation approach to modeling nonstationary spatial covariance structure has been considered by a number of authors since the early work presented 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 $Z_{it} = Z(x_i, t)$ are available at N sites $x_i, i = 1, \ldots, N$, typically in R^2 and at T points in time $t = 1, \ldots, T$. $\mathbf{X} = [\ \underline{X}_1 \quad \underline{X}_2\ ]$ represents the matrix of geographic locations. We now write the underlying spatio-temporal process as

$$Z(x, t) = \mu(x, t) + \nu(x)^{1/2} E_t(x) + \varepsilon(x, t), \tag{12}$$

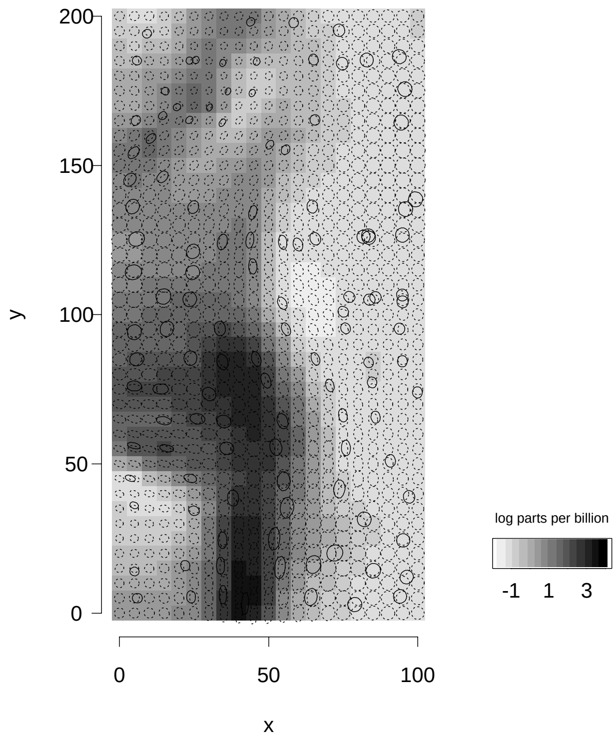

Figure 1. Estimated kernels of the process-convolution model for the Piazza Road data. Solid ellipses represent the kernels at the sampling sites and dotted ellipses the extension to a regular grid according to the random field prior model. The images show the corresponding posterior mean estimates for the dioxin concentrations. This figure was provided by Jenise Swall.

where $\mu(x,t)$ is the mean field, $\nu(x)$ is a smooth function representing spatial variance, and $E_t(x)$ is a zero mean, variance one, second–order continuous spatial Gaussian process, i.e. $\mathrm{cov}(E_t(x), E_t(y)) \to 1$ as $x \to y$. $\varepsilon(x,t)$ represents measurement error and/or very short scale spatial structure which is assumed Gaussian and independent of E_t.

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

$$\mathrm{cor}(E_t(x), E_t(y)) = \rho_\theta(\|f(x) - f(y)\|), \tag{13}$$

where $f(\cdot)$ is the 1:1 transformation that expresses the spatial nonstationarity and anisotropy, and ρ_θ belongs to a parametric family with unknown parameters θ.

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 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 (1998) 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 $\times$ parameter 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, $\xi_i = f(x_i)$, 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.

Iovleff and Perrin (2000) 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, $\xi_i = f(x_i)$, 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 $f(x_i)$ is indeed bijective; i.e., that 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. (2000) and Schmidt and O'Hagan (2000) have recently, and independently, proposed similar Bayesian modeling approaches for inference about 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. (2000).

