SPATIAL COVARIANCE ESTIMATION BY
SCALED-MEETING SCALING AND BIORTHOGONAL GRIDS

Paul Sampson

Tech. Report No. 102

SIAM INSTITUTE FOR MATHEMATICS AND SOCIETY

HEALTH SCIENCES PROGRAM
RAND CORPORATION

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY

DEPARTMENT OF STATISTICS
UNIVERSITY OF BRITISH COLUMBIA
SPATIAL COVARIANCE ESTIMATION BY
SCALED-METRIC SCALING AND BIORTHOGONAL GRIDS

by

Paul Sampson
University of Washington

Tech. Report # 102
October 86

Prepared in part under support to SIMS from the United States Environmental Protection Agency
Department of Statistics, University of British Columbia, 2021 West Mall, Vancouver, B.C. V6T 1W5
Spatial Covariance Estimation by
Scaled-Metric Scaling and Biorthogonal Grids

Paul D. Sampson

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

August, 1986

ABSTRACT

A new, nonparametric approach is proposed for global estimation of the spatial covariance structure of a random function $X(z)$ observed repeatedly at a finite number of sampling stations $z_i$, $i=1, 2, \ldots, n$, in the plane. The true covariance structure is assumed to be neither isotropic nor stationary. Using a new variant of multidimensional scaling, termed scaled-metric scaling, one may compute a two-dimensional representation for the sampling stations in which interpoint distances approximate sample estimates of the spatial dispersions $\text{Var}(X(z_i)-X(z_j))$. (These variances are usually fitted by parametric models for the variogram.) The methods of minimum-Laplacian interpolation and biorthogonal grids, introduced by Bookstein in the field of morphometrics, are applied to compute and depict a smooth mapping of the geographic representation of the sampling stations onto the scaled-metric scaling representation. This yields a graphical representation of the global spatial covariance structure and a method for estimating $\text{Var}(X(a)-X(b))$ for any two unsampled locations $a$ and $b$ in the geographic plane. Preliminary applications to solar radiation data and rainfall acidity data are discussed.

KEYWORDS: Dispersion; Interpolation; Kriging; Multidimensional scaling; Nonstationarity; Variogram.

Prepared in part under support to SIMS and the Department of Statistics, University of British Columbia, from the United States Environmental Protection Agency.

Department of Statistics, University of British Columbia, 2021 West Mall, Vancouver, B.C. V6T 1W5
1. INTRODUCTION

The development of nonparametric procedures for interpolating observed spatial covariances of a random function sampled at a finite number of locations has lagged well behind the development of interpolation methods for the expected value of the underlying function. Among the latter are finite element methods, stiff lamina methods, natural neighbour interpolation, kriging, splining, and local regularization (see the discussions in Ripley 1981, Sibson 1981, Watson 1984, and Weerahandi and Zidek 1985). Kriging is distinguished by its ability to directly provide estimates of the variance of interpolation error at arbitrary locations. The kriging and regularization methods depend explicitly on the spatial covariance or variogram functions.

Most approaches to modeling spatial covariance structure have been parametric, and have assumed isotropy and/or stationarity. Fundamental results are given by Matérn (1960). The best known models are those for the variogram originating in Matheron's theory of "regionalized variables;" cf. Matheron (1963, 1971), Delhomme (1978), Journel and Huijbregts (1978), and Stein (1984). The common assumption of a spatially stationary variogram in kriging analyses was called the "intrinsic dispersion law" by Matheron (1963). Recent work on variogram estimation includes a number of papers in a NATO ASI volume (Verly, David, Journel, and Marechal 1984), and a thesis of Stein (1984) covering parametric and nonparametric estimation. Although almost all previous work assumes stationarity, there is little reason to expect spatial covariance structures to be stationary over many regional/spatial scales of interest. Cressie (1985a) considers scaling of variograms to accommodate a form of heterogeneity (see also Cressie 1986). 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 affects weather patterns, and thereby the transportation of pollutants. Empirical spatial covariances that reflect paths of pollutant transportation (for data averaged over appropriate temporal scales) should therefore be nonstationary insofar as the landscape varies from place to place. Analyses of
Escoufiier, Camps, and Gonzalez (1984) and Chami and Gonzalez (1984), in particular, suggest that stationarity should not be assumed in studies of atmospheric pollution data.

The lack of attention given to estimation of nonstationary spatial covariances is certainly justified in kriging's native field of geostatistics. Applications in that field primarily concern observation of a single realization of a random function, and so it is hard to proceed with any analysis without assuming some type of stationarity. But, whenever nonparametric estimates of covariance between pairs of sampling stations can be estimated from replicate observations in time, one need not make such assumptions.

A number of different approaches to time replicated atmospheric deposition data have recently appeared. Finkelstein (1984) computes a year's average value for each sampling station and then applies standard geostatistical methods (assuming isotropy and stationarity) to estimate the sample semi-variogram. Bilonick (1984) similarly applies standard geostatistical methods to estimate the sample semi-variogram of pH deposition for observations, month by month, on a network of sampling stations. He then fits a parametric model to the average of these monthly sample semi-variograms. Other recent kriging-type analyses of atmospheric deposition pay more attention to temporal structure but still utilize stationary variogram models; see, for example, Eynon and Switzer (1983), and Egbert and Lettenmaier (1986). Only Chami and Gonzalez (1984) appear to have considered nonstationary covariance structure.

This paper presents a new approach to the estimation and graphical depiction of spatial covariance. The method is nonparametric and assumes neither isotropy nor stationarity. Supposing that the random function is sampled at a fixed number of sampling stations, the fundamental idea is to relate the geographic coordinates of the sampling stations to a different set of coordinates for the stations. The new coordinates represent spatial covariance as expressed through the variogram, \( \text{Var}(X_i - X_j) \), where \( X_i \) and \( X_j \) represent observations at stations \( i \) and \( j \). (We will refer to \( \text{Var}(X_i - X_j) \) as spatial dispersion.) Implementation of this idea requires two tools. The first is a new variant of metric multidimensional scaling (MDS), which I call scaled-metric scaling (SMS). SMS generates a coordinate representation of the sampling stations where
interpoint distances reflect (smoothed versions of) the sample variogram. The second tool is the method of minimum-Laplacian interpolation, borrowed from the field of morphometrics (Bookstein 1978a,b; Bookstein, Sampson, and Hurley, in prep.), and its visualization by biorthogonal grids, borrowed likewise. This interpolation provides a mapping between the representations of the stations in the two coordinate systems.

Together, these tools provide estimates of the covariance between observations at any two locations in the geographic sampling region. Estimates of covariance between observations at two observed stations are smoothed versions of the original sample estimates. Estimates at unsampled locations are consistent with the estimates at the sampling stations. Biorthogonal grids provide a graphical depiction of the spatial covariance (second-order structure) of the random function.

This methodology was specifically motivated by problems of environmental monitoring: to map the spatial distