Cahiers de Géostatistique, Fascicule 3
Compte–rendu des Journées de Géostatistique,
25–26 Mai 1993, Fontainebleau, pp. 35–46

MODELING OF HETEROGENEOUS SPATIAL CORRELATION STRUCTURE BY SPATIAL DEFORMATION

Pascal MONESTIEZ, Paul D. SAMPSON * and Peter GUTTORP *

Laboratoire de Biométrie,
I.N.R.A.
84140 Montfavet

(*) Department of Statistics,
University of Washington,
Seattle, WA. USA

RESUME. Soit $Z(x,t)$ un champ aléatoire où x représente les coordonnées géographiques dans $\mathcal{R}^2$ et t le temps. On suppose connue une série d'observations du champ aléatoire sur un ensemble fini de sites de monitoring, $x_i, i = 1,2,\ldots,n$. Les résultats présentés concernent une classe de modèles basés sur la déformation spatiale du support du champ aléatoire, $x \rightarrow y \in \mathcal{R}^k, k \geq 2$, afin de rendre $Z(y(x),t)$ stationnaire et isotrope en y. La méthode comprend l'estimation de la déformation $y(x)$ au travers d'un algorithme de multidimensional scaling appliqué aux sites de monitoring, ainsi qu'une estimation du modèle de variogramme du champ aléatoire dans l'espace déformé. Nous nous intéresserons plus particulièrement au choix du paramètre controllant la déformation de l'espace par validation sur un jeu de stations test. L'exemple présenté concerne l'analyse de 20 ans de pluviométrie sur 75 sites du Languedoc Roussillon.

ABSTRACT. Let $Z(x,t)$ denote a random field as a function of geographic coordinates $x \in \mathcal{R}^2$ and time t. We assume repeated observations of the random field at a finite number of monitoring sites, $x_i, i = 1,2,\ldots,n$. We will discuss recent results concerning a modeling approach based on a deformation of the spatial domain of the random field, $x \rightarrow y \in \mathcal{R}^k, k \geq 2$, for which $Z(y(x),t)$ becomes a stationary and isotropic random field when considered as a function of y. The method involves the estimation of the deformation $y(x)$ by metric non-linear multidimensional scaling of the monitoring sites x_i together with estimation of a variogram model for the random field represented in the transformed spatial domain. The focus of this paper is a demonstratiom of a validation analysis for the choice of a parameter controlling the smoothness of the deformation $y(x)$. We present an analysis of 20 years of rainfall data from the Languedoc Roussillon region of France.

Table des Matières

1 Introduction.

The analysis—modelling and estimation—of the spatial covariance structure underlying spatial-temporal environmental processes has been the subject of increasing research in recent years. Knowledge of the spatial covariance of a random field is fundamental in spatial estimation or kriging and in the design of optimal monitoring networks. We have been concerned primarily with environmental monitoring data whose stochastic characteristics are driven by atmospheric processes; for example, rain fall, acid precipitation, and air pollution. But the issues and methods we discuss here are important also to data analysis in many other environmental fields including, for example, hydrology, geophysics, and oceanography. Particularly important is the fact that the spatial covariance structures underlying such processes are almost always nonstationary (or, nonhomogeneous) over the spatial scales of interest. By this we mean that the stochastic relationship (covariance or correlation) between time series of observations of, say, rainfall monitored at two sites x_a and x_b separated by a geographic vector $h = x_a - x_b$ will depend on the absolute location of these two monitoring sites—how they are located with respect to various topographic or orographic features such as coastlines and mountain ranges, and how they are located with respect to the flow of weather systems. A pair of sites separated by a vector of length $|h|$ oriented along the path of weather systems will have observations more highly correlated than a pair of sites separated by a vector of length $|h|$ oriented perpendicular to the path of weather systems. Furthermore, the strength of that correlation may depend, for example, on how close the sites are to coastlines.

The fact that environmental processes display correlation structures that are not spatially stationary has been recognized in some fields of study for a number of years. See, for example, Journel and Huijbregts [5] for discussion in the "geostatistics" literature and Thiébaux [13], [14] for discussion in the meteorological literature. However, only recently have models been proposed for the analysis and estimation of nonstationary spatial covariance structure. For a review, see Guttorp and Sampson [3]. The methods introduced to date have left a number of important practical issues unresolved, and as a result, they are not yet widely used in applied studies. In this paper we begin to address some of the most important practical problems underlying a methodology introduced by Sampson and Guttorp [11], [3] and further developed by Monestiez and Switzer [9]. Mardia and Goodall [7] have also considered this approach. In particular, we discuss the generality of the model, the problem of choice of the dimensionality of the model, and demonstrate a validation analysis for the choice of a smoothing parameter critical in fitting the model to data.

2 The nonstationary spatial covariance model.

We consider space-time processes representable as

$$Z_G(x,t) = \mu(y(x),t) + \sigma(y(x)) \times Z(y(x),t), \tag{1}$$

where x denotes geographic coordinates and t time. $\mu(y(x),t)$ represents the mean field generally varying in space and often also in time because of periodic structure or long-term trends. $\sigma^2(y(x))$ is the variance field and $Z(y,t)$ is a zero mean, unit variance, space-time process that is stationary in time t and both stationary and istropic as a function of "location" $y(x)$.

The expression of the process given in (1) indicates the fact that we aim to determine a new coordinate system $y = y(x)$ in which the spatial covariance structure of Z is stationary

or homogeneous. While the geographic coordinate space is typically R^2, the dimension of the new space of coordinates will be two or more with $y(x)$ representing a smooth mapping or deformation of the geographic coordinates. An interpretation of this coordinate system is given below in terms of a real example.

Note that $Z(x,t)$ is not stationary in x. But in terms of the new coordinate system $Z(y,t)$ has a stationary, isotropic covariance structure, which may be represented with a (parametric) covariance model $c(h)$ where h denotes distance between monitoring sites in the y space. The covariance model must be positive definite. Very broad families of covariance functions may be considered using, for example, results presented in Matérn [8]. In practice we have been using mixtures of a small number (one, two, or at most three) exponential or gaussian covariance functions. For example, a mixture of two exponential correlation functions can be written

$$\rho(h) = \alpha_1\rho_0(b_1h) + \alpha_2\rho_0(b_2h), \quad h > 0,$$

where $\rho_0(h) = exp\{-h\}$ and $\alpha_1 + \alpha_2 \leq 1$. Note, however, that interpretation of these as nested correlation models corresponding to processes operating at different spatial scales (b_1 and b_2) is generally not convenient because this correlation model applies to the process represented in a transformed coordinate system. In particular, the "range" parameter often discussed in geostatistics would be difficult to interpret. (For this reason, we do not usually employ the "spherical" covariance or variogram model common in geostatistics.)

One often hopes to determine a scale of measurement on which the process Z can be assumed to have constant variance. However, this is not always possible and the model (1) accomodates a heterogeneous variance field. In this case it is most convenient to separate the analysis of the variance field from analysis of the correlation field, although this approach is not pursued in the example presented below.

Let $Z_{it} = Z_G(x_i,t)$ denote the obserations at monitoring sites x_i, $i = 1,\ldots,N$, at times t, $t = 1,\ldots,T$. Let $\bar{z}_i$ denote the empirical mean of the Z_{it} over t for site x_i and let c_{ij} denote the empirical covariance for the pair of sites (i,j):

$$c_{ij} = \frac{1}{T}\sum_{t=1}^{T}(Z_{it} - \bar{z}_i)(Z_{jt} - \bar{z}_j). \tag{2}$$

The c_{ij} may be considered as "similarities" relating the sampling sites. Adopting a framework closer to that in geostatistics, we may also carry out analysis in terms of a transformation of these covariances to "dissimilarities" or "distances"

$$\delta(x_i,x_j) = var(Z(x_i,t) - Z(x_j,t)). \tag{3}$$

The corresponding empirical quantities can be computed as

$$d_{ij} = c_{ii} + c_{jj} - 2 \times c_{ij}. \tag{4}$$

We have called this the *spatial dispersion function* ([11], [3]). In the geostatistics literature (where variances and covariances are *not* computed over replicates in time) stationarity is often assumed and $\delta(x_i,x_j) = \gamma(x_i - x_j)$ is called the *variogram*. Our assumption of an isotropic correlation structure in the transformed space means that we can write

$$\delta(x_i,x_j) = \gamma(|y_i - y_j|). \tag{5}$$

This equation defines the class of models we consider, with the variogram function $\gamma()$ based on the families of mixtures of covariance functions described above. Because transformation of the geographic space is viewed as a multidimensional scaling problem using the *dispersion* function as a metric or distance measure on the monitoring sites, we refer to the coordinate space of $y's$ as the *D-space* in contrast to the geographic *G-space*.

3 Semiparametric estimation by least squares.

Current approaches to estimating the deformation model (5) have been framed in terms of the minimization of a weighted or simple sum of squares with respect to the parameters θ underlying the variogram (mixture) model and the image points $\{y_i\}$:

$$\min \; S^2(y_1, ..., y_n, \theta) = \sum_{i<j} [d_{ij} - \gamma_\theta(|y_i - y_j|)]^2. \tag{6}$$

(In fact, [9] and [7] express the optimization in terms of the covariances c_{ij} rather than the dispersions d_{ij}.) This is conveniently minimized using an alternating least squares calculation, minimizing first with respect to variogram parameters θ for given coordinates $\{y_i\}$, and then with respect to the coordinates for given variogram parameters, iterating to convergence. We refer to this as semiparametric estimation because the deformation is essentially represented nonparametrically in terms of image coordinates for the monitoring sites conditional on a parametric variogram model. An analytic form extending the mapping of the site coordinates, $x_i \rightarrow y_i$, to the entire geographic domain of interest has been computed only *post-hoc* after calculation of the $\{y_i\}$ using either thin-plate splines [3], [7], or "bi-kriging" with a linear drift [9].

As noted above, our model, (1) and (5), assumes a mapping of the original space of geographic coordinates: $x \rightarrow y(x)$. In fact, we assume this mapping to be continuous and bijective so that two distinct geographic locations are not mapped into the same point in D-space. Sampson and Guttorp ([11], [3]) further assume the mapping is differentiable, or a diffeomorphism, in representing it with thin-plate splines and visualizing it using biorthogonal grids [12]. However, when estimating the coordinates y_i as parameters of the model by a least-squares criterion such as (6), the y_i are not constrained so that the relationship between the original x_i and estimated y_i, extended to the whole domain of analysis by some reasonable form of interpolation, can be represented by such a mapping. In fact, because of sampling error in the spatial dispersions d_{ij} on which the estimation is based, computed mappings $y(x)$ will often "fold." That is, they will not be bijective. We provide an example below.

To obtain a bijective relationship it has been necessary to *smooth* the y_i as a function of the geographic coordinates x_i. This has been done using either a kernel smoothing procedure [9] or smoothing splines [3]. The kernel smoothing procedure is used in the following example.

The kernel smoothing procedure computes a transformation of the y_i effectively shrinking them toward the original geographic coordinates:

$$y_i' = x_i + \frac{\sum_{j=1}^N (y_j - x_j) \times exp\,(-\frac{1}{2}|x_i - x_j|^2/s_j^2)}{\sum_{j=1}^N exp\,(-\frac{1}{2}|x_i - x_j|^2/s_j^2)}. \tag{7}$$

One could minimize criterion (6) defined in terms of the y_i', but the smoothing becomes ineffective if the y_i are absent from the criterion and so themselves are not constrained to fit the dispersions. Supposing that a set of y_i' are accurate estimates of the coordinates

in a "true" model, the y_i deviate from these due to noise or uncertainty in the empirical spatial dispersions. We therefore considered a minimization criterion defined in terms of the y_i but balancing contributions from both the y_i and the y_i' as defined above.

$$\min \; S^2(y_1, ..., y_n, \theta) = \sum_{i<j} [d_{ij} - \gamma_\theta(|y_i' - y_j'|)]^2 + \lambda \sum_{i<j} [d_{ij} - \gamma_\theta(|y_i - y_j|)]^2. \qquad (8)$$

In fact, the solution in y_i' is not very sensitive to the value of λ over a range giving adequate weight to both parts. The value of λ was set at 0.5 for all minimizations reported here. Note, however, that although the quantity in (8) was used for minimization, only the value of the first part of (8) is reported for comparisons in the following sections.

The value of smoothing parameters s_j in (7) could be kept constant for all j, but this is not effective in dealing with varying density of monitoring sites. The s_j were therefore specified to be locally proportional to the distances to the nearest neighbor and the smoothing parameter s reported in the example below is the mean value of these s_j.

4 Example: A rainfall network in southern France.

As an example we present the result of fitting the spatial dispersion model to rainfall data in southern France. Time series of ten-day rainfall totals over more than 20 years were collected in the region of Languedoc Roussillon. The monitoring sites, pictured in

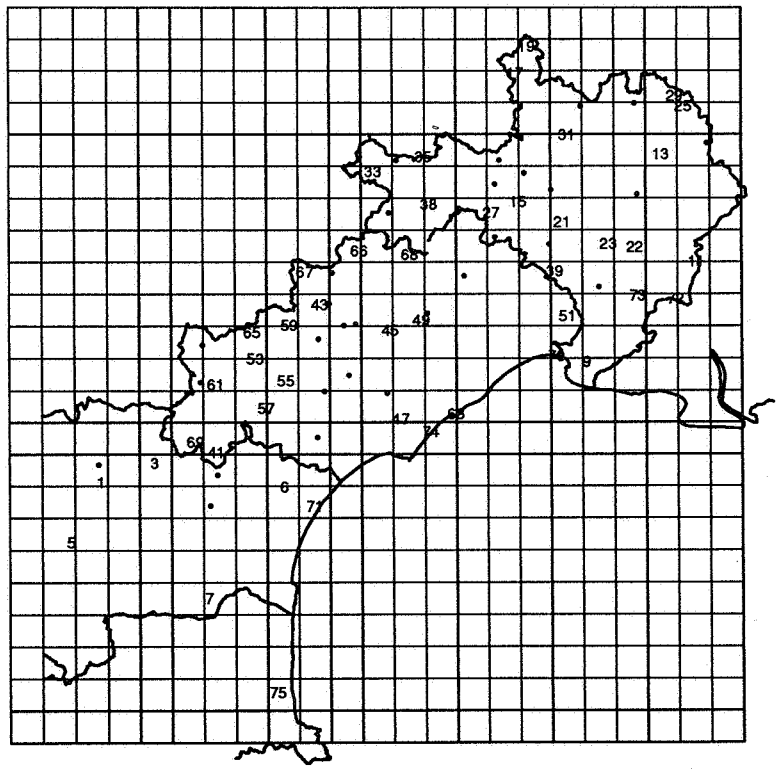

Figure 1: Map of the studied region of Languedoc Roussillon. The numbers denote the 45 rain gauge sites; the dots mark the 30 sites used for validation of the fitted model. The regular grid has cells of 10 km by 10 km and will be used to represent the deformation.

Figure 1, belong to three French departments: Aude, Herault and Gard, and cover an area from the Mediterranean littoral to the southeast mountain range of the Massif Central.

For purposes of model validation the monitoring sites were partitioned into two sets. One set of 45 rain gauges used for fitting the model for the spatial dispersion function and a second set of 30 rain gauges, located approximately inside the convex hull of the first set, was reserved for validation.

Preliminary analysis and transformation of the time series was necessary in order to obtain site means and variances that were reasonably homogeneous over time and over the whole region. First, the time series were deseasonalized by subtraction of a periodic seasonal term. We considered both local seasonal terms for each site as well as a common seasonal adjustment for all sites in the region. The differences among sites were small and so we chose a common seasonal adjustment for all sites. This also simplifies reconstruction of the series after prediction of rainfall at unmonitored sites.

Second, modeling an altitude effect was necessary to correct for local variation in site means and variances. After deseasonalization, the cumulated rains over ten-day periods show site means clearly linearly proportional to the corresponding standard errors, suggesting a multiplicative correction term. Referring to equation (1), we now write:

$$Z_G(x,t) = \mu(t) + K(h(x)) \times Z(y(x),t), \tag{9}$$

where $\mu(t)$ is a periodic term, $h(x)$ the altitude at location x, and $Z(y(x),t)$ a standardized series for altitude zero.

The term $K(h(x))$ was deduced from the regression

$$log(Z_G(x,t) - \mu(t)) = a + b\ h(x) + c\ \sqrt{h(x)} + \varepsilon \tag{10}$$

and has the form

$$K(h(x)) = \exp\left(b\ h(x) + c\ \sqrt{h(x)}\right). \tag{11}$$

This expression represents a local adjustment and has no general meaning. The corrected

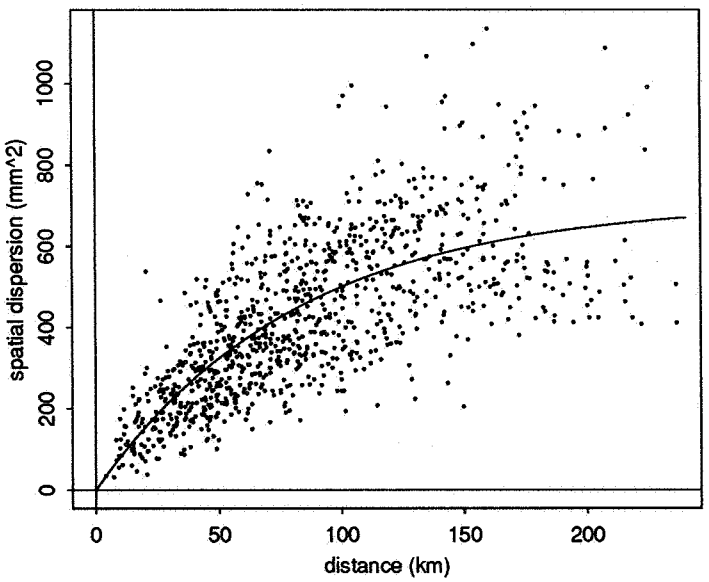

Figure 2: Plot of the geographic inter-site distances $|x_i - x_j|$ in km vs empirical spatial dispersions d_{ij} and the fitted exponential variogram model.

series $Z(x,t)$ are no longer mean-centered as in (1), but after altitude correction, site means and variances are reasonably homogeneous over space and the resulting $Z(x,t)$

may be interpreted as potential rain over the region at hypothetical sea level (ignoring the fact that the topographic variations are partly responsible for the fact of rainfall as well as for variation in mean amount).

Figure 2 shows a plot of geographic inter-site distance vs the corresponding empirical spatial dispersions d_{ij} — to which one would fit a variogram model if one assumed stationarity and isotropy for $Z(x,t)$. There is considerable scatter in this figure and it seems appropriate to apply our algorithm in order to reduce it.

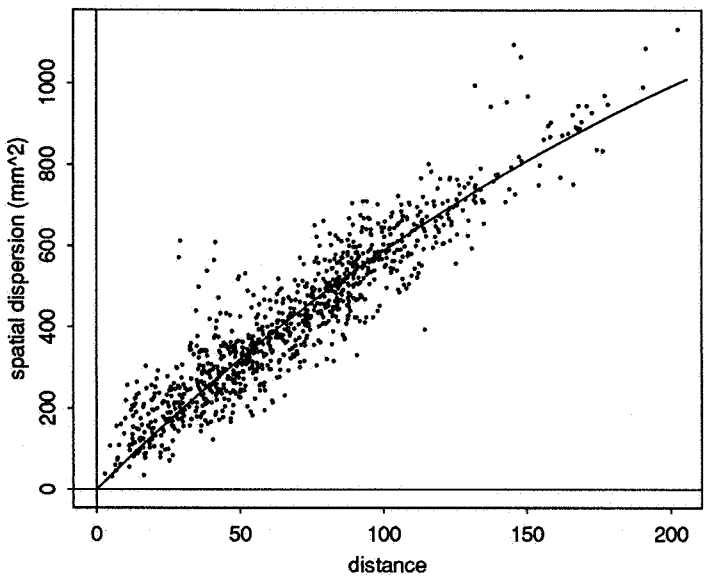

Figure 3: Plot of the D-plane inter-site distances $|y_i - y_j|$ vs empirical spatial dispersions d_{ij} and the fitted exponential variogram model.

We first execute the alternating optimization algorithm to fit the spatial dispersion model by minimization of (6) without smoothing. Figure 3 presents a plot of distance in the D-plane vs the empirical spatial dispersions. The vertical scatter about the fitted exponential variogram is now dramatically reduced.

Figure 4 shows the corresponding coordinates y_i and the mapping $x_i \rightarrow y_i$ for the 45 sites (fixing sites 6 - Narbonne and 22 - Nimes as reference points). The mapping represented by the arrows clearly shows some spatial coherence for a majority of sites, but it is also clear that some site displacements will generate foldings for any interpolation of these pointwise correspondences to the entire region.

Several procedures might be considered to "solve" the problem of folding. One solution is to use a higher dimensional D-space. However, that approach would result in a model fitting the observed spatial dispersions even more closely than the fit represented in Figure 3. Because we believe that Figure 3 represents, in fact, some degree of overfitting of the variogram, we use instead the smoothing algorithm described the previous section. First we empirically adjust the smoothing parameter until the foldings, manifest in depictions of interpolated mappings computed as bi-thin-plate splines, completely vanish. We achieve this with a smoothing parameter of approximately $s = 40km$ (as defined at the end of Section 3); both the spatial dispersion plot and the corresponding interpolated mapping of the 45 sites (computed using thin-plate splines) are displayed in figure 5.

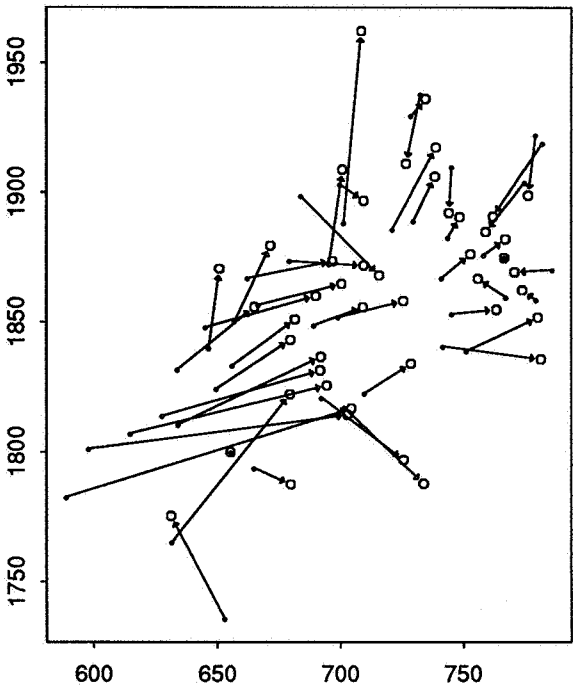

Figure 4: Plot of the spatial deformation from the x_i (black dots) to the y_i (circles). The coordinate system is a Lambert two standard parallel projection with units in km.

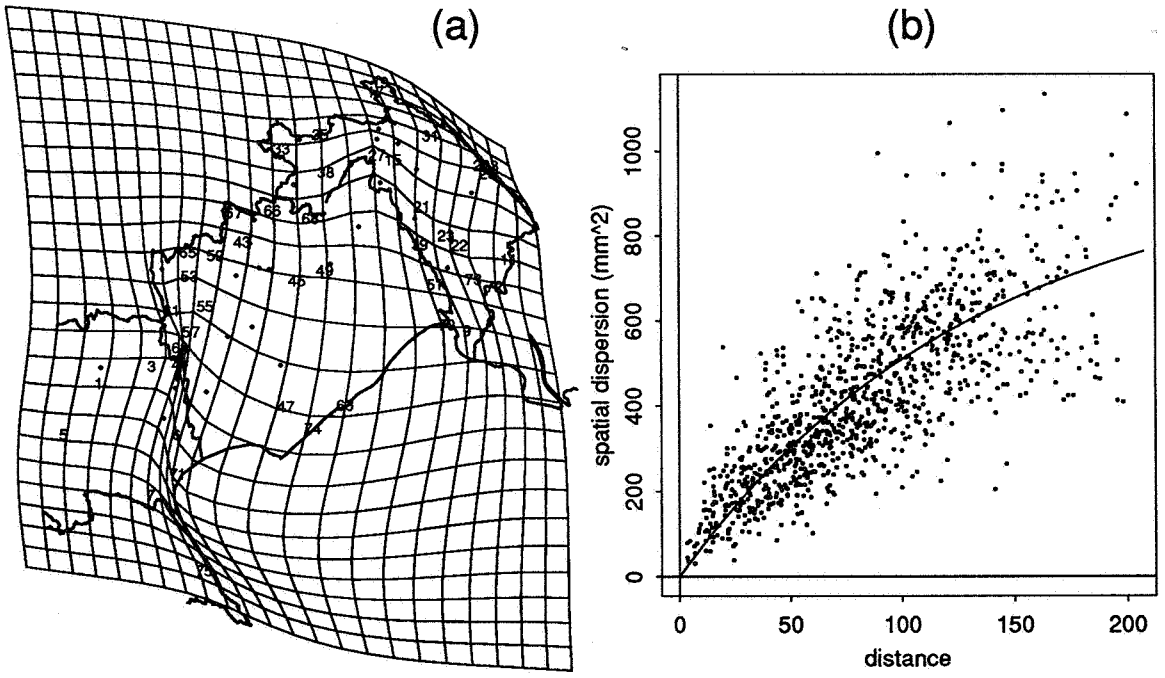

Figure 5: (a) Image of Figure 1 grid and department contours in the D-plane for the model fitted with smoothing parameter $s = 40km$. Site locations correspond to the y_i'. The deformation of the contours and the nodes of the square grid was computed using a thin-plate spline mapping the x_i in Figure 1 into the y_i'. (b) Inter-site distances $|y_i' - y_j'|$ vs empirical spatial dispersions d_{ij} and the fitted exponential variogram model.

5 Choice of smoothing parameter by validation on a test set.

There are two important issues in the specification of a fitted model: choice of dimension for the D-space representation and the amount of smoothing. In Figure 6 we show, in contrast to Figure 5, the thin-plate spline deformation computed from the model fitted with smoothing parameter $s = 20km$. Small folding in the mapping is evident. The square grid in Figure 1 is folded over on itself in the region encompassing sites 61, 57, 3, and 6. Increasing this parameter "smooths out" this fold (as shown in Figure 5) at the expense of increasing slightly the scatter in the plot of dispersion versus distance. We consider here calculations relevant to the choice of a smoothing parameter. We note first that the smoothing parameter must be chosen large enough to eliminate folds in the mapping (as diagnosed by visual inspection of figures such as Figures 5 and 6). If this is not achieved by the values of s suggested by the following calculations, then a higher-dimensional model for the D-space coordinates is necessary.

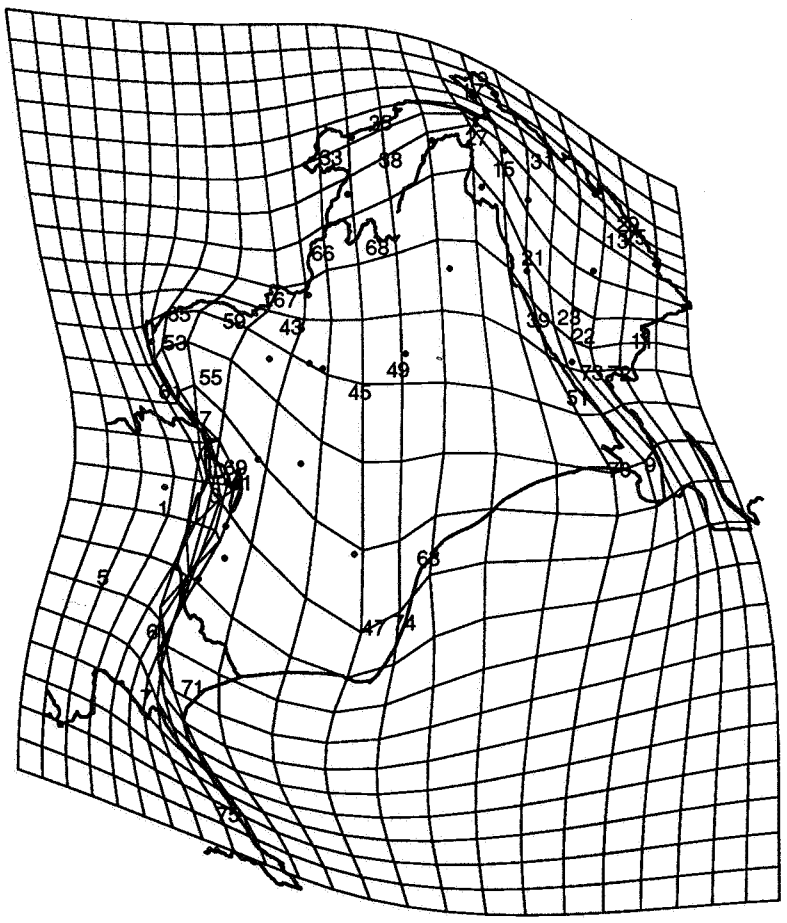

Figure 6: Image of Figure 1 grid and department contours in the D-plane for the model fitted with smoothing parameter $s = 20km$. Site locations correspond to the y_i'.
The deformation of the contours and the nodes of the square grid was computed using a thin-plate spline mapping the x_i in Figure 1 into the y_i'.

We have assumed that a smooth deformation is capable of describing the nature of the non-stationarity in the covariance structure. One can imagine topographic or orographic features that might sharply alter the orientation of the principal directions of strongest

and weakest spatial correlation. It must therefore be assumed that the spatial density of the monitoring network is sufficiently dense to capture the spatial features of interest. In general terms, if, for example, topography is the important factor influencing the local spatial correlation structure, then the monitoring network must *comprehensively* sample the variation in topography.

Assuming the monitoring sites are in fact *redundant* with respect to the spatial scale of nonstationarity, we may consider classical methods of validation and cross-validation to choose a smoothing parameter.

1. Set aside a subset of (at least two) sites chosen at random. Estimate the spatial dispersion (covariance) model on the remaining sites and use it to predict the intersite dispersions (covariances) involving those sites initially set aside. Choose s to "optimize" this prediction by some weighted least squares criterion.

2. Set aside pairs of sites, two at a time, and compute corresponding predicted intersite dispersion (covariances). Choose s to minimize the average prediction error over all pairs of sites.

The second of the two cross-validation methods is highly computationally intensive. Here we present only the results of a validation analysis of the first type.

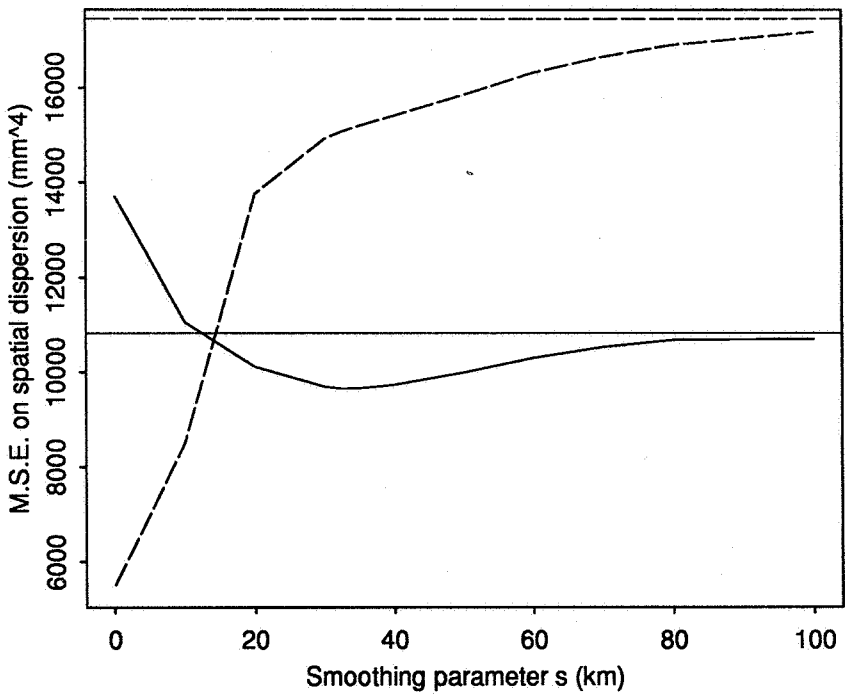

Figure 7: Mean Square Errors of spatial dispersion vs smoothing parameter s. Dashed lines represent the MSE of fit among all pairs of the 45 sites. Solid lines represent the MSE of prediction among all pairs of the 30 test sites. Horizontal lines, dashed and solid, represent the corresponding MSEs for an isotropic model, i.e., no deformation.

In our analysis of the Languedoc Roussillon rainfall data, 30 of the 75 gauge sites were set aside from the beginning for validation. These were chosen randomly, subject to the condition that they lied inside or near the convex hall of the entire set. Their locations are identified by dots in Figure 1. For each of 12 values of the smoothing parameter s,

the spatial dispersion model (5) was fitted to the 45 remaining sites. We interpolated the correspondences of the geographic sites x_i with the deformed locations y_i using thin-plate splines (as in [11] or [12]) in order to compute a mapping $y(x)$ everywhere and so predict the values of the spatial dispersion (or variogram) for the 435 pairs determined from the 30 test sites. Figures 5 and 6 show the interpolated locations of the test sites for the models fitted with smoothing parameters $s = 40$ km and $s = 20$ km, respectively.

We computed mean squared errors (MSE) to compare the predicted spatial dispersions among the 30 test sites with the observed values. The solid curve in Figure 7 shows this MSE as a function of the smoothness parameter. It can be seen that the MSE first drops, reaching a minimum for a smoothness parameter of approximately 35 km, and then increases. Overfitting is thus suggested for models resulting from no smoothing, or smoothing with a parameter $s < 35$ km. The horizontal solid line represents the MSE for a stationary isotropic model; that is, for a model with infinite smoothing (no deformation).

The dashed lines in Figure 7 represent the mean square errors of fit for the dispersions among all pairs of the 45 sites used to fit the model. As expected, the mean square error of fit increases as the smoothing parameter s increases. The asymptotic mean square error of fit for infinite smoothing, denoted by the horizontal dashed line, happens to be at a higher level than the asymptotic mean square error of prediction. But this is due, in part, to differences in the ranges of dispersions among the 45 fitting sites and among the 30 validation sites.

It happens to be the case that the value of s we chose to produce a smooth, non-folding mapping, is near the minimum for the mean square error of prediction. Although this need not be the case, it is comforting as it is the result one would expect if a smooth two-dimensional deformation model is adequate to represent the spatial dispersion structure. If such a validation exercise suggested a model with substantial folding, then a higher-dimensional deformation model may be called for.

References

[1] Bookstein, F.L. (1989): Principal warps: Thin-plate splines and the decomposition of deformations. *I.E.E.E. Trans. Patt. Anal. Mach. Intell.* 11, 567-585.

[2] Bookstein, F.L. (1991): *Morphometric Tools for Landmark Data*, Cambridge University Press.

[3] Guttorp, P., and Sampson, P.D. (1992): Methods for estimating heterogeneous spatial covariance functions with environmental applications. In *Handbook of Statistics, Vol. 12: Environmental Statistics* (G.P. Patil and C.R. Rao, editors), Elsevier, (to appear).

[4] Hastings, W.K. (1970): Monte Carlo sampling methods using Markov Chains and their applications. *Biometrika* 57, 97-109.

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

[6] Mardia, K.V., Kent, J.T., and Walder, A.N. (1991): Statistical shape model in image analysis. In *Proceedings of the 23rd Symposium on the Interface between Computing Science and Statistics*, pp. 550-557. Interface Foundation, Fairfax Station.

[7] Mardia, K.V., and Goodall, C.R. (1993): Spatial-temporal analysis of multivariate environmental monitoring data. To appear in *Multivariate Environmental Statistics.* N.K. Bose, G.P. Patil, and C.R. Rao, eds., North Holland, New York.

[8] Matérn, B. (1986) *Spatial Variation.* (2nd Ed.) Lect. Notes in Statistics (Vol. 36). Springer Verlag.

[9] Monestiez, P., and Switzer, P. (1991): Semiparametric estimation of nonstationary spatial covariance models by metric multidimensional scaling. Technical Report, Dept of Statistics, Stanford University.

[10] Monestiez, P., Sampson, P.D., and Switzer, P. (1992): Multidimensional scaling and nonstationary spatial covariance modeling. In *Distancia '92: Congres International sur Analyse en Distance*, S. Joly and G. Le Calve, eds., Rennes, France, 425-428.

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

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

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

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