Computational Issues in Fitting Spatial Deformation Models for Heterogeneous Spatial Correlation

Peter Guttorp **Wendy Meiring** **Paul D. Sampson**

NRCSE

Technical Report Series

NRCSE-TRS No. 006

Computational Issues in Fitting Spatial Deformation Models for Heterogeneous Spatial Correlation

W. Meiring

Geophysical Statistics Project
National Center for Atmospheric Research
Boulder, CO 80307

P. Guttorp and P. D. Sampson

Statistics Department and NRCSE
University of Washington
Seattle, WA 98195

Abstract

Environmental monitoring networks are recording pollutant levels, weather, and a myriad of other factors. It is often of interest to estimate these values at locations where records are not available. Many spatial estimation procedures rely on spatial covariance models. Assumptions of spatial isotropy or stationarity may be violated due to factors such as topography and local emissions structures. In this paper we discuss computational issues for a heterogeneous (spatially non-stationary) model for spatial correlations between point monitoring sites. The modeling procedure involves deforming the geographic space into a new space (D-space) where inter-site correlations depend only on distances. Correlations between unmonitored sites are then estimated as a function of distance in the D-space. The estimation of the D-space locations of the monitoring sites, and of the parameters of the isotropic D-space variogram model is a difficult multidimensional problem. The dimensionality increases with the number of monitoring sites. We use examples to review the modeling approach and illustrate computational complexities. We indicate directions for future work necessary for modeling massive data sets.

1 Introduction

Spatial correlation estimates are frequently used in the estimation of values of environmental processes at spatial locations. Assumptions such as isotropy and spatial stationarity are often made. There may be local influences in the correlation structure however. For example, there may be pollutant sources that impact a small geographic region, or there may be differences in topography which influence the spatial process. Sampson and Guttorp (1992) and Guttorp and Sampson (1994) developed a heterogeneous spatial correlation modeling approach using spatial deformations. Other heterogeneous spatial correlation modeling approaches have been developed by Haas (1990 a,b; 1995), Loader and Switzer (1992), and Oehlert (1993).

In this paper we consider computational issues in the implementation of the spatial deformation approach, and directions for future work. In section 2, we review the estimation problem and the estimation approach developed by Sampson and Guttorp (1992) for $\Re^p \rightarrow \Re^d$ deformations, assuming that the spatial correlation is an isotropic function of distance in a d-dimensional deformation of the geographic space. We concentrate on the special case of $p = d = 2$ in describing the estimation procedure that we currently use, although higher-dimensional extensions are immediate. In section 3, we describe the optimization procedures and computational difficulties that we have encountered. Other approaches have been used for the estimation of the deformations, and section 4 references some of these, including the approach of Smith (1996). Section 5 describes directions for future work, including higher-dimensional mappings and challenges in analyzing satellite data.

Ten-day aggregated precipitation data for November and December at 20 sites in a French precipitation network will be used as an illustrative example in this paper. Figure 1 shows the site locations. This is a subset of the sites used in Monestiez et al. (1993) and Meiring et al. (1996).

2 Review of estimation problem

Suppose observations $Y(x_i, t)$ are collected at T time points $t \in \{1, \ldots, T\}$ at each of N monitoring sites $\{x_i : i = 1, \ldots, N\}$, which may be irregularly located in space. Suppose that $Z(x_i, t)$ represent the data that result from standardization by the site specific temporal variances. In this paper we assume that the observations are independent in time, and that the spatial correlation is constant in time.

We model spatial association in terms of the variance of the spatial increments, which we call the *spatial dispersion function*, defined as

$$D(u, w) \quad = \quad \mathrm{var}[\, Z(u,t) \; - \; Z(w,t) \,] \tag{1}$$

for sites u and w. We will only use the terminology *variogram* (c.f., Cressie, 1991) when we are referring to dispersion models which depend only on the Euclidean distance between sites (isotropic models), and will otherwise use the term "dispersion". A dispersion of 2 corresponds to zero correlation on the variance-standardized scale.

The estimation approach introduced by Sampson and Guttorp (1992) models correlation structure as an isotropic function of distances after a deformation of the geographic coordinate system. We refer to the geographic space or coordinate system as the G-space (or G-plane for $p = 2$) and the deformation space as the D-space (or D-plane for $d = 2$). The model is of the form

$$d_{ij} \quad = \quad \gamma_\theta \left(\| f(x_i) - f(x_j) \| \right) + e_{ij} \tag{2}$$

where d_{ij} is the sample dispersion between geographic sites x_i and x_j, γ_θ is a valid isotropic variogram with parameters θ, and $f(x_i)$ and $f(x_j)$ are coordinates in the D-space, corresponding to geographic locations x_i and x_j; e_{ij} is an error term. The form of γ_θ must be chosen—for example γ_θ might be an exponential variogram model with nugget—and the variogram parameters θ and the D-plane coordinates must be estimated. In the example in this paper, we use an exponential D-plane variogram with nugget. For the variance-one spatial field this is of the form

$$D(u,w) = \begin{cases} a_0 + (2 - a_0)\{1 - \exp(-t_0\, h_{uw})\} & \text{if } h_{uw} > 0 \\ 0 & \text{otherwise} \end{cases}$$

where a_0 is the nugget effect representing small scale variability and measurement error, t_0 is a scaling parameter, and h_{uw} is the distance between sites u and w.

We currently use a penalized weighted least squares criterion in estimating the variogram parameters and D-space coordinates of the monitoring sites. If there are N monitoring sites, the criterion to be minimized is

$$\sum_{j=2}^{N} \sum_{i=1}^{j-1} \left[\frac{d_{ij} - \widehat{d_{ij}}}{\widehat{d_{ij}}} \right]^2 + \lambda \, \mathrm{BEP}, \tag{3}$$

where d_{ij} is the sample dispersion, and $\widehat{d_{ij}}$ is the fitted dispersion between the i^{th} and j^{th} monitoring sites, calculated as $\widehat{d_{ij}} = \gamma_{\hat\theta}\left(\left\| \hat f(x_i) - \hat f(x_j) \right\| \right)$. BEP is proportional to the "bending energy" for the transformation $f(.)$ of the G-plane to the D-plane and λ is a smoothness parameter controlling the variance-bias trade-off.

For $p = d = 2$ we compute the bending energy penalty as the sum of the bending energy for two $\Re^2 \to \Re$ thin-plate spline mappings:

$$f_1 : (x, y) \to x_D \quad \text{and} \quad f_2 : (x, y) \to y_D$$

where (x, y) is a geographic location, and (x_D, y_D) is the corresponding D-plane location. The bending energy of each mapping f_i is

$$\int_{\Re^2} \left[\left(\frac{\partial^2 f_i}{\partial x^2} \right)^2 + 2 \left(\frac{\partial^2 f_i}{\partial x \partial y} \right)^2 + \left(\frac{\partial^2 f_i}{\partial y^2} \right)^2 \right] dx\, dy. \tag{4}$$

It may be calculated as a quadratic form in the new coordinates or as a quadratic form in the spline coordinates (c.f., Wahba, 1990). We use the formulae of Mardia et al. (1991), namely

$$[f_i(x_1), \ldots, f_i(x_N)]\, \mathbf{B}\, [f_i(x_1), \ldots, f_i(x_N)]^T, \tag{5}$$

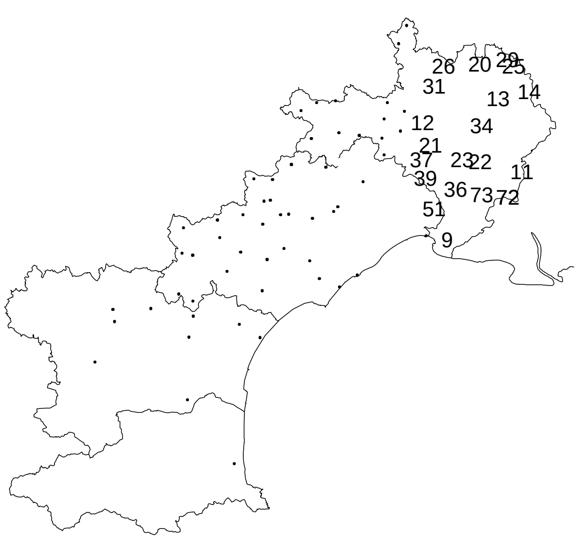

Figure 1: Network of precipitation monitoring sites in a region of France. Data from the 20 numbered sites are used in this paper.

where

$$\begin{aligned} \mathbf{B} &= [(I-A)\,K\,(I-A)]^{-} \\ K_{ij} &= \begin{cases} 0 & \text{if } i=j \\ h_{ij}^2 \log\left(h_{ij}^2\right) & \text{if } i \neq j \end{cases}, \\ P_{i.} &= \left[\begin{array}{ccc} 1 & x_i(1) & x_i(2) \end{array}\right], \\ A &= P(P'P)^{-1}P', \end{aligned}$$

and $[\,]^{-}$ denotes the Moore-Penrose generalized inverse (see for example Pringle and Rayner, 1971). The expression given by (5) is proportional to (4).

Meiring et al. (1996) illustrate the variance-bias trade-off in choosing λ in criterion (3) and present a cross-validation study of choosing λ. For small λ, one may fit the sample dispersion estimates too closely, not allowing for variability in these estimates. As λ increases, the model tends to a stationary spatial correlation structure with elliptical contours of constant dispersion. Figure 2 illustrates the effect of changes in λ on the fitted dispersions and deformations using an exponential variogram model with nugget as the D-plane variogram for the precipitation data. (The exponential variogram is almost linear in the range of dispersions shown.) Pairwise dispersions between site 72 and other sites are highlighted, indicating substantial bias in the estimation of these for $\lambda = 1000$.

The cross-validation simulation experiment of Meiring et al. (1996), minimizes the cross-validation criterion

$$\sum_{i \in \mathcal{N}} \sum_{j \in \mathcal{N}-\{i\}} \left(\frac{d_{ij} - \gamma_{\hat{\theta}_{i\lambda}}(\hat{f}_{i\lambda}(x_i) - \hat{f}_{i\lambda}(x_j))}{\gamma_{\hat{\theta}_{i\lambda}}(\hat{f}_{i\lambda}(x_i) - \hat{f}_{i\lambda}(x_j))} \right)^2 \quad (6)$$

with respect to λ. The estimates subscripted by $i\lambda$ are based on observations from all sites except x_i, using λ as the smoothing parameter. $\mathcal{N} = \{1, \ldots, N\}$. The mean square error between the fitted and the true dispersions was lower when the fitted dispersions were based on λ chosen by cross-validation, than when choosing λ by a simpler stopping rule, or than when using a homogeneous spatial dispersion model corresponding to an infinitely large value of λ. Due to the complexities of optimizing criterion (3), cross-validation may prove too computationally intensive for use in routine analyses. This points to the need for efficient estimation routines and careful consideration of computational issues.

3 Implementation and computational issues

For a specific choice of λ, we use an algorithm which alternates between estimation of the D-plane coordinates for fixed variogram parameters, and estimation of the variogram parameters for fixed D-plane coordinates. Two sites are fixed at their geographic coordinates in order to fix the scale, orientation and location of the D-plane configuration. The matrix $\mathbf{B}$ in (5) depends only on the geographic locations of the monitoring sites so remains constant throughout the optimization.

Estimation of the few variogram parameters for fixed monitoring site locations is substantially less complex than the optimization over the monitoring site locations. However, standard optimization considerations such as scaling and starting values are vitally important. Figure 3 shows the objective surface as a function of variogram parameters for an exponential variogram with nugget, keeping the monitoring sites at their geographic locations. Such plots may be considered to decide on the starting values for variogram parameter estimation, and also on the scaling of the coordinates so that the local minimum does not appear to fall along a sharp ridge in the objective surface. The vertical axis on the left of Figure 3 represents the t_0 parameter scaled by 10^{-3} to make it more comparable with the nugget parameter a_0 on the horizontal axis. In terms of the actual t_0 coordinates on the right vertical axis the objective surface would appear to have an extremely long ridge.

In the optimization step for the D-plane monitoring site locations, the objective surface is highly complex and is often multimodal with points of inflection. For $\Re^2 \to \Re^2$ mappings with N monitoring sites the surface is defined over a $2N-4$-dimensional parameter space (after fixing two monitoring sites sites at their G-plane locations in order to fix the orientation, scale and location of the D-plane up to reflection of the entire plane). Figure 4 shows multimodality in the objective surface for $\lambda = 0$. Contours are plotted of the objective surface for a grid of locations of site 23 with the variogram parameters fixed at their optimized values and the other monitoring sites fixed at their (optimal) D-plane locations, which are indicated by solid triangles. The geographic and D-plane locations for site 23 are indicated. Note that peaks in the objective surface at other monitoring sites prevent sites from getting too close in the D-plane. However, there may be multiple local minima around the D-plane, as indicated by the dotted lines corresponding to an objective value of 3.5.

Problems of multi-modality are reduced as λ increases, since the mapping approaches an affine deformation in the limit. Figure 5 separately displays the weighted least squares and bending energy components of criterion (3) over a grid of site locations for site 72. Site 72 was chosen for this figure since it is relatively isolated, allowing clearer presentation. The objective surface of the weighted least squares component is highly

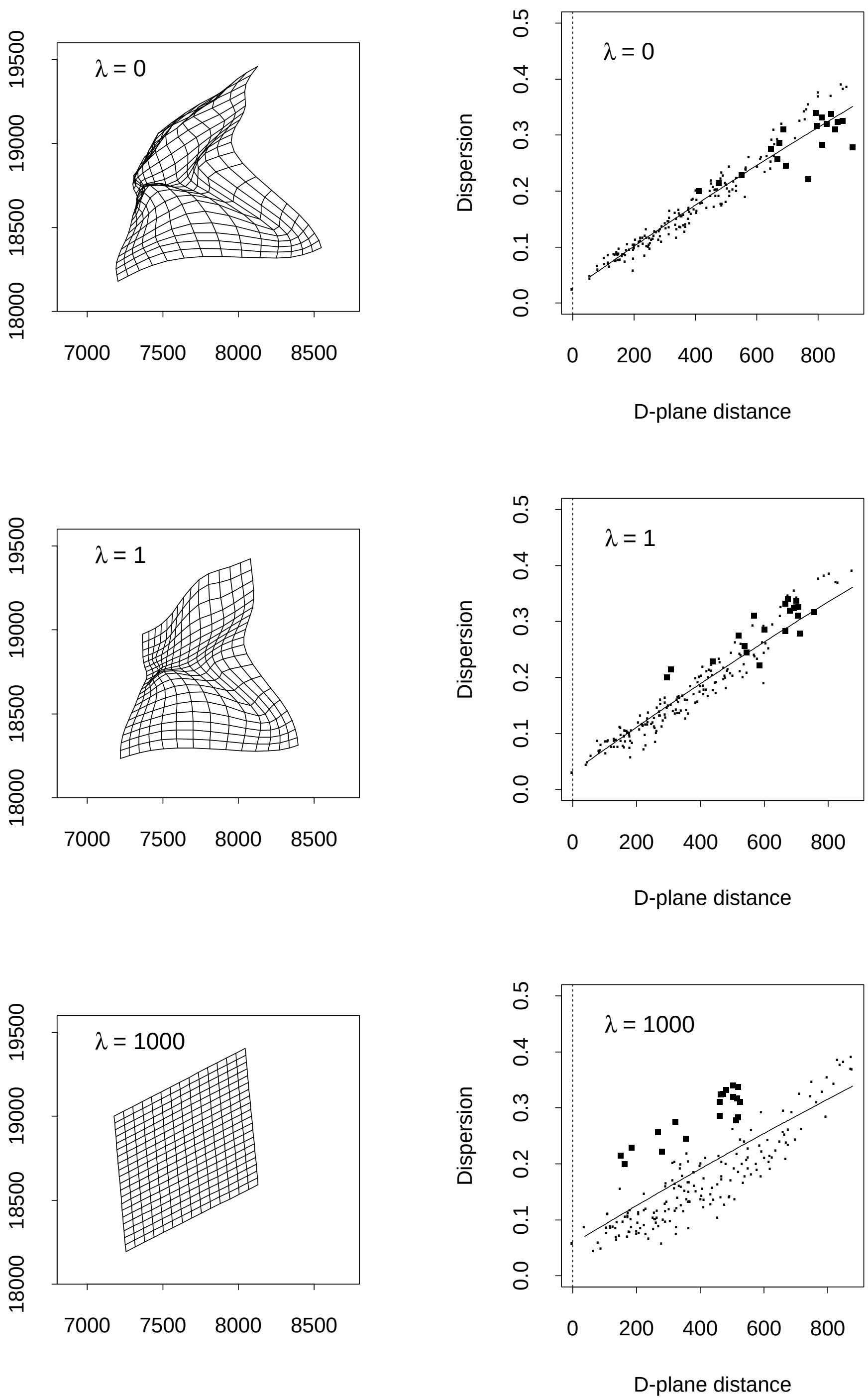

Figure 2: Deformations and fitted variograms for different values of λ for 20 sites in a French precipitation network. Dispersions involving site 72 are highlighted.

complex, whereas the surface of the bending energy component is much simpler. As λ increases, the bending energy penalty component begins to dominate the optimization. The final solution corresponds to a weighting between the optimum of the bending energy component and the weighted least squares component, with the weighting determined by λ. Similar plots for $\lambda = 1000$ showed that the optimized D-plane location for site 72 corresponded with the minimum of the bending energy component, at which stage we had fitted a stationary dispersion model.

As with many cases of multi-modality, it is difficult to create diagnostics for the existence of multiple modes, or to explore the objective surface to search for these modes. We have demonstrated plots of the objective surface over a grid of D-plane locations for a monitoring site while keeping the variogram parameters and D-plane locations for the other sites fixed at their optimized values. However, such plots are not sufficient as diagnostics since the set of N (conditional) surfaces does not represent all the modes for the $2N-4$ dimensional space.

