METHODS FOR ESTIMATING HETEROGENEOUS SPATIAL COVARIANCE FUNCTIONS WITH ENVIRONMENTAL APPLICATIONS

by

Peter Guttorp
Paul D. Sampson

TECHNICAL REPORT No. 236
August 1992

Department of Statistics, GN-22
University of Washington
Seattle, Washington 98195 USA

Methods for Estimating Heterogeneous Spatial Covariance Functions With Environmental Applications

Peter Guttorp and Paul D. Sampson*
Department of Statistics, University of Washington

ABSTRACT

Estimation of spatial covariance is important to many statistical problems in the analysis of environmental monitoring data. In this Chapter we review several different methods for spatial covariance estimation from monitoring data, with emphasis on methods for heterogeneous models. We briefly describe some applications, and outline how these methods can be extended to space-time and multivariate models.

KEY WORDS: Spatial estimation, network design, space-time process.

1. Introduction

A fundamental problem in environmental research is how properly to assess the spatial structure of pollution fields at various scales. While much attention has been given to spatial mean structures, not nearly as much effort has gone into studying higher-order stochastic structure; specifically, soatuak covariance of environmental processes.

In order to assess the severity of an environmental insult, the researcher typically has access to monitoring data from a relatively sparse network of stations, while assessment of either the actual level or the mean level (averaged both temporally and spatially) is needed over unobserved locations. Another question of practical importance is how to redesign an existing network, by removing redundant stations and adding stations in areas that are poorly represented. Common to these problems is a need to estimate the spatial covariance of the process, both at monitored and unmonitored locations. We will focus on applications to acidic precipitation, since this has been our field of concentration for quite a while. However, most of the methods discussed in this Chapter are of general use for situations where a phenomenon is monitored regularly over time at fixed locations.

The process being monitored can usually be characterized as a space-time random field $Z(x,t)$. We suppose that data are available from each of N sampling stations at the same T points in time. Let $Z_{it} = Z(x_i,t)$ denote the observation taken at location x_i at time t, $i=1,2,...,N$, $t=1,2,...,T$. We consider models of the form

$$Z(x,t) = \mu(x,t) + E_\tau(x,t) + E_\varepsilon(x,t),$$

where $\mu(x,t)$ represents a mean field which may include temporal and spatial trends, $E_\tau(x,t)$ is a mean zero space-time process which is L_2-continuous in space, and $E_\varepsilon(x,t)$ represents measurement error and small scale spatial variability. The process E_τ is assumed to be temporally stationary with a spatial

*Drs. Guttorp and Sampson have been working on problems in acidic deposition since 1985. Among their recent joint publications in this area are Sampson and Guttorp (1991, 1992), Sampson et al. (1991), and Guttorp et al. (1992a,b). Dr. Sampson participated in the first study that resulted in unequivocal proof that emission changes can result in deposition changes (Vong et al., 1987,1988). Dr. Guttorp wrote the chapters on *Spatial statistics in ecology* and *Spatial statistics in environmental sciences* in the recent National Research Council report on *Spatial Statistics and Digital Image Analysis* (Baddeley et al., 1991).

covariance function that depends smoothly on geographic coordinates. In some applications (e.g., Haslett and Raftery, 1989; Oehlert, 1992) it is possible to further decompose this field into a spatially pervasive long-memory temporal process, and a component depicting short-term and limited range space-time features.

We assume that the distribution of the measurement error process E_ε is independent of location x and time t, as well as being independent of the process E_τ. In terms of covariance, the presence of the measurement error process induces a discontinuity in the spatial covariance, in that

$$\mathrm{Var}\, Z(x,t) - \lim_{y \to x} \mathrm{Cov}(Z(x,t),Z(y,t)) = \mathrm{Var}\, E_\varepsilon(x,t).$$

This nonzero quantity on the diagonal, as it were, is called the **nugget effect** in geostatistics (e.g., Journel and Huijbregts, 1978).

There are two types of approaches to the specification and analysis of spatial covariance structure for space-time monitoring data generated by models such as this. One approach, which may be called the geostatistical approach, concentrates on a single snapshot or time-slice of the field at one time t. The other, call it the meteorological approach, assumes that there are several time-slices, either considered as independent realizations of the field, or at least as a temporally stationary realization. With the exception of Sections 2 and 3, we will largely pursue the meteorological approach in this Chapter. Haslett (1989, Section 2) gave an extensive discussion of this distinction.

In this paper we will largely ignore the (nontrivial) problem of modeling the mean field μ. While this may often be the focus of the research, the particular aspect of interest to us is the estimation of the spatial covariance. Among recent work on estimating mean acidic deposition fields are Le and Petkau (1988), Loader and Switzer (1992), and Oehlert (1992). We shall henceforth assume (without changing our notation) that the estimated mean has been removed from the data. In addition we assume that the space-time covariance function is **separable** in the sense that it factors into a product of a spatial covariance and a temporal covariance. This is a fairly strong assumption, but in our experience the temporal dependence of the spatial covariance is often mainly a seasonal dependence, and can be dealt with, if necessary, by analyzing the seasons separately (see Section 10.1 for an alternative treatment).

In the analysis of most space-time processes underlying environmental studies, there is little reason to expect spatial covariance structures to be homogeneous over the spatial scales of interest. Certainly, the heuristic arguments underlying assumptions of stationarity in time series do not readily apply to spatial processes. For example, in studies of the dispersion of atmospheric pollutants, landscape or topography affects weather patterns, and thereby the transportation of pollutants. Empirical spatial covariances may be computed using time series of residuals from some form of estimated location (and possibly time) dependent process means. These covariances may reflect paths of pollutant transportation (for data averaged over appropriate temporal scales), and should therefore be heterogeneous insofar as the landscape varies from place to place.

Prior to the 1980's, heterogeneity was most often addressed (when recognized) by partitioning a spatial field into relatively homogeneous regions. Anisotropy in the covariance structure was recognized in both the geostatistical literature (Journel & Huijbregts, 1978) and the meteorological literature (Thiébaux, 1977, 1991) and treated as a slowly varying function of the azimuth (cf. Section 2 below).

In what follows we describe several different approaches to estimating the spatial covariance function. The simplest assumption, namely that the covariance structure is homogeneous and isotropic, is the basis of much of the geostatistical work. We outline, very briefly, some of the methods applicable to this model in Section 2. We also indicate how one may extend this to an anisotropic, but still homogeneous, model. In Sections 3 and 4 we present two approaches to the estimation of heterogeneous covariances by methods of smoothing. The first is a purely geostatistical approach, while the second follows the meteorological path. Section 5 outlines an empirical Bayes approach to incorporating prior assumptions as to the structure of the covariance, while in Section 6 we present a general method of decomposition of

the covariance in terms of certain orthogonal functions. A newer approach, using a deformation of the geographic domain into a space in which the methods of Section 2 apply, is presented in Section 7. This method combines flexibility with ease of interpretation. It is generally difficult to display covariance fields, and in Section 8 we present a method that we find useful for this purpose. The use of covariance estimates in estimation and network redesign is the topic of Section 9, while extensions to genuine space-time covariance models and to multivariate processes are outlined in the final Section.

2. Homogeneous Approaches

In the theory of temporal stochastic processes, it has been quite useful to study **stationary** processes, which, in the strict sense of the word, are invariant under temporal shifts, i.e., whose distribution does not depend on the time origin. If in addition such processes have covariance functions that decrease with increasing separation, statistically valuable tools like the central limit theorem will apply (see, e.g., Brillinger, 1981, for a theoretical discussion). In the case of spatial processes, the property of invariance under shifts is usually called **homogeneity**. Matters are especially simple if the distribution of the process is also invariant under rotations, in which case it is called **isotropic**.

Define the **dispersion** between two sites located at x and y, respectively, by

$$\delta(x,y) = \mathrm{Var}(Z(x,t)-Z(y,t)) = \mathrm{Var}\,Z(x,t) + \mathrm{Var}\,Z(y,t) - 2\,\mathrm{Cov}\,(Z(x,t),Z(y,t)). \qquad (2.1)$$

Knowing the variance field, one can compute the covariance from the dispersion and vice versa. If the spatial field is homogeneous, δ is a function only of the vector $x-y$, rather than the absolute locations x and y. If the field is also isotropic, δ is a function only of the length of this vector. Hence we can write

$$\delta(x,y) = h(|x-y|).$$

The function h is often called the **variogram** (terminology originating in the field of geostatistics). In the atmospheric science literature this is called the **structure function**. The choice of variogram h must be made so as to ensure that the corresponding covariance function is valid, i.e., nonnegative definite. Some popular choices of variograms are (cf. Cressie, 1991, Ch. 2.3.1):

Linear $\quad h(s) = \sigma_0^2+\sigma^2 s$

Exponential $\quad h(s) = \sigma_0^2+\sigma^2(1-\exp(-\alpha s))$

Gaussian $\quad h(s) = \sigma_0^2+\sigma^2(1-\exp(-\alpha s^2))$

Spherical $\quad h(s) = \begin{cases} \sigma_0^2+\sigma^2(1.5\alpha s-0.5(\alpha s)^3 & 0<s\le\alpha \\ 0 & \text{otherwise.} \end{cases}$

Here σ_0^2 is the nugget effect. For isotropic homogeneous spatial random fields, contours of constant spatial dispersion, relative to a fixed reference point, are concentric circles.

The assumptions of homogeneity and isotropy may be natural in some geological applications, such as mining. More importantly, they may be necessary in order to estimate the variogram. If there is only one realization of the process, rather than the repeated observations typical for monitoring networks, the variogram has to be estimated empirically using average squared distance between values of process at similar distances, i.e., averages over all pairs of points that are within a bandwidth parameter of being the same distance apart. As in many other statistical applications, the bandwidth parameter determines the tradeoff between bias and variability. A specific variogram model can then be fitted to this empirical variogram using (weighted) least squares or other methods (Cressie, 1991, Sections 2.4 and 2.6,). For example, Cressie et al. (1990) fitted a spherical variogram to yearly hydrogen deposition from 19 stations in the UAPSP acid precipitation monitoring network in 1982, and separately for 1983 data. The two fits

were rather different, and neither fit the corresponding empirical variogram very well. Using more data from the same network, and a temporal perspective, Guttorp et al. (1992b) deduced substantial deviations from isotropy and homogeneity (cf. Section 9.1).