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

$$L\left(\{z_{it}|\mathbf{\Sigma}\}\right) = |2\pi\mathbf{\Sigma}|^{-(T-1)/2} \exp\left\{-\frac{T}{2}\mathrm{tr}\mathbf{\Sigma}^{-1}\mathbf{S}\right\} \tag{14}$$

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

$$s_{ij} = \frac{1}{T}\sum_{t=1}^{T}\left(z_{it} - \bar{z}_i\right)\left(z_{jt} - \bar{z}_j\right) \tag{15}$$

and the true covariance matrix is parameterized as $\mathbf{\Sigma} = \Sigma\left(\theta, \nu_i, \xi_i\right)$, with $\Sigma_{ij} = \left(\nu_i\nu_j\right)^{1/2}\rho_\theta\left(\|\xi_i - \xi_j\|\right)$, and $\xi_i = f\left(x_i\right)$. 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 $\mathbf{\Xi} = [\xi_1, \ldots, \xi_N]' = \left[\begin{array}{cc} \underline{\Xi}_1 & \underline{\Xi}_2 \end{array}\right]$, Damian et al. (2000) use a prior of the form

$$\pi\left(\mathbf{\Xi}\right) \propto \exp\left\{-\frac{1}{2\tau^2}\left[(\underline{\Xi}_1 - \underline{X}_1)'\mathbf{K}(\underline{\Xi}_1 - \underline{X}_1) + (\underline{\Xi}_2 - \underline{X}_2)'\mathbf{K}(\underline{\Xi}_2 - \underline{X}_2)\right]\right\}, \tag{16}$$

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 "non-smoothness" 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 transformation, so that the prior is flat over the space of all affine deformations and is thus improper. $\mathbf{K}$ is consistent with the generalized covariance structure of an intrinsic random function corresponding to a thin-plate spline (Kent and Mardia, 1994).

The parameter space is high-dimensional 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. (2000) 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.

Damian et al. (2000) present an analysis of time series of 10-day aggregated rainfall data at 39 sites in the Languedoc-Roussillon region of southern France, data also analyzed by Meiring et al. (1997). Figure 2 presents one illustration from a reanalysis of a subset of 36 monitoring sites for which differences in site variances are modest, permitting a model with constant variance, $\nu(x) \equiv \nu$. The processes governing variation and covariation in measured rainfall are presumably influenced by the Pyrennées mountains to the extreme south-west, the Cevennes mountain range running parallel to the north-west boundary of the region, the Rhone valley to the east, and the Mediterranean border to the south-east. The figure derives from a subsample of 250 out of 125,000 realizations from the Metropolis-Hastings generated Markov chain (after convergence) representing the posterior distribution of D-plane coordinates ξ_i for the monitoring sites under the model described above. Figure 2(b) depicts the mean coordinates of this posterior distribution of the ξ_i, with a thin-plate spline interpolation used to map the square grid drawn in Figure 2(a). The thin-plate spline provides, in fact, the Bayes estimates of D-plane coordinates for unmonitored sites. Circular 0.90 equal-correlation contours centered on sites 38, 59, 61, and 74 are back-transformed into the contours shown in 1(a). These illustrate the fundamental notion of nonstationarity in spatial covariance structure with, in this case, strong anisotropy in the northeast-southwest direction parallel to the limit of the Cevennes mountain range, and relatively strong spatial correlation indicated by the contraction in the region just west of the mouth of the Rhone river.