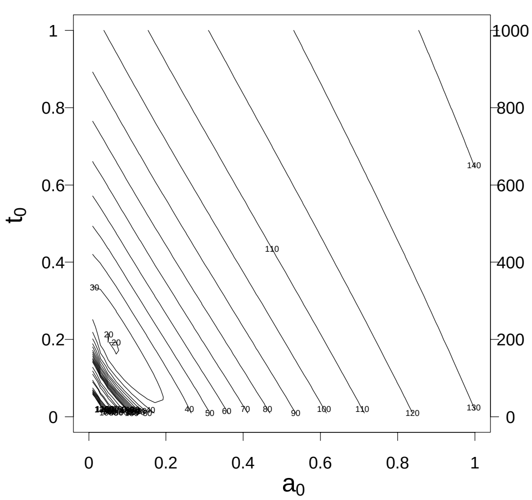

Figure 3: Objective surface over a grid of variogram parameter values, corresponding to coordinates fixed at their geographic locations. The vertical axis on the left represents the t_0 parameter scaled by 10^{-3} to make it comparable with the scale of the nugget parameter. The vertical axis on the right represents the actual coordinates of t_0.

We begin the optimization of the locations for the monitoring sites at their geographic locations, since we model the spatial dispersions as a smooth function of geographic location. For this reason we are perhaps most interested in modes that are near the original geographic locations, which are least likely to result in a mapping that folds (is not bijective). Repeating the optimization with different initial locations for the monitoring sites, may aid in investigating multi-modality. However, we would not expect a reasonable transformation mapping if the starting values were too far from the original geographic locations.

Clearly, the complexity of the optimization also depends on how well-defined the minima are in the objective space. In Guttorp et al. (1997), we illustrated that the spatial correlation of pre-whitened and demeaned transformed hourly ozone residuals, is diurnally varying. The optimization routines converge relatively quickly for the (day to day) correlation structure of the afternoon hours, but not for the morning hours. For the morning hours all of the spatial correlations are close to zero (positive and negative) and the fitted variogram is relatively flat. During the morning hours, the optimization procedure cannot distinguish between a flat variogram with a nugget close to 2 (corresponding to zero correlation) or a fitted variogram with a low nugget, but a sharp rate of increase at distances less than the minimum intersite distance.

3.1 Current optimization approach

Our current implementation uses the NPSOL 4.0 nonlinear optimizer developed by Gill, Murray, Saunders and Wright (1986), or the routines E04UCF and U04UEF in the NAG Fortran Library. These optimizers use a sequential quadratic programming (SQP) approach (see Gill et al. 1981, and references therein). They are designed to solve problems of the form

$$\min_{v \in \Re^n} F(v) \quad \text{subject to} \quad l \leq \left\{ \begin{array}{c} v \\ A_L v \\ c(v) \end{array} \right\} \leq u, \qquad (7)$$

where $F(v)$ is the objective function as a function of the variables over which we are optimizing. Both steps of the alternating algorithm can be formulated in this way. In the variogram parameter estimation step of the alternating algorithm, v is the vector of variogram parameters. In the D-plane step of the algorithm, v is a vectorized form of the site coordinates. We have only upper and lower bound constraints on the variables over which we optimize, so A_L and $c(v)$ are set to zero. In the variogram parameter estimation step of the algorithm, the

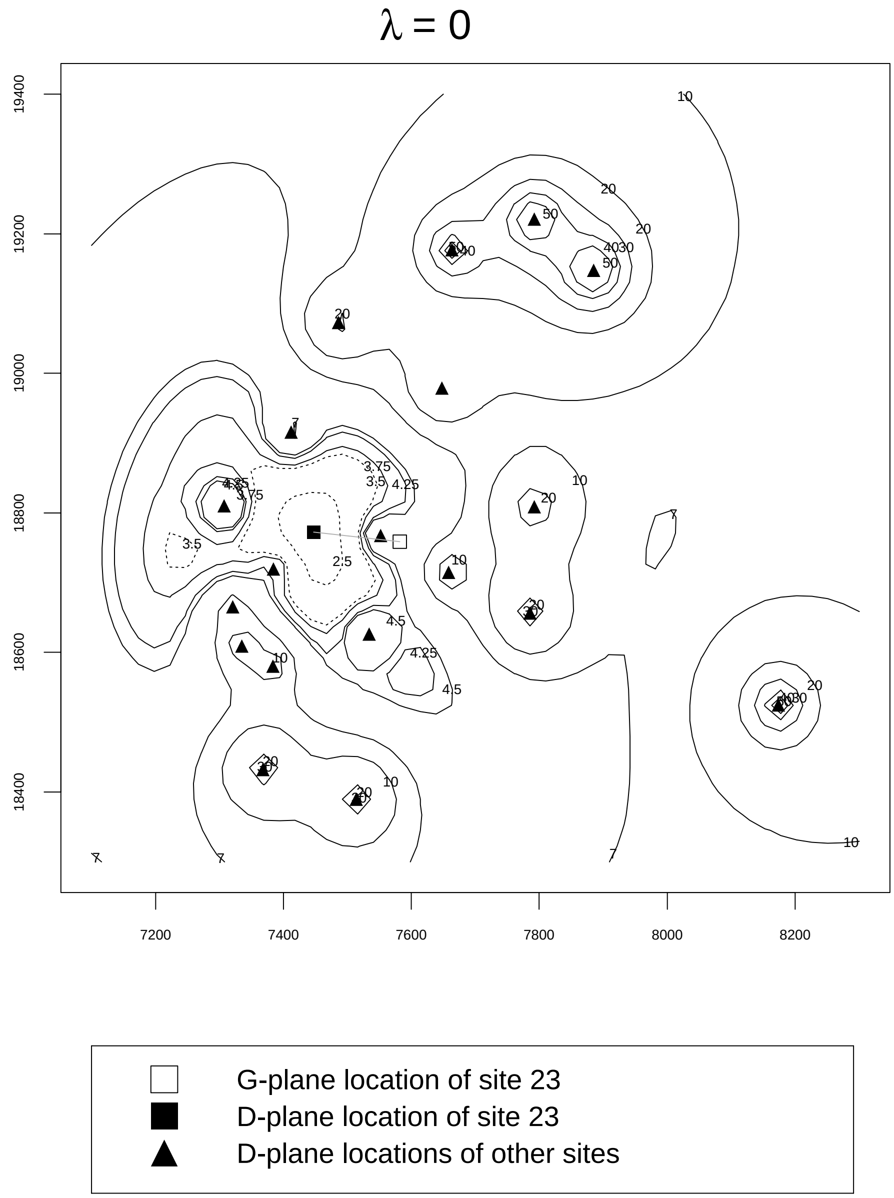

Figure 4: Objective surface for $\lambda = 0$ over a grid of locations for site 23, keeping other sites fixed at their D-plane locations. Two local minima below 3.5 are indicated by the dotted lines.

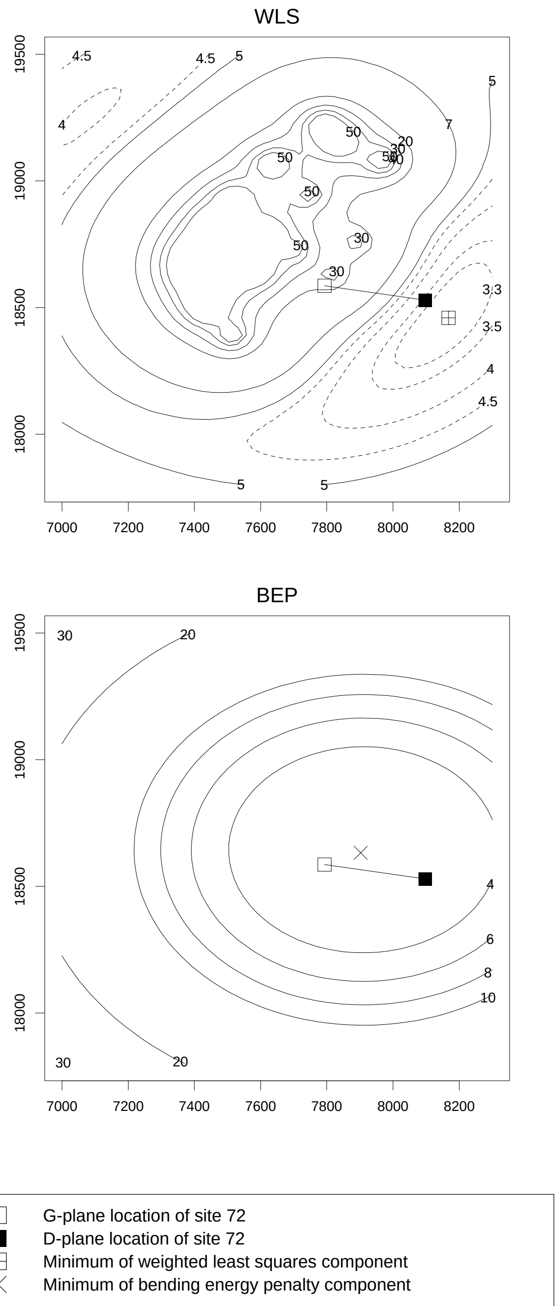

Figure 5: Surfaces over a grid of values for site 72, separately for the weighted least squares and bending energy penalty components for criterion (3). The geographic location and the D-plane location for site 72 are indicated, as are the minima for each of the weighted least squares and bending energy penalty components. The D-plane location is seen to be between the minima of the two individual components, with the weighting determined by λ, which equals 1 in this example.

linear constraints consist of bounds on the variogram parameters. For example the nugget effect a_0 is constrained to lie between 0 and 2. In the D-plane step of the algorithm, the linear constraints are equality constraints on the two site locations which are fixed.