In environmental applications isotropy and homogeneity are usually reasonable assumptions only over small areas. Orographic effects, synoptic wind patterns, and the irregular mesoscale structure of precipitation fields all contribute to anisotropy in regional patterns of acidic precipitation. A first step towards a more realistic model is to drop the isotropy assumption, while maintaining the spatial homogeneity. A simple way of doing this is to divide the directions up into zones, and estimate the variogram separately in each directional zone (so-called zonal anisotropy). Many homogeneous variogram models have elliptical anisotropy (also called geometric anisotropy), i.e., the contours of constant dispersion relative to a reference point are concentric ellipses. We can write

$$h(u) = \gamma((u'Bu)^{1/2}),$$

where u is the vector $x-y$ and γ an appropriately chosen monotone function (here we do not consider hole effects or periodic variograms which are not monotone). In this case a linear transformation of the station coordinates, namely

$$x \to f(x) = Ax,$$

where $A'A=B$, results in a problem with isotropic covariance structure. Thus, most variogram models commonly used in geostatistical practice assuming homogeneity can be expressed as

$$\delta(x,y) = h(x-y) = \gamma(|f(x)-f(y)|). \tag{2.2}$$

In other words, a linear transformation of the original map yields a map on which dispersion corresponds to distance.

Let d_{ij} be the empirical dispersion between the observed values of the field at monitoring locations x_i and x_j. In the case of monitoring data this can be estimated using the sample covariance between the monitoring stations and equation (2.1). It may then be reasonable to separate out the two estimation problems: the linear map A and the variogram γ. For a given class of variogram functions, depending on a low-dimensional parameter θ, one would minimize over A and θ, either simultaneously or separately in an alternating iterative fashion, the quantity (Cressie, 1985)

$$\sum_{i,j} \left[\frac{d_{ij} - \gamma(|A(x_i - x_j)|;\theta)}{\gamma(|A(x_i - x_j)|;\theta)} \right]^2. \tag{2.3}$$

Under the assumption of homogeneity there is a Fourier approach to covariance estimation. For simplicity we consider only one time-slice t with observations $Z(x_i,t)$, $i=1,\ldots,N$. Write

$$\sigma^2\rho(u) = \mathrm{Cov}(Z(x,t),Z(x+u,t))$$

where $\sigma^2=\mathrm{Var}\,Z(x,t)$, and $\rho(u)$ is the spatial correlation function for separation vector u. The **normalized wave-number spectrum** $\phi(w)$ is the Fourier transform of the correlation function, i.e.,

$$\phi(w) = \int \rho(u)\exp(i(u\cdot w))dw.$$

As in the time series case there is a spectral representation of the process Z (e.g., Bartlett, 1966, Section 6.50), namely

$$Z(x,t) = \int \exp(i(x\cdot w))d\zeta(w)$$

where $\zeta(w)$ is a complex-valued process orthogonal in both components of w with

$$\mathrm{E}|\zeta(w)|^2 = \sigma^2\phi(w).$$

This suggests basing an estimate of ϕ (and, by back-transformation, ρ) on the discrete Fourier transform of the data. Brillinger (1985) expands on this. If the observations are on a regular grid (as would be the

case, for example, for some remote sensing applications) it is possible to use FFT methods. However, for landbased irregularly spaced monitoring data the corresponding estimation problem is difficult. The methods of Brillinger (1972) apply.

3. Moving Windows

Although most environmental processes are heterogeneous in both mean and covariance structure, the process may often be considered approximately homogeneous over subregions. With this in mind Haas introduced a technique called "moving-window kriging" in which the spatial covariance structure (represented in terms of a variogram model) is estimated in a circular window about each target point at which a spatial (kriging) estimate is to be computed (Haas 1990a,b). This approach, which provides an alternative to partitioning the overall region of interest into a small number of non-overlapping subregions, appears to be the only method suggested to deal with heterogeneity in the context of a single realization of a spatial process.

Haas designed his procedure for estimation of annual wet acid deposition using data from the NADP/NTN monitoring network. At each point of a relatively fine grid on which spatial estimates are desired he selects a window containing at least 35 sites. He then enlarges the radius of the window, as necessary, until there is at least one pair of sites represented in each of the distance bins over which a variogram is to be estimated (generally with at least two pairs in the shortest spatial lag bin of 50 km), and so that the nonlinear regression estimation of an isotropic, spherical variogram model converges. See Haas (1990b) for further discussion of the rationale underlying these heuristics. Haas uses only isotropic variogram models, although his general approach could accommodate anisotropic models as presented here in Section 2.

Haas' approach is intuitively attractive, albeit computationally demanding. (In Haas, 1990b, he computed variogram parameters by nonlinear least squares in windows centered on each of 2000 grid points—and solved 2000 kriging systems for the point estimates!) Perhaps the major source of concern in this procedure is the ad hoc approach to choice of window sizes and the difficulty in assessing the uncertainty in the estimated parameters of the spatially varying variogram models. Cross-validatory choice of window size does not appear feasible. Nonetheless, for point estimation with single realizations the general approach is quite sensible. This is a purely spatial (or geostatistical) approach, rather than a meteorological approach.

4. A Kernel Smoothing Method

Oehlert (1992) introduced a kernel method for estimating spatial covariance, based on the idea that spatial covariance (or dispersion) should vary relatively smoothly. Define a kernel κ for two pairs of locations (x_1,x_2) and (x_3,x_4) by

$$\kappa((x_1,x_2),(x_3,x_4)) = \begin{cases} \exp(-(d(x_1,x_3)^2+d(x_2,x_4)^2)/\lambda_\kappa^2) & \text{if } x_1 \neq x_2 \text{ and } x_3 \neq x_4, \text{ or } x_1=x_2 \text{ and } x_3=x_4 \\ 0 & \text{otherwise} \end{cases}$$

where $d(x,y)$ is the Euclidean distance between x and y, and λ_κ is a smoothing parameter. Given appropriately detrended and deseasonalized residuals and a corresponding sample spatial covariance matrix $S=[S_{ij}]$, based on monitoring N stations at locations $x_1, \ldots, x_N$, the smoothed covariance matrix $\hat{S}$ is given by

$$\hat{S}_{ij} = \frac{\sum_{k,l} S_{kl}\kappa((x_i,x_j),(x_k,x_l))}{\sum_{k,l}\kappa((x_i,x_j),(x_k,x_l))} .$$

Notice that this method smooths variances and covariances separately, but using the same smoothing parameter. A small value of λ_κ essentially returns the raw matrix S. For the purposes in Oehlert's paper, namely estimation of spatial means and trend slopes, the choice of λ_κ was not critical. He proposed to

choose it so that a plot of the rescaled variances (observed divided by smoothed) were linear on a chi-square plot, with degrees of freedom chosen appropriately for Gaussian theory. It is not clear whether or not this procedure would always yield a valid (i.e., positive definite) covariance structure. When applied to annual sulfate data from several acid rain monitoring networks in the northeastern United States and southeastern Canada (APIOS-C, MAP3S, NADP/NTN, and UAPSP), the resulting covariance estimate showed distinct, but weak, covariance structure, with strongest covariance along the NE/SW axis, consistent with the prevailing southwesterly winds, and weakest along a NNW/SSE axis.

5. An Empirical Bayes Approach

Loader and Switzer (1992) suggested an empirical Bayes type of shrinkage approach to smoothing the sample spatial covariance matrix S. Specifically, assuming an underlying Gaussian distribution for monitoring data, they assigned the inverted Wishart conjugate prior distribution for the unknown true $N\times N$ covariance matrix Σ with density

$$f_\Sigma(\Sigma|\tilde{C},m) = \frac{|\tilde{C}|^{m/2}|\Sigma|^{(m+p+1)/2}}{2^{mp/2}\Gamma_p(m/2)} \exp(-\tfrac{1}{2}tr(\Sigma^{-1}\tilde{C})) \tag{5.1}$$

with prior mean $C = \tilde{C}/(m-p-1)$, and $\Gamma_p(t) = \pi^{p(p-1)/4}\prod_{j=1}^{p}\Gamma(t-\tfrac{1}{2}(j-1))$. The conditional distribution of S given Σ is then a Wishart distribution and the posterior mean of Σ has the form

$$E(\Sigma|S) = \lambda S + (1-\lambda)C \tag{5.2}$$

where $\lambda = n/(n+m-p-1)$.

In practice, the prior mean C is specified by fitting a parametric model to the observed S_{ij}. In analyzing monthly average sulfate concentrations from the UAPSP network over the two year period 1982-83 they fitted a simple exponential model, $C(x,y) = \sigma^2\exp(-\alpha s)$ where $s = |x-y|$. The parameter m of the prior distribution, which specifies the shrinkage parameter λ, was estimated under an empirical Bayes paradigm from the marginal density of S after integrating with respect to the prior.

Extension of this approach to covariances involving unmonitored sites is somewhat problematic. For a single unmonitored location x, they followed a suggestion of Switzer (1989) in estimating the vector of covariances by $\sigma(x) = (\sigma(x,x_1), \ldots, \sigma(x,x_N))$,

$$\hat{\sigma}(x) = \hat{\Sigma}C^{-1}c(x),$$

where $\hat{\Sigma}$ is the $N\times N$ posterior mean given in (5.2) and C and $c(x)$ are the $N\times N$ matrix of covariances and the N-vector of covariances, respectively, from the specified a priori parametric model. Note that when taking $x = x_i$, one of the actual monitoring sites, this simply returns the ith column of $\hat{\Sigma}$, so the extension is faithful to the empirical Bayes estimates for the observed locations.

This approach does not lend itself easily to providing covariance estimates for an arbitrary number of unmonitored locations. Loader and Switzer suggested a sequential approach which, fortunately, is invariant to the order of the sequence of new sites, but for which the specification of the diagonal elements (variances) for the new sites can cause problems with positive-definiteness. This framework provides no general mechanism for estimation of covariances involving an arbitrary pair of unmonitored sites.

A disturbing aspect of this extension to unmonitored sites is that it does not affect spatial point estimates. Assuming known mean and covariance fields, $\mu(x,t)$ and $\sigma(x,y)$, respectively, the Gaussian-based maximum likelihood estimate of the random field $Z(x,t)$ at an unmonitored site x is

$$\hat{Z}(x,t) = \mu(x,t) + \sigma(x)'\Sigma^{-1}(Z(\cdot,t)-\mu(\cdot,t)).$$

Substituting the estimators given above we find that

$$\hat{\sigma}(x)'\hat{\Sigma}^{-1} = c(x)'C^{-1},$$

i.e., the estimated value at location x depends only on the (prior) parametric covariance estimate, not on the empirical Bayes estimate $\hat{\Sigma}$. The variance of the estimated value, however, does depend on $\hat{\Sigma}$.

For the application to 19 UAPSP acid deposition monitoring sites in the eastern U.S., plots of interpolated covariances displayed essentially no local anisotropy. This stands in contrast to the nature of the spatial covariance structure estimated using similar data in Sections 4 and 9.1. Switzer, in a recent technical report (Monestiez and Switzer, 1991), written subsequent to the article by Loader and Switzer (despite the publication dates), used a multi-dimensional scaling modeling approach similar to that introduced by Sampson and Guttorp (1992) and discussed in Section 7.

6. Empirical Orthogonal Functions

The first, and seemingly the most general, published analytic approach to heterogeneity in space-time processes is provided by an analysis of a space-time random field in terms of a Karhunen-Loève decomposition (see, e.g., Adomian, 1983, Section 2.10) or **empirical orthogonal functions** (EOF's; see, e.g., Preisendorfer, 1988). While the focus of most of the applications of this approach has been on the interpretation of these EOF's (or spatial principal components), Obled and Creutin (1986; see also Creutin and Obled, 1982) demonstrated that an EOF analysis could be used for spatial estimation and that it readily accommodated anisotropic and even heterogeneous covariance or correlation structure. We will begin here with the simpler analysis in terms of spatial EOF's or principal components as presented, for example, in Preisendorfer (1988).

The Karhunen-Loève decomposition of the (mean-centered) field $Z(x,t)$ in terms of spatial and temporal components is

$$Z(x,t) = \sum_{k=1}^{\infty} F_k(x) A_k(t) \tag{6.1}$$

where $F_k(x)$ represents the kth spatial component (EOF) and $A_k(t)$ represents the temporally varying amplitude of the kth spatial component. Most of the atmospheric science literature addresses this decomposition only for a finite number of monitoring sites x_i, $i=1,2,\ldots,N$ and observation times t, $t=1,2,\ldots,T$. In this case the decomposition evaluated at the sampling sites x_i involves a sum with N_T terms, $N_T = \min(N,T)$. Let Z denote the $N\times T$ matrix of observations $[Z(x_i,t)]$. With the usual scaling conventions we compute the $N\times N_T$ matrix $F = [F_{ik}] = [F_k(x_i)]$ as eigenvectors of the sample spatial covariance matrix, $S = \frac{1}{T} Z Z'$ so that

$$F'F = I.$$

These components are thus analytically orthogonal in space. The $N_T\times T$ matrix of temporal components $A = [A_{kt}] = [A_k(t)]$ are statistically orthogonal or uncorrelated and satisfy

$$\frac{1}{T} A'A = \Lambda,$$

where Λ is the diagonal matrix of eigenvalues of S. Because these coefficients of the F_i in the expansion of $Z(x,t)$ are uncorrelated, this is a spectral representation.

The spatial covariance matrix S is then represented in terms of the EOFs by the usual spectral decomposition

$$S = F'\Lambda F = \sum_{k=1}^{N_T} \lambda_k F_k F_k'. \tag{6.2}$$

The sample estimates for station pairs, S_{ij}, can be approximated (or smoothed) by retaining a smaller number of spatial components. Note that this type of model for the observed spatial covariances, unlike

the nonparametric model of Oehlert (Section 4 above) and the empirical Bayes approach of Loader and Switzer (Section 5 above) makes no explicit reference to spatial distance and takes no account of the relative spatial locations of the observations. For this reason it does not provide estimates of spatial covariance at unmonitored locations without extending or spatially interpolating the EOF's which are computed only at the monitoring sites.

Extension of an EOF analysis in order to address spatial estimation appeared in Cohen and Jones (1969) and Buell (1972, 1978). Here we present the development of Creutin and Obled (1982) and Obled and Creutin (1986). In the general case the Karhunen-Loève expansion represents $Z(x,t)$ for arbitrary spatial location x and time t as the sum in equation (6.1) with the eigenfunctions $F_k(x)$ being solutions of the Fredholm integral equation

$$\int S(x,y)F_k(x)dx = \lambda_k F_k(y)$$

with the analytic orthogonality constraint

$$\int F_k(x)F_l(x) = \delta_{kl},$$

where δ_{kl} is the Kronecker delta function.

Generalizing the finite decomposition of the sample covariance matrix S given in (6.2) above, we can represent the covariance function as

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

A solution to this general problem requires a numerical approximation to the Fredholm integral equation. Given N sampling sites we can estimate at most N eigenfunctions. Obled and Creutin solved this by approximating the true field $Z(x,t)$ (and the eigenfunctions) in terms of a vector space spanned by an elementary set of generating functions $e_1(x), e_2(x), \cdots, e_p(x)$. The simplest approach (yielding the solution proposed by Cohen and Jones, 1969) is to take the generating functions to be piecewise constant non-overlapping functions with $e_i(x)$ set equal to one on the Thiessen polygon P_i containing x_i and zero elsewhere. In this case, since the $e_i(x)$ are orthogonal, the numerical problem to be solved remains a symmetric $N \times N$ eigenproblem as in the case of the principal component solution above, but with weights on the elements of the covariance matrix reflecting the areas of the associated Thiessen polygons. Other generating functions, like the facet-like linear ones used by Obled and Creutin (1986), considerably complicate the definition of the eigenproblem to be solved. See their work for further details, as well as Preisendorfer (1988, Section 2d).

Creutin and Obled applied this EOF approach to spatial estimation for a network of 101 rainfall monitoring stations and 81 rainfall events over a period of 20 years. They fitted the EOF model to 73 stations and reserved 28 stations for testing the model. In their analysis they argue that it is appropriate to retain 39 components in the expansion. The resulting spatial covariance model is certainly complex. It would appear to be difficult to study and interpret the structure of this spatial covariance model.

Some insight can be gained by considering the vector

$$F^p(x_i) = (\sqrt{\lambda_1}F_1(x_i), \sqrt{\lambda_2}F_2(x_i), \cdots, \sqrt{\lambda_p}F_p(x_i))$$

as coordinates for a representation of location x_i in p dimensions. Then

$$\begin{aligned} |F^p(x) - F^p(y)|^2 &= \sum_{k=1}^{p} \lambda_k (F_k(x) - F_k(y))^2 \\ &= \sum_{k=1}^{p} \lambda_k F_k(x)^2 + \sum_{k=1}^{p} \lambda_k F_k(y)^2 - 2\sum_{k=1}^{p} \lambda_k F_k(x)F_k(y). \end{aligned}$$

The right-hand side is the p-dimensional EOF model for the spatial dispersion, i.e., the dispersion can be

estimated by

$$\hat{\delta}(x,y) = \left|F^p(x) - F^p(y)\right|^2. \tag{6.2}$$

Thus the spatial covariance or spatial dispersion structure can be understood in terms of distances between stations represented as points in a high-dimensional Euclidean space. We will contrast this model with equation (2.2) and a related model in Section 7 below.

7. Multidimensional Scaling

Sampson and Guttorp (1992) proposed an approach to estimation of heterogeneous spatial covariance that allows estimates between any two points of the field, whether monitored or not. Recall from equation (2.2) that most homogeneous variograms are of the form

$$\delta(x,y) = \gamma(\left|f(x) - f(y)\right|)$$

where f is a linear transformation (the notation in this Chapter is somewhat different from that in Sampson and Guttorp, 1992). In order to produce a large class of heterogeneous dispersions, we allow f to vary over the class of smooth invertible functions. The basic concept is that dispersion is a smooth function of geographic coordinates and is locally anisotropic. By transforming the geographic map appropriately we get a new representation where dispersion is a monotone function of distance, i.e., where the dispersion is homogeneous and isotropic. The EOF representation (6.2) is of this form, with $f \equiv F^p$ being a mapping from R^2 to R^p, and $\gamma(s)=s^2$. The approach of Sampson and Guttorp described in the remainder of this Section is more flexible. We suspect that it will lead to far simpler (lower-dimensional) models in most cases, but extensive comparisons have not yet been carried out.

The first step in implementing this idea is to transform the geographic map of the monitoring stations—the "G-plane"—to a map (possibly in dimension higher than 2) where dispersion increases (approximately) monotonically with distance—the "D-image." This can be accomplished using nonmetric multi-dimensional scaling (MDS). The D-image configuration of stations with coordinates $\tilde{x}_i$ is chosen so that interpoint distances $h_{ij}=\left|\tilde{x}_i - \tilde{x}_j\right|$ minimize the stress measure

$$\min_{\psi} \frac{\sum_{i<j} (\psi(d_{ij}) - h_{ij})^2}{\sum_{i<j} h_{ij}^2}, \tag{7.1}$$

where the minimum is taken over all monotone functions ψ, so that $\psi(d_{ij})$ represents a least squares monotone regression of the h_{ij} on estimated spatial dispersions d_{ij} (Mardia, Kent and Bibby 1979). In order to ensure the validity of the resulting dispersion estimate when the variance field is not constant, we compute the d_{ij} based on the **correlation dispersion**, defined by

$$\delta^*(x,y) = 2 - 2\,\mathrm{Corr}\,(Z(x),Z(y)),$$

which can be estimated from the empirical correlation matrix.