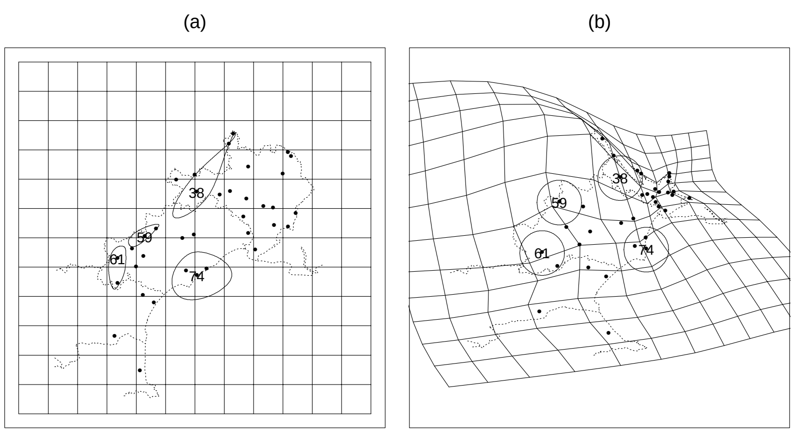

Figure 2. (a) Geographic map of 36 rainfall monitoring sites in the Languedoc-Roussillon region of southern France with 0.90 correlation contours around sites numbered 38, 59, 61, and 74, respectively, according to the estimated posterior mean deformation model depicted in 1(b). See text for explanation of geographic reference points. (b) Mean of 250 Markov chain Monte Carlo samples from the posterior distribution of the D-plane coordinates of the 36 monitoring sites in 2(a). The deformed grid represents a thin-plate spline interpolation mapping the square grid drawn on 2(a). The circles represent 0.90 correlation contours centered at sites 38, 59, 61, and 74, respectively.

For evaluation purposes, the model was refitted using only 22 of the 36 sites, the remaining 12 sites being used to assess the predictions of the time series of 108 (log-transformed) 10-day aggregate rainfall observations. Figure 3 illustrates performance of the method at two of the 12 sites, one in the northeast region (3(a, c)) and one in the lower central region along the Mediterranean coast (3(b, d)). In the northeast part of the region, where spatial correlation is higher and the monitoring sites more dense, there is less uncertainty in the posterior D-plane location of the monitoring site (3(a) vs. 3(b)). The point predictions are, as expected, more accurate (3(c) vs. 3(d)). The scatters show that the log-transformed rainfall are not truly Gaussian and, correspondingly, there is systematic error in prediction of zero rainfall observations.

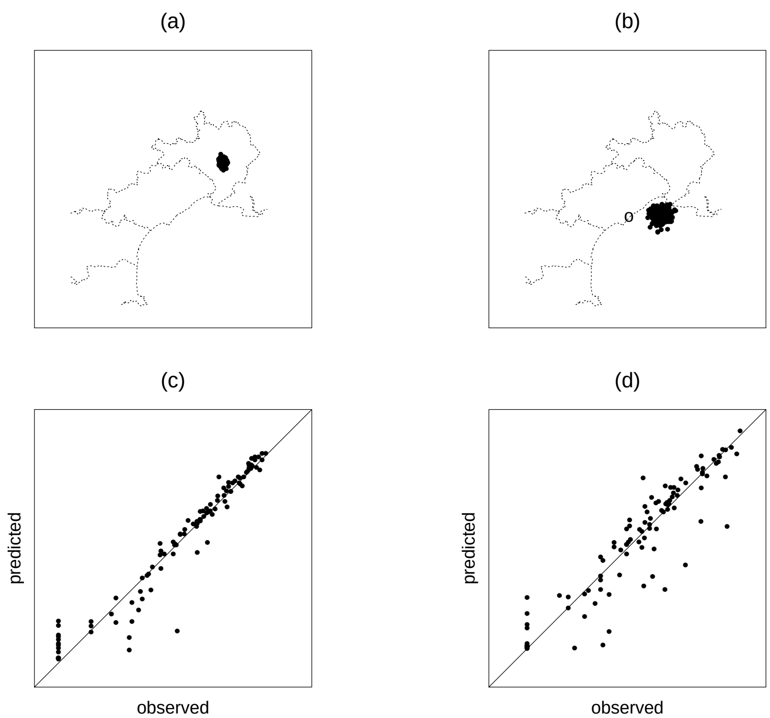

Figure 3. (a,b): scatters of a sample of 250 D-plane locations (relative to a Procrustes superposition of the D-plane configurations on the geographic map) from the posterior distribution for two monitoring sites not used in fitting the deformation model. The "O" in panel (b) indicates the original G-plane location of the second site. The scatter in panel (a) covers the original G-plane location of the first site. (c,d): corresponding scatter plots of observed versus posterior mean predictions of log-rainfall observations at these two sites.

Schmidt and O'Hagan (2000) work with the same Gaussian likelihood, but use a general Gaussian process prior for the deformation. The effect of this is mainly to replace $\mathbf{K}$ with a full rank covariance matrix in the prior $\pi(\Xi)$ for the D-plane coordinates. Schmidt and O'Hagan (2000) also differ from Damian et al. (2000) 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.

6. DISCUSSION

Although this review covers a substantial literature, the recent methodologies are still not mature in that 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 such as kriging. The Bayesian methods (Higdon et al. (1999); Fuentes (2000); Damian et al. (2000); Schmidt and O'Hagan (2000)) 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. The details of the MCMC estimation are of similar computational magnitude, the process-convolution approach putting a spatial random field prior on two parameters (ellipse foci) per monitoring site, and the spatial deformation approach requiring a random field prior on essentially two parameters (D-plane coordinates) per monitoring site. In principle, the spatial deformation model may also be applied to purely spatial problems as was the process-convolution model. 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 article has focused on essentially nonparametric approaches to the modeling of nonstationary spatial covariance structure. 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 2D deformations, and Das' (2000) development of deformation models for the covariance structure of atmospheric processes defined on the sphere. In addition, parametric models appropriate for the characterization of certain point source effects have been introduced by Hughes-Oliver and colleagues (Hughes-Oliver et al. 1998, Hughes-Oliver and Gonzalez-Faria 1999).

Acknowledgements

The U.S. Environmental Protection Agency supported this work through cooperative agreement CR825173-01-0 with the University of Washington. It has not been subjected to the Agency's peer and policy review and therefore does not necessarily reflect the views of the Agency and no official endorsement should be inferred. The authors wish to thank Jenise Swall for kindly providing Figure 1.

References

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

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

Damian, D., Sampson, P.D., and Guttorp, P. (2000). Bayesian estimation of semi-parametric non-stationary spatial covariance structures. *Environmetrics*, in press.

Das, B. (2000). Global Covariance Modeling: A Deformation Approach to Anisotropy. Ph.D.dissertation, Department of Statistics, University of Washington (http://www.cgd.ucar.edu/stats/manuscripts/das.ps).

Fuentes, M. (2000). A new high frequency kriging approach for nonstationary environmental processes. *Environmetrics*, in press.

Gneiting, T., and Schlather, M. (2001).Space-time covariance models. In: *Encyclopedia of Environmetrics*, New York, Wiley, to appear.

Guillot, G., Senoussi, R., and Monestiez, P. (2001). Definite positive covariance estimator for nonstationary random fields. In: *GeoENV 2000: Third European Conference on Geostatistics for Environmental Applications*, P. Monestiez, D. Allard, and R. Froidevaux, eds., Kluwer, Dordrecht, to appear.

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, pp.761-768.

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

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

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

Iovleff, S., and Perrin, O. (2000). Estimating a non-stationary spatial structure using simulated annealing, (http://www.math.chalmers.se/ olivier/paper/recuit.ps).

Kent, J.T., and Mardia, K.V. (1994). Link between kriging and thin plate splines. In: *Probability, Statistics and Optimization.* F.P. Kelly, ed.New York, Wiley, pp.325-339.

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.

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, pp.162-173.

Nott, D., and Dunsmuir, W.T.M. (1998). Analysis of spatial covariance structure for Sydney wind patterns. Report S98-6, Department of Statistics, University of New South Wales, Sydney, Australia
(http://www.maths.unsw.EDU.AU/statistics/preprints/s98-6.pdf).

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. (2000). Large spatial prediction problems and nonstationary random fields. Geophysical Statistics Project, National Center for Atmospheric Research, Bolder, CO, research report (http://www.cgd.ucar.edu/stats/manuscripts/krig5.ps).

Obled, C., 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.

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, 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.

Preisendorfer, R.W. (1988). *Principal component analysis in meteorology and oceanography.* Elsevier, Amsterdam.

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.M., and O'Hagan, A. (2000). Bayesian inference for nonstationary spatial covariance structure via spatial deformations. Research Report No.498/00, Department of Statistics, University of Sheffield (http://www.shef.ac.uk/ st1ao/ps/spatial.ps).

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