We encountered a number of optimization problems while implementing NPSOL. Several of these were related to it being a gradient based optimizer. When we ran NPSOL without specifying analytical derivatives, the D-plane coordinate solution was particularly poor for large values of λ in the bending energy penalty. This may be due to the automatic choice of finite difference intervals in the approximation, possibly resulting in the algorithm stepping over the minimum when evaluating steps in the appropriate search direction. We thus specified the gradients with respect to the coordinates analytically, and we usually specify the gradients with respect to the variogram parameters analytically as well.

4 Alternative approaches

The concept of isotropy as a function of distance in a deformation of the geographic space (allowing local correlation structure as a function of location in the geographic space), is the most important idea in this heterogeneous spatial correlation modeling approach. We have described an estimation approach using a penalized weighted least squares criterion and two thin-plate spline mappings. Other estimation approaches have been suggested, and continued work is needed to investigate approaches which will be computationally feasible for massive data sets.

Our original computational approach used a nonmetric multidimensional scaling algorithm for the calculation of D-plane coordinates, but this calculation, followed by possible smoothing of the implied G-plane to D-plane mapping, did not represent the optimization of any objective criterion (Sampson and Guttorp, 1992). It would be natural to consider maximum likelihood or Bayesian approaches. Smith (1996) proposes a maximum likelihood approach calculating mappings using only subsets of the radial basis functions underlying the thin-plate spline. Much work remains to be done here.

5 Discussion

The exponential variogram models used in this paper are monotonically increasing and tend to an asymptote corresponding to zero correlation as the distances between sites increases. Atmospheric phenomena may exhibit negative correlations between regions; for example the surface pressure in a region in the eastern tropical Pacific is negatively correlated with that of Djakarta in Indonesia (Wilks, 1995, page 59). These negative correlation patterns are known as teleconnections. Hole function variograms are non-monotonic and allow negative correlations. We have used hole function variograms in Sampson et al. (1994). Further study is needed on properties and interpretation of the deformations when the variogram is non-monotone.

We have considered mainly the situation where the spatial correlation is an isotropic function of distance in a two dimensional deformation of the two dimensional geographic plane. There are instances where this may not be reasonable. For example, sites on either side of a pollutant plume may be more strongly correlated with each other than they are with sites within the pollutant plume. Two-dimensional plane deformations based on monotone variograms may exhibit folding in these situations. We are considering the applicability and interpretation of some extensions of our current methodology for these problems. The simplest approach is to consider non-monotonic variograms. Another approach is to use $\Re^d \rightarrow \Re^p$ mappings where p is larger than d. We expect that these higher dimensional mappings will also prove useful by allowing more local spatial structure in the deformation than is currently accommodated by the global bending energy penalty when λ is set sufficiently high to eliminate any large scale folding of the mapping.

One of the most difficult problems in geostatistics is how to estimate the nugget effect if one doesn't have collocated observations. This is also one of the most important problems, since the nugget is important when estimating at spatial locations that are close to monitoring sites. In the absence of collocated sites, there is no information about the correlation of the spatial process at distances smaller than the smallest intersite distance. Small-scale variability and measurement error cannot be distinguished without collocated observations. In practice we still estimate the nugget, but sensitivity to this estimate should be investigated. This is also an essential consideration when considering the choice of D-plane variogram, since different classes of variograms may give similar fits over the range of the data, but yet give very different estimates of the nugget. This once again highlights the need for statisticians to be involved in monitoring network design.

Underlying any application of the deformation model for spatial correlation structure is the implicit assumption that the monitoring sites are sufficiently densely located so as to represent the heterogeneity in the underlying ("true") correlation structure as represented by the nonlinear mapping. This model is motivated by the assumption that factors such as predominant wind directions and topography influence the empirical correlation in measurements of environmental processes at different points in space. We do not currently know how best to model these effects explicitly, but we expect that they will be manifest in the nonlinear deformations. So it is important that the monitoring sites adequately represent the spatial variation in these factors presumed to underly the spatial covariance structure. Thus, for example, if a mountain range is expected to influence not just the mean, but the spatial covariance structure of a space-time process, then we should not extrapolate the spatial covariance structure over the mountain from a sample of monitoring sites on either side of the range.

Systematic biases due to monitoring network design may be avoided by taking advantage of remote sensing data—for those environmental parameters that can be measured from space. Satellites are providing massive data sets. They may alleviate some data sparsity problems (subject to consideration of the size of the "footprint" that satellite observations represent), but they present many new computational challenges due to the sheer volume of data and due to the fact that satellite paths have particular distributions in time and space—in contrast to the regular temporal sampling from fixed monitoring networks. The development of methods for fitting these spatial deformation models from satellite data, and for some applications, the integration of satellite data with surface monitoring data, pose challenging research problems.

We have outlined our current estimation approach and computational considerations in its implementation. Current challenges include the development of better computational algorithms that will make cross-validation exercises more feasible on a routine basis and the development of tools to guide the optimization in view of the multimodality of the objective surfaces for this high-dimensional estimation problem. In addition, development of procedures permitting the application of the spatial deformation model to massive remote sensing data sets would be extremely valuable.

Acknowledgements

The authors are grateful for funding from the Electric Power Research Institute through contract #8010-01; from the NCAR Geophysical Statistics Project which is sponsored by the National Science Foundation under Grant #DMS93-12686; and from the National Research Center for Statistics and the Environment which is funded by the United States Environmental Protection Agency. Although the research described in this article has been funded in part by the United States Environmental Protection Agency through agreement CR825173-01-0 to the University of Washington, it has as not been subjected to the Agency's required peer and policy review and therefore does not necessarily reflect the views of the Agency and no official endorsement should be inferred. We thank Pascal Monestiez of INRA for providing us with the precipitation data.

References

[1] Cressie, N.A.C. (1991). *Statistics for spatial data.* Wiley-Interscience.

[2] Gill, P.E., W. Murray, M.A. Saunders, and M.H. Wright (1986). User's Guide for NPSOL (Version 4.0): A Fortran Package for Nonlinear Programming. *Technical Report SOL* 86-2. Systems Optimization Laboratory, Department of Operations Research, Stanford University.

[3] Gill, P.E., W. Murray, and M.H. Wright (1981). *Practical Optimization*. Academic Press.

[4] Guttorp, P., and P.D. Sampson (1994). Methods for Estimating Heterogeneous Spatial Covariance Functions with Environmental Applications. *In* Patil, G.P. and C.R. Rao (eds), *Handbook of Statistics XII: Environmental Statistics*, Elsevier/North Holland, New York, pp. 663-690.

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

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

[7] 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.

[8] Loader, C., and P. Switzer (1992). Spatial covariance estimation for monitoring data. *In* A. Walden and P. Guttorp, eds., *Statistics in the Environment and Earth Sciences.* Edward Arnold, London, 52-70.

[9] Mardia, K.V., J.T. Kent, and A.N. Walder (1991). Statistical shape models in image analysis. *Computing Science and Statistics. Proceedings of the 23rd Symposium on the Interface*, pp. 550-557.

[10] Meiring, W., P. Monestiez, P.D. Sampson, and P. Guttorp (1996) Developments in the modelling of nonstationary spatial covariance structure from space-time monitoring data. Submitted to *Proceedings of the 5th International Geostatistics Congress.* E.Y. Baafi and N. Schofield (eds).

[11] Monestiez, P., P.D. Sampson, and P. Guttorp (1993). Modelling of heterogeneous spatial correlation structure by spatial deformation. *Cahiers de Geostatistique*, Fascicule 3, Compte Rendu des Journees de Geostatistique, 25-26 May 1993, Fontainebleau. Published by the École Nationale Supérieure des Mines de Paris.

[12] NAG Fortran Library. Numerical Algorithms Group, Limited.

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

[14] Pringle, R.M., and A.A. Rayner (1971). *Generalized inverse matrices with applications to statistics.* Griffin's statistical monographs and courses, no. 28. New York, Hafner Pub. Co.

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

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

[17] Sampson, P.D., P. Guttorp, and W. Meiring (1994). Spatio-temporal analysis of regional ozone data for operational evaluation of an air quality model. *Proceedings of the Section on Statistics and the Environment, American Statistical Association.*

[18] Smith, R.L. (1996). Estimating Nonstationary Spatial Correlations. *Unpublished manuscript on the Spatial Statistics Preprint Service at* http://hp2.niss.rti.org/organization/personnel/granville/list.html

[19] Wahba, G. (1990). *Spline models for observational data.* SIAM:PA.

[20] Wilks, D.S. (1995). *Statistical Methods in the Atmospheric Sciences.* Academic Press.