A plot of MDS distances h_{ij} versus empirical dispersions d_{ij} indicates the quality of this spatial model. To this plot we fit a general model for isotropic homogeneous dispersions (i. e., a variogram model), since by construction the measurements, when referred to the D-image rather than the G-plane, correspond to a homogeneous and isotropic random field. Here it is possible to use any valid variogram model, as well as the spectral approach of Section 2. In the spirit of nonparametric estimation we note, following Matérn (1986, Section 2.3), that every isotropic correlation function in R^n, continuous except at the origin, can be written in the form

$$\rho(h) = \sigma_0^2 + G(0) + \int_0^\infty \Lambda_k(hw)\,dG(w) \tag{7.2}$$

where

$$\Lambda_k(v) = k!(2/v)^k J_k(v),$$

J_k is the Bessel function of the first kind with $k=(n-2)/2$, and $G(x)$ is an arbitrary distribution function of a non-negative random variable. Results of Böhning (1982) generalize to produce an algorithm for the fitting of this model. Sampson and Guttorp (1992, Appendix) applied the algorithm to the special case of a mixture of Gaussian variogram models. The Gaussian model is obtained from (7.2) by choosing

$$dG(w) \propto w^{n-1}\exp(-w^2/4a^2)$$

(Matérn, 1986, Section 2.4). The mixture of Gaussian variograms is the general form for a variogram model that is valid in any dimension (Schoenberg, 1938, pp. 817 ff.). In Guttorp et al. (1992a) a simple exponential variogram was used. A mixture of exponential variograms yields the class of completely monotone variogram functions (Matérn, 1986, Section 2.4; he also outlines how this is obtained from 7.2). An improved algorithm for these mixture problems is given by Lesperance and Kalbfleisch (1992).

To complete the estimation procedure we must compute a smoothly invertible function f that maps the G-plane into the D-image. Generally (except for applications with extremely high covariances, such as that in Sampson and Guttorp, 1992) an interpolating function that maps geographic coordinates $x_i, i=1,\ldots,n$ into the coordinates $\tilde{x}_i$ returned by the MDS algorithm will not be invertible. Although it is natural to initialize the MDS algorithm with the geographic coordinates, the final coordinates $\tilde{x}_i$ may be quite jumbled in comparison to the geographic configuration because of noise in the empirical d_{ij}. In effect, the MDS algorithm fits the d_{ij} too accurately, in that the mean squared error in the variogram scatter of h_{ij} versus d_{ij} is less than the sampling error in the d_{ij}.

To correct both these problems we compute the function f using thin-plate smoothing splines (Wahba, 1990), one for each D-image coordinate, and write $f_d(x)=x^*$. In the limit, as the spline smoothing parameter gets arbitrarily large, the mapping $f_d:\mathrm{R}^2\to\mathrm{R}^2$ is linear and the resulting spatial dispersion model becomes homogeneous (as in Section 2). We choose the smoothing parameter to ensure that the spread in the plot of $h_{ij}=|x_i^*-x_j^*|$ against d_{ij} is similar in size to the sampling variability of the d_{ij}, as computed for the Gaussian case.

We estimate the spatial variance field $V(x)$ separately using another thin-plate smoothing spline, $f_v:\mathrm{R}^2\to\mathrm{R}^1$. The smoothing parameter is chosen so that the difference between raw and smoothed variances at each station behaves like the standard error of the raw variances for a Gaussian situation.

The composition of the computed smoothing spline f_d, the fitted variogram γ, and the estimated variance field f_v yields our model for the spatial covariance structure. In order to estimate the dispersion between two points x and y (monitored or unmonitored), one first computes their locations in the D-image, $x^*=f_d(x)$ and $y^*=f_d(y)$, and then estimates their correlation dispersion by $\gamma(|x^*-y^*|)$. In order to compute the corresponding covariance estimate $\hat{\mathrm{Cov}}(x,y)$ on the original (non-standardized) scale of measurement, let $\hat{V}_x=f_v(x)$. Then the estimated covariance is given by

$$\hat{\mathrm{Cov}}(x,y) = (\hat{V}_x\hat{V}_y)^{1/2}(1-\tfrac{1}{2}\gamma(|x^*-y^*|)).$$

Cross-validation or resampling methods can be used to estimate the accuracy of these estimates.

Monestiez and Switzer (1991) noted that the separation of the scaling from the estimation of the variogram was unnatural, and developed a technique in which the nonmetric scaling is replaced by a metric scaling. Given a parametric class of variogram models $\gamma(h;\theta)$, an initial set of parameters θ, and an initial configuration of points x_i^*, their algorithm has two steps

A-STEP

Move the N points x_i^* to minimize the least squares criterion

$$\sum_{i,j}(d_{ij}-\gamma(|x_i^*-x_j^*|;\theta))^2.$$

B-STEP

For the new pattern of points estimate the value of θ from the new distances using the same least squares criterion.

Return to the A-step, unless the changes in both A- and B-steps are below a convergence criterion.

Cressie's weighted least squares criterion (2.3) could readily be substituted in the Monestiez-Switzer algorithm. Monestiez and Switzer also suggested that a three-stage algorithm, including the fitting of mean and variance field as a C-step, but noted that this may encounter problems of instability.

8. Display of Covariance Fields

Display and interpretation of the resulting mapping f is important in our analysis. We have found the method of **biorthogonal grids** introduced by Bookstein (1978) in the context of morphometrics, valuable in this regard (Sampson et al., 1991). To describe this method, it is again useful to start with the homogeneous case. Linear or affine transformations of the geographic plane, $f(x)=Ax$, appropriate for a homogeneous spatial dispersion structure, are characterized by a single pair of principal axes given by the eigenvectors of $B=A'A$. The geographic direction or axis corresponding to the largest eigenvalue is the direction in which the geographic plane is (relatively) most stretched. The degree of such stretching, i.e. the ratio of distance in the deformed image to distance in the geographic plane, is given precisely by the square root of the largest eigenvalue (the first singular value of A), which is called the first **principal strain**. Given two locations x_a and x_b such that $|x_a-x_b|=c$, the spatial dispersion is greatest when x_a and x_b fall on a line parallel to this principal axis since this maximizes $|Ax_a-Ax_b|$. Conversely, for a given geographic separation, spatial dispersion is least (spatial correlation greatest) when x_a and x_b are aligned parallel to the principal axis of smallest eigenvalue. The ratio of the two principal strains is a measure of anisotropy. The affine derivative matrix A, together with the variogram function γ, completely characterize homogeneous spatial dispersion models with elliptical anisotropy.

We describe smooth heterogeneous spatial dispersion structures as those for which the principal axes of the affine derivative matrix of the mapping f vary smoothly from place to place. For any differentiable transformation, $f=(f_1(x),f_2(x))$, we can evaluate the affine derivative matrix,

$$\begin{bmatrix} \partial f_1/\partial x_1 & \partial f_2/\partial x_1 \\ \partial f_1/\partial x_2 & \partial f_2/\partial x_2 \end{bmatrix},$$

at any point $x=(x_1,x_2)$. These derivatives are easily computed for the thin-plate spline mappings used in the approach of Section 7. The (local) principal axes computed from this matrix are differentials at 90°. It is a theorem that through almost every point of a differentiable transformation pass just two differentials which are at 90° both before and after the transformation (see Bookstein, 1978). The integral curves of these differentials, sampled finitely, form a curvilinear grid whose intersections are at 90° in both images. These are called the biorthogonal grids for the transformation because there are two such grids, one in each image (or plane), corresponding curve for curve and intersection for intersection. They provide a useful depiction of the spatial variation in the (local) principal axes and principal strains of the nonlinear transformation. For discussion of the way in which these curves reflect the meteorological process(es) driving the random field being modeled, see the applications in Section 9.1 below and in Sampson et al. (1991).

9. Some Applications

There are many different applications in environmental science that require estimated spatial covariances. In this Section we discuss, very briefly, two such cases, namely how to estimate unobserved values of a spatial field, and how to redesign an existing monitoring network.

9.1. Spatial estimation

One of the main uses of monitoring data is for estimating the level of the pollutant at an unobserved site, or, more generally, over an area with or without monitoring stations in it. The kriging approach in geostatistics (pioneered by Krige, 1951, and developed by Matheron, 1963b; cf. Cressie, 1991), and the parallel theory of objective analysis in meteorology (due to Gandin, 1963; cf. Thiébaux and Pedder, 1986), were developed with this as their main goal. Using least squares ideas, the optimal linear predictor of the field at a point x, based on monitored values at points $x_1, \ldots, x_N$ (and assuming a known constant mean) is

$$\hat{Z}(x) = \sum \lambda_i z_i$$

Since we are here referring to observations corresponding to one point in time, the time argument has been dropped. The weights λ_i are computed from the covariances; in vector form

$$\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_N)' = \Sigma^{-1}\, \sigma(x)$$

where Σ is the estimated covariance for the monitored sites, and $\sigma(x)$ is the estimated vector of covariances between the monitored sites and the site x. These are estimated from the entire record of monitoring data. Provided that the variance field of $Z(x)$, Σ and $\sigma(x)$ are known (rather than estimated) the variance of the estimated value is

$$\mathrm{Var}\, Z(x) - \sigma(x)'\, \Sigma^{-1}\, \sigma(x).$$

It is common in geostatistical applications to assume an unknown mean which may be either constant or parametrized as a varying function of, for example, longitude and latitude. In these cases unbiasedness constraints are imposed on the linear estimator, and the resulting kriging system of linear equations for estimation of the weights λ can be expressed in terms of the dispersion function, rather than the covariance function. Indeed, the variance of the spatial process (and hence the covariance function) need not be assumed finite (Matheron, 1963a). In environmental monitoring applications, as well as in most meteorological and climatological applications, it is usually possible to estimate the mean field from historical data or from a theoretical model. We therefore limit ourselves to the case of a specified mean. For further discussion of this issue, see Haslett (1989, Section 2.1).

In order to illustrate this methodology, we briefly review results in Newman et al. (1992). Here the spatial covariance was estimated based on logarithms of hydrogen deposition for the UAPSP wet deposition network, based on four-weekly summaries of winter data from 1982-1987 using 17 stations.

***Figure 9.1 about here ***

For an analysis of the estimated correlation dispersion, Figure 9.1 shows how a regular geographic grid gets deformed in the D-image, according to the computed thin-plate spline mapping f_d. Figure 9.2 displays the biorthogonal grid corresponding to the estimated correlation dispersion. This analysis is based on residuals from a space-time mean model containing spatially smooth periodic seasonal factors and station-by-station smooth temporal trends, also smoothly spatially interpolated.

*** Figure 9.2 about here ***

The curves running generally SW-NE in Figure 9.2 indicate the directions of the local principal axes of smallest principal strain, i.e. the directions of weakest dispersion or strongest correlation. Four different line types encode the value of the principal strain with solid curves representing the smallest principal strains and dotted curves representing the largest principal strains. The correlation structure is heterogeneous because these grids are in fact curved, and because the degree of spatial correlation and local anisotropy varies from place to place. Near the point where Massachusetts, Vermont and New York meet, the principal strain ratio is approximately 2.05/0.81 = 2.5. In Virgina this ratio is 1.98/0.58 = 3.4, in

central Illinois it is 1.48/0.96 = 1.7, while in Alabama the principal axes have rotated slightly and the strain ratio is 1.83/1.05 = 1.7. In addition, there is a generally increasing trend in the variance field from west to east, adding to the heterogeneity of the covariance field. The structure appears similar to that reported by Oehlert (1992; cf. Section 4) for annual sulfate deposition, although the heterogeneity of the correlation structure is not discussed or depicted in his paper. This deformation is consistent with the prevailing wintertime meteorology.

*** Figure 9.3 about here ***

The effect of the deformation is indicated in Figure 9.3, with the raw correlation dispersions plotted against geographic distance on the left, and against D-image distance on the right. The scatter is appreciably tighter, and the improvement in terms of root mean squared error is 13%.

Figure 9.4 shows the estimated hydrogen field (after taking antilogarithms) for March of 1983 based on an isotropic covariance structure for the residuals (left) and the heterogeneous covariance approach from Section 7 (right). In both approaches a Gaussian variogram was judged appropriate.

*** Figure 9.4 about here ***

While much of the structure results from the (common) estimated mean field, and thus is grossly similar, there are some differences in details. In particular, the field obtained from the isotropic approach is flatter, and has less directional structure.

A clearer image of the difference between the estimates is shown in Figure 9.5, where the kriging weights resulting from the two estimators are shown.

*** Figure 9.5 about here ***

These weights are relative to a location in West Virginia, marked with an X on the figures. The importance of the SE-NW axis of strong covariance is clear.

9.2. Redesign of monitoring networks

An important purpose of a monitoring network is to detect potential changes in key environmental parameters. However, the designer of a long term monitoring network cannot fully foresee all of the benefits that may be derived from the network by its future users. Environmental engineers, resource developers, biologists, human health agents, etc., will need the data for a variety of purposes, some of which will not even have been identified. In addition to the detection problem mentioned above, there is a need for inference about changes in areal averages, and about the location of extreme changes. Caselton and Zidek (1984) suggested dealing with these multiple objectives by an approach which may be suboptimal in specific cases, but has overall merits for these types of networks. Recent applications to redesign of acid precipitation networks are Guttorp et al. (1992a) and (1992b). We discuss redesign rather than design, since statistical considerations are unlikely to be given much emphasis in the early stages of designing a monitoring network.

Let Z denote a random field of measurable quantities indexed by potential site labels i. We decompose Z into the gauged sites $G=(Z_i, i\in \mathrm{M})$ and the ungauged sites U. One approach is to choose M in order to maximize the amount of information in G about U. Caselton and Zidek (1984) suggested Shannon's index of information transmission $I(U,G) = \mathrm{E}(\log(f(U\mid G)/f(U)))$, where $f(U\mid G)$ is the conditional density of U given G and $f(U)$ is the a priori density of U. Assuming that the random field to be monitored results in multivariate normal measurements, this information criterion is just $I(U,G) = -\frac{1}{2}\log|I-R|$, where I is the identity matrix and R the diagonal matrix with elements the squared canonical correlation coefficients between the sites (variables) in U and those in G. These correlation

coefficients must be computed from some form of model for the spatial correlation structure such as that proposed in Section 7. Guttorp et al. (1992b) demonstrated this type of analysis.

A different but related approach adopts the general viewpoint that the purpose of the network is to reduce uncertainty—uncertainty about the values to be observed on the random field, and uncertainty in the parameters of the underlying models that specify the distribution of the random field. Caselton et al. (1992) proposed this approach and discussed its relation to the information criterion. The uncertainty about a random quantity X (e.g., random field) is taken to be the entropy of its distribution,

$$H(X) = E[-\log f(X)/h(X)]$$

where $f(x)$ is the density of X and $h(x)$ is a possibly improper reference density representing "complete ignorance." Here, entropy is calculated conditional on available data D. In Guttorp et al. (1992a) this rationale is applied in the context of augmenting an existing acid deposition monitoring network with new sites. It is shown that the appropriate optimization criterion, called *MEAS* (for measurement), is the uncertainty about the additional sites to be gauged, or

$$MEAS = E[-\log(f(G|D)/h(G))|D].$$

Using a hierarchical Bayesian approach, with the mean and spatial covariance matrix for the random field described by the normal-inverted Wishart conjugate prior distribution, it is shown that this can be evaluated in terms of the entropy of multivariate t-distributions. More specifically, the optimization criterion can be taken as

$$|\Psi_{add\,|\,curr}|,$$

the determinant of the prior mean covariance matrix of the sites (variables) to be *add*ed to the monitoring network conditional on the *curr*ently gauged sites. In their analysis, Guttorp et al. (1992a) adopted an empirical approach in specifying the mean Ψ of the inverted-Wishart prior for the spatial covariance matrix (over all the gauged and ungauged sites) using a spatial covariance model computed as described in Section 7.

10. Temporal and Multivariate Extensions

In this Section we comment on extensions of the analysis of heterogeneous spatial correlation structure to accommodate temporal correlation and multivariate data. Neither or these areas has been addressed in the literature to date.

10.1. Models for space-time correlation structure

The definition and interpretation of spatial covariance structure for environmental processes is closely linked to the temporal scale of the analysis. The discussion above has essentially assumed that replicate observations in time are independent, an assumption that is certainly inappropriate for many applications. For example, air quality data such as ozone are often recorded hourly and manifest short lag temporal correlation structure as well as very clear diurnal and weekly patterns. All of these components must be accommodated in any procedure that attempts to estimate ozone concentration at a point in time and a location (point or grid cell) in space. Geostatistical methods for dealing with temporal structure, such as those introduced by Seguret (1989), and Wackernagel (Wackernagel, 1988; Rouhani and Wackernagel, 1990), can now be considered in concert with current approaches to modeling heterogeneous spatial covariance structure. Temporal periodicities (such as the diurnal and weekly cycles noted above) may be considered part of a "trend" and filtered out of an analysis (Seguret, 1989), or they may be incorporated as components in the correlogram or temporal variogram (Rouhani and Wackernagel, 1990). For those components not filtered out one may use the framework of Rouhani and Wackernagel (1990) who modeled the cross-variogram relating observations at sites i and j,

$$\gamma_{ij}(\tau) = E\left[(Z_i(t)-Z_i(t+\tau))\times(Z_j(t)-Z_j(t+\tau))\right],$$

as a sum of cross-variograms at different temporal scales

$$\gamma_{ij}(\tau) = \sum_{u=1}^{S} \gamma_{ij}^{u}(\tau) = \sum_{u=1}^{S} b_{ij}^{u} g_u(\tau).$$

In this decomposition the elementary temporal variogram functions $g_u(\tau)$ must be conditionally negative definite and the matrices of coefficients b_{ij}^u, for fixed u, must be positive semi-definite. Periodic behavior, if not filtered out, is represented with "hole function" variograms. Analyses of ozone air quality might utilize one, two, or three components, representing short-term non-periodic covariance, diurnal, and weekly cycles, respectively.

The matrices of coefficients $B_u = [b_{ij}^u]$, called the **coregionalization matrices**, represent components of the spatial covariance structure at different temporal scales. Indeed, the spatial covariance matrix V can be expressed as a mixture of covariance structures at these different temporal scales

$$V = \sum_{u=1}^{S} B_u.$$

In order to compute spatial estimates at locations x that are not monitored, it is necessary to extend this model to express the components b_{ij}^u more generally as functions of spatial location, $b^u(x_i,x_j)$, permitting extrapolation to arbitrary locations x_a and x_b. The modeling approach of Section 7 can provide this necessary extension, although it has not yet been implemented in this context.

An extension of the spectral theory in Section 2 to space-time processes which are homogeneous in space and stationary in time is straightforward. Under the assumption of separable covariance, the spectrum factors into a product of a spatial and a temporal spectrum, each of which may be analyzed separately. However, if the assumption of separability is dropped the joint space-time spectrum can be a useful parameter function. To the extent that spectral estimation is straightforward (for example, if the spectrum has a given parametric form or if observations are available on a regular grid), these estimates can be applied in the D-space. North and Nakamoto (1989) use this to compare rain estimation designs for spatially homogeneous regions.

10.2. Multivariate models

Most environmental analysis problems are multivariate. In the study of acidic deposition three primary chemical species, hydrogen, sulfate and nitrate, are usually addressed. In some situations there may be only one particular variable of primary interest (e.g., rainfall), but even in these cases there are generally one or more important spatial covariates. Multivariate problems are discussed in the geostatistical literature under the heading of co-kriging (always assuming homogeneity). General modeling for heterogeneous covariance structure has not yet been addressed. We discuss here two different aspects of this problem.

For the multivariate case we will introduce a superscript and denote observations by $Z^k(x_i,t)$ for monitoring sites $i=1,2,\ldots,N$, times $t=1,2,\ldots,T$, and variables $k=1,2,\ldots,K$. From these data we may compute sample covariances $S_{ij,kl}$ corresponding to variables k and l measured at sites i and j, respectively. If all of the measured variables are driven by the same environmental process, then it may be reasonable to expect that they will all have the same spatial covariance structure. Write the observation at time t as a $K\times N$ random matrix $Z_t = [Z_i^k]$. Then,

$$\mathrm{Cov}(Z_i^k, Z_j^l) = S_{ij}\, C_{kl},$$

where S_{ij} are the elements of a spatial covariance matrix underlying all the variables and C_{kl} is the among-variables covariance matrix. In matrix notation the $NK\times NK$ covariance matrix factorizes as

$$\mathrm{Cov}(Z) = S \otimes C. \qquad (10.1)$$

Mardia and Goodall (1992) adopt this approach and, assuming that this factorized covariance structure holds for some power transformation of the measured variables Z^k, carry out maximum likelihood estimation of the transformation as in Box and Cox (1964). Notice that the factorized model above specifies a symmetric cross-covariance structure,

$$\mathrm{Cov}(Z_i^k, Z_j^l) = \mathrm{Cov}(Z_j^k, Z_i^l).$$

However, such cross-covariances are not necessarily symmetric, and for many applications this assumption is inappropriate.

Assuming that the structure (10.1) is valid one may use the techniques of Sections 7 and 8 to analyze and depict the spatial covariance structure underlying the matrix S. This approach clearly provides a simpler model for the full set of covariances and, if judged an appropriate model, a useful basis for further calculations such as co-kriging.

The first generalization to consider is the usual suggestion for an EOF analysis of multivariate spatial data (see, e.g., Preisendorfer, 1988, Chap 4). A principal components analysis is carried out on the $T \times NK$ data matrix consisting of rows given by the elements of the matrices Z_t as defined above. In this way all of the spatial covariances and cross-covariances are expressed in terms of a reduced number, say p, of dimensions and these covariances can be interpreted as in Section 6, as distances among points in p dimensions. However, now every site is represented by K points, one for each measured variable.

Along the lines of Section 7, we may instead consider spatial cross-dispersions

$$d_{ij}^{kl} = \mathrm{Var}(Z_i^k - Z_j^l),$$

allowing us to write

$$d_{ij}^{kl} = \mathrm{Var}(Z_i^k) + \mathrm{Var}(Z_j^l) - 2\,\mathrm{Cov}(Z_i^k, Z_j^l).$$

In the homogeneous case this has been called the pseudo cross-variogram (Myers 1991). As before, our modeling will be most sensible in terms of standardized variables so that the right-hand side is $2 - 2\,\mathrm{Corr}(Z_i^k, Z_j^l)$. The d_{ij}^{kl} may be considered pseudo-distances on the sites indexed by i and j (''pseudo'' since the corresponding $N \times N$ matrix is not generally symmetric). Consider the case of a bivariate spatial field, $K=2$. Although the $N \times N$ matrix $d^{12} = [d_{ij}^{12}]$ is not symmetric, we have $d^{12} = d^{21'}$ and the $2N \times 2N$ matrix

$$\begin{bmatrix} d^{11} & d^{21'} \\ d^{21} & d^{22} \end{bmatrix}$$

is symmetric. This suggests applying the multi-dimensional scaling approach to the $2N$ entries represented in this matrix, each monitoring site represented once for each of the two variables. There are now two maps or deformations to consider, one from the original geographic coordinates to the coordinates of the sites associated with measurement of field 1, $Z^1(x,t)$, and one to the coordinates of the sites associated with measurement of field 2, $Z^2(x,t)$. Interpolated mappings (using thin-plate splines or other procedures) could be used to evaluate the spatial cross-dispersion for sites other than those at which the two spatial fields were monitored.

In this case the deformation associated with measurement of field 1 is computed to represent both the spatial correlation structure of field 1 and the spatial cross-correlation structure from field 1 to field 2. Similarly, the deformation associated with field 2 represents the spatial correlation structure of field 2 together with the spatial cross-correlation structure from field 2 to field 1. A comparison of these two mappings, expressed as a comparison between the two multi-dimensional scaling representations of the sites (using a Procrustes-metric) would provide a basis for judging the appropriateness of the factorized model which assumes a common mapping for both the within-field spatial correlation structures and the between field cross-correlation structures.

This approach does not require the two fields to be monitored at the same sites. Field 1 may be monitored at N_1 sites and field 2 monitored at N_2 sites and the entire analysis carried out on the basis of the resulting square matrix of order N_1+N_2. It should be interesting to see what insight these procedures provide when applied to real data.

Acknowledgements

We are grateful to John Haslett for several comments that improved the presentation in this Chapter. Ken Newman did most of the analysis reported in Section 9.1, and helped with some of the figures. We also thank Gary Oehlert and Colin Goodall for providing preprints of their papers.

References

Adomian, G. (1983): *Stochastic systems*. Academic Press, New York.

Baddeley, A., J. Besag, H. Chernoff, P. Clifford, N. A. Cressie, D. J. Geman , B. Gidas, L. S. Gillick, N. Green, P. Guttorp, T. Kadota, A. Lippman and J. Simpson (1991): *Spatial Statistics and Digital Image Analysis*. National Academy of Sciences Press, Washington.

Bartlett, M. S. (1966): *Stochastic Processes*, 2nd ed. Cambridge University Press, Cambridge.

Böhning, D. (1982): Convergence of Simar's algorithm for finding the maximum likelihood estimate of a compound Poisson process. *Annals of Statistics* **10**, 1006-1008.

Bookstein, F. L. (1978): *The Measurement of Biological Shape and Shape Change*, Lecture Notes in Biomathematics (Vol. 24). Springer-Verlag, New York.

Box, G. P., and D. R. Cox (1964): An analysis of transformations. *J. R. Statist. Soc.*, Ser. B, **36**, 211.

Brillinger, D. R. (1972): The spectral analysis of stationary interval functions. In *Proc. Sixth Berkeley Symp. Math. Statist. Prob.* (L. M. Le Cam, J. Neyman, and E. L. Scott, editors), pp. 483-513. University of California Press, Berkeley.

Brillinger, D. R. (1981): *Time Series: Data Analysis and Theory*, expanded edition. Holden-Day, San Francisco.

Brillinger, D. R. (1985): Fourier inference: Some methods for the analysis of array and nongaussian series data. *Water Res. Bull.* **21**, 743-756.

Buell, E. C. (1972): Integral equation representation for factor analysis. *J. Atmos. Sci.* **28**, 1502-1505.

Buell, E. C. (1978): The number of significant proper functions of two-dimensional fields. *J. Appl. Meteor.* **17**, 717-722.

Caselton, W. F., and J. V. Zidek (1984): Optimal monitoring network designs. *Statist. Prob. Letters* **2**, 223-227.

Caselton, W. F., L. Kahn, and J. V. Zidek (1991): Quality data networks that minimize entropy. In *Statistics in Environmental and Earth Sciences* (A. Walden and P. Guttorp, editors), pp. 10-38. Edward Arnold, London.

Cohen, A., and R. H. Jones (1969): Regression on a random field. *J. Amer. Statist. Assoc.* **64**, 1172-1182.

Cressie, N. (1985): Fitting variogram models by weighted least squares. *Math. Geol.* **17**, 565-586.

Cressie, N. (1991): *Statistics for Spatial Data*. Wiley, New York.

Cressie, N., C. A. Gotway, and M. O. Grondona (1990): Spatial prediction from networks. *Chemometrics* **7**, 251-271.

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

Gandin, L. S. (1963): *Objective Analysis of Meteorological Fields*. Jerusalem: Israel Program for Scientific Translations.

Guttorp, P., N. D. Le, P. D. Sampson, and J. V. Zidek (1992a): Using entropy in the redesign of an environmental monitoring network. To appear in *Multivariate Environmental Statistics* (N.K. Bose, G.P. Patil, and C.R. Rao, editors). North Holland, New York.

Guttorp, P., P. D. Sampson and K. Newman (1992b): Nonparametric estimation of non-stationary spatial covariance structure with application to monitoring network design. In *Statistics in Environmental and Earth Sciences* (A. Walden and P. Guttorp, editors), pp. 39-51. Edward Arnold, London.

Haas, T. C. (1990a): Kriging and automated variogram modeling within a moving window. *Atm. Env.* **24A**, 1759-1769.

Haas, T. C. (1990b): Lognormal and moving window methods of estimating acid deposition. *J. Amer. Statist. Assoc.* **85**, 950-963.

Haslett, J. (1989): Space time modelling in meteorology—a review. *Bull. Internat. Statist. Inst.* **51**, book 1, 229-246.

Haslett, J., and A. E. Raftery (1989): Space-time modelling with long-memory dependence: Assessing Ireland's wind resource (with discussion). *J. R. Statist. Soc.*, Ser. C, **38**, 1-50.

Journel, A. G., and C. J. Huijbregts (1978): *Mining Geostatistics*. Academic Press, New York.

Krige, D. G. (1951): A statistical approach to some basic mine valuation problems on the Witwatersrand. *J. Chem. Metall. Mining Soc. S. Africa* **52**, 119-139.

Le, N. D., and A. J. Petkau (1988): The variability of rainfall acidity revisited. *Can. J. Statist.* **16**, 15-38.

Lesperance, M. L., and J. D. Kalbfleisch (1992): An algorithm for computing the nonparametric MLE of a mixing distribution. *J. Amer. Statist. Assoc.* **87:** 120-126.

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

Mardia, K. V., and C. R. Goodall (1992): Factorized models in spatial modeling. To appear in *Multivariate Environmental Statistics* (N.K. Bose, G.P. Patil, and C.R. Rao, editors). North Holland, New York.

Mardia, K. V., Kent, J. T., and Bibby, J. M. (1979): *Multivariate Analysis*. Academic Press, New York.

Matérn, B. (1986): *Spatial Variation* (2nd ed). Lecture Notes in Statistics (Vol. 36). Springer-Verlag, New York.

Matheron, G. (1963a), Principles of geostatistics, *Economic Geology*, 58, 1246-1266.

Matheron, G. (1963b): *Traité de Geostatistique Appliquée, Tome II: Le Krigeage*. Editions Technip, Paris.

Monestiez, P., and P. Switzer (1991): Semiparametric estimation of nonstationary spatial covariance models by metric multidimensional scaling. SIMS Tech. Rep. 165, Department of Statistics, Stanford University.

Myers, D. E. (1991): Pseudo-cross variograms, positive definiteness, and cokriging. *Math. Geol.* **23**, 805-816.

Newman, K., P. Guttorp, and P. D. Sampson (1992): A non-stationary spatial covariance structure for estimation of acid deposition. Technical Report, Department of Statistics, University of Washington.

North, G. R., and S. Nakamoto (1989): Formalism for comparing rain estimation designs. *J. Atm. Ocean. Tech.* **6**, 985-992.

Obled, Ch. and J. D. Creutin (1986): Some developments in the use of empirical orthogonal functions for mapping meteorological fields. *J. Appl. Meteor.* **25**, 1189-1204.

Oehlert, G. W. (1992): Regional trends in sulfate wet deposition. To appear, *J. Amer. Statist. Assoc.*.

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

Rouhani, S., and H. Wackernagel (1990): Multivariate geostatistical approach to space-time data analysis. *Water Resourc. Res.* **26**, 585-591.

Sampson, P. D. and P. Guttorp (1991): Power transformations and tests of environmental impact as interaction effects. *American Statistician* **45:** 83-89.

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

Sampson, P. D., S. Lewis, P. Guttorp, F. L. Bookstein, and C. Hurley (1991): Computation and interpretation of deformations for landmark data in morphometrics and environmetrics. *Proceedings of the 23rd Symposium on the Interface between Computing Science and Statistics*, pp. 534-541. Interface Foundation, Fairfax Station.

Schoenberg, I. J. (1938): Metric spaces and completely monotone functions. *Ann. Math.* **39**, 811-841.

Seguret, S. A. (1989): Filtering periodic noise by using trigonometric kriging. In *Geostatistics*, vol.1 (M. Armstrong, editor), pp. 481-491. Kluwer Academic, Dordrecht.

Switzer, P. (1989): Non-stationary spatial covariances estimated from monitoring data. In *Geostatistics*, vol.1 (M. Armstrong, editor), pp. 127-138. Kluwer Academic, Dordrecht.

Thiébaux, H. J. (1977): Extending estimation accuracy with anisotropic interpolation. *Mon. Weather Rev.* **105**, 691-699.

Thiébaux, H. J. (1991): Statistics and the environment: The analysis of large-scale earth-oriented systems. *Environmetrics* **2**, 5-24.

Thiébaux, H. J., and M. A. Pedder (1986): *Spatial objective analysis with applications in atmospheric science.* Academic Press, London.

Vong, R. J., L. Moseholm, D. S. Covert, P. D. Sampson, J. F. O'Loughlin, M. N. Stevenson, R. J. Charlson, W. H. Zoller, and T. V. Larson (1987): Statistical Analysis of Spatial and Temporal Variations in Urban Rainwater Chemistry: Changes in pH and Sulfate Associated with the Closure of a Copper Smelter. In *Acid Rain: Scientific and Technical Advances* (R. Perry, R.M. Harrison, J.N.B. Bell, and J.N. Lester, editors), pp. 90-97. Selper , London.

Vong, R. J., L. Moseholm, D. S. Covert, P. D. Sampson, J. F. O'Loughlin, M. N. Stevenson, R. J. Charlson, W. H. Zoller, and T. V. Larson (1988): Spatial and Temporal Variations in Urban Rainwater Chemistry: Changes in pH and Sulfate Associated with the Closure of a Copper Smelter. *J. Geophys. Res.* **93D6**, 7169-7179.

Wackernagel, H. (1988): Geostatistical techniques for interpreting multivariate spatial information. In *Quantitative analysis of mineral and energy resources* (C. F. Chung, editor), pp. 393-409. Reidel, Hingham.

Wahba, G. (1990): *Spline Models for Observational Data.* Society for Industrial and Applied Mathematics, Philadelphia.

Figure Captions

FIGURE 9.1. Geographic map of the UAPSP wet deposition network (left) and the deformation of a regular grid into the D-image (right) for hydrogen ions using winter data, computed using the method of Section 7.

FIGURE 9.2. Biorthogonal grid depicting the estimated spatial dispersion for hydrogen ions using winter data from the UAPSP wet deposition network. Line types encode the values of the principal strains of the local affine derivative matrix for the thin-plate spline mapping with solid lines indicating the directions of strongest correlation (smallest principal strain). Solid, long dashed, short dashed, and dotted lines are determined by splitting the range of principal strains (0.49 to 2.05) at 0.75, 1.00, and 1.55.

FIGURE 9.3. Raw correlation dispersions as a function of geographic distance (left) and D-image distance (right).

FIGURE 9.4. Estimated hydrogen field for March of 1983. The left panel is based on an isotropic Gaussian variogram, while the right is using a heterogeneous covariance estimate.

FIGURE 9.5. The kriging weights for the isotropic model (left) and the heterogeneous model (right). The weights are for estimating deposition in March, 1983, at the (unobserved) site marked with an X.

Figure 1

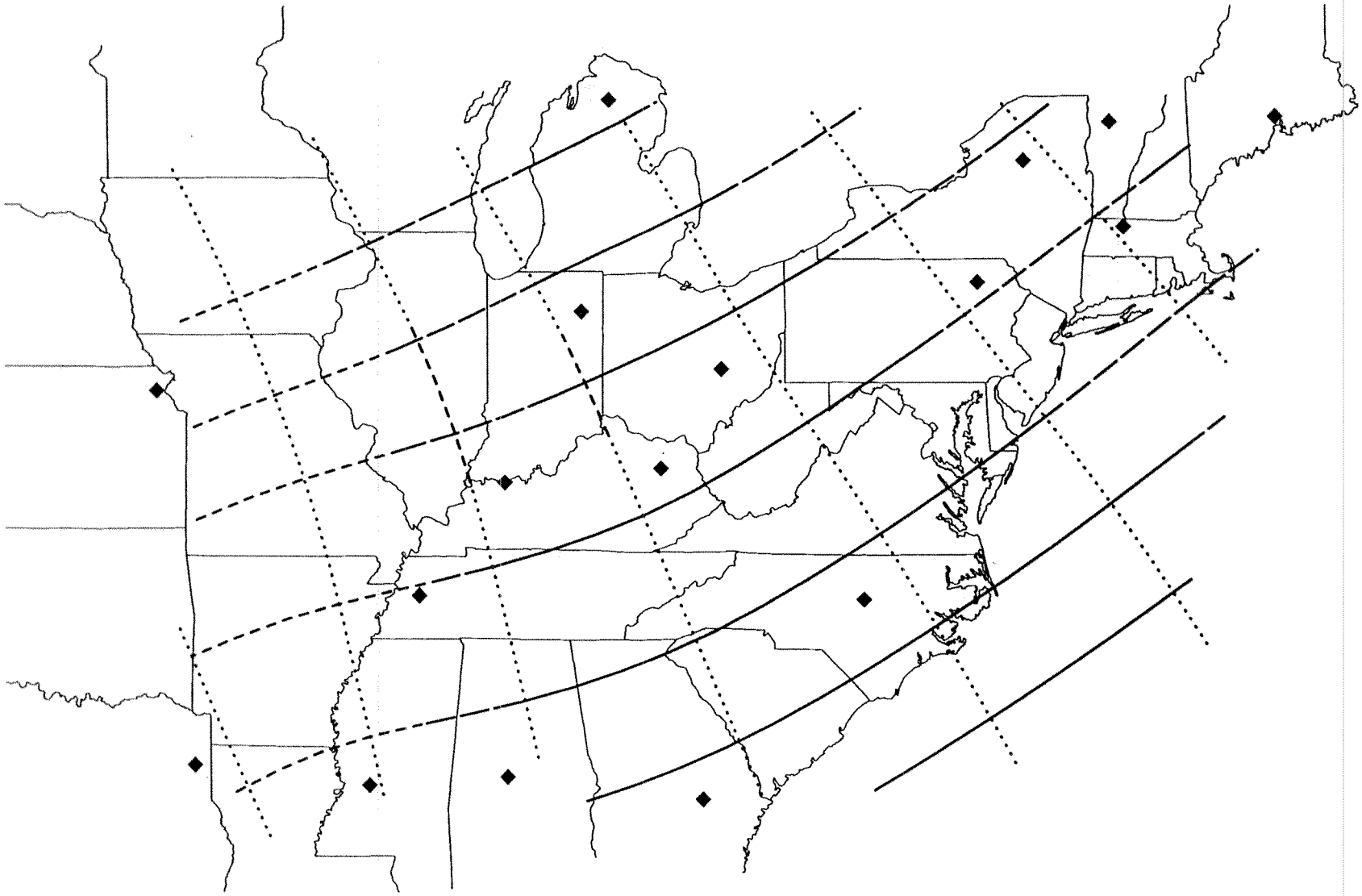

Figure 2

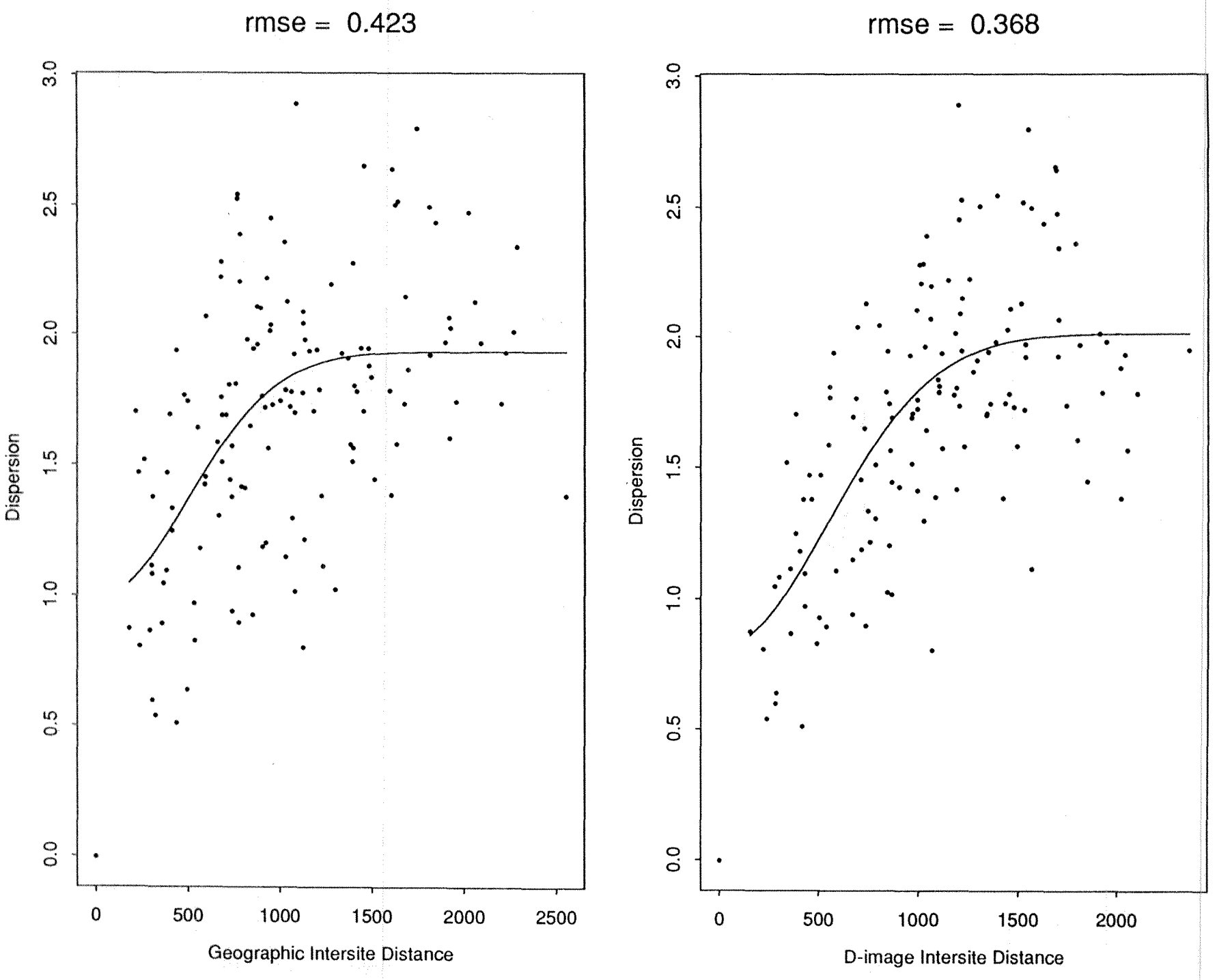
rmse = 0.423
Dispersion
0.0
0.5
1.0
1.5
2.0
2.5
3.0
0
500
1000
1500
2000
2500
Geographic Intersite Distance
rmse = 0.368
Dispersion
0.0
0.5
1.0
1.5
2.0
2.5
3.0
0
500
1000
1500
2000
D-image Intersite Distance

Figure 3

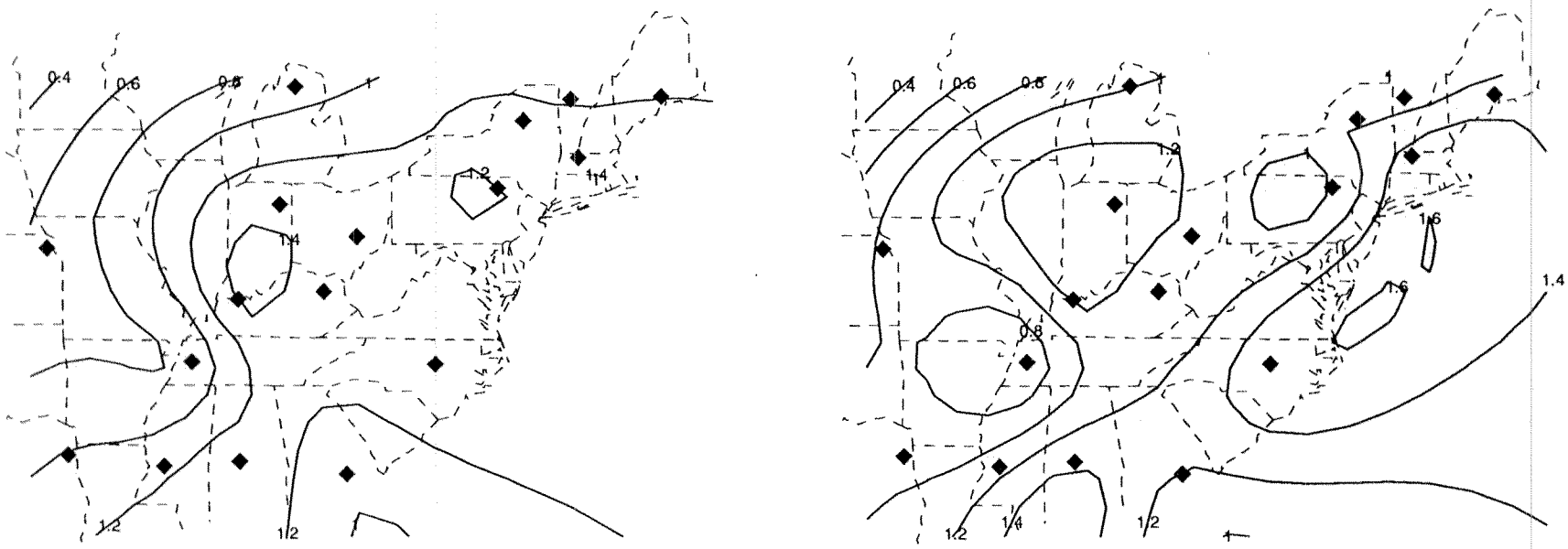

Figure 4

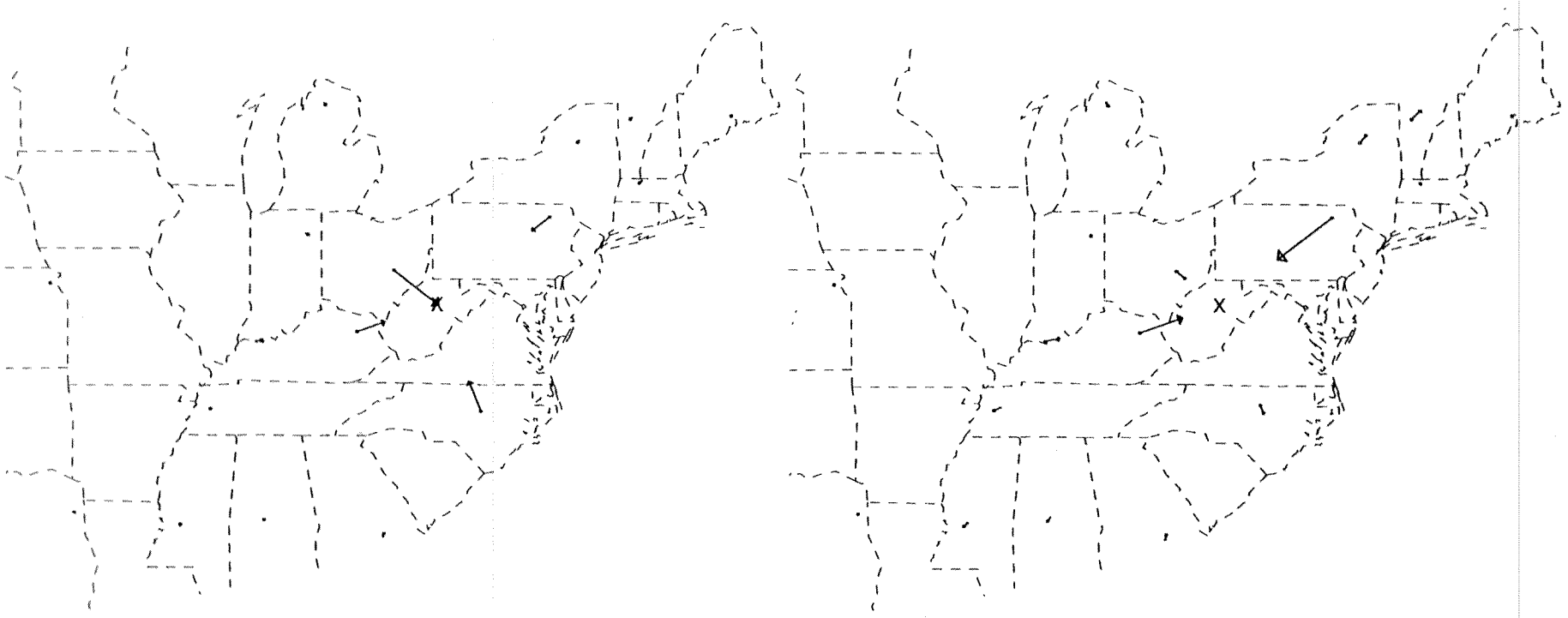

Figure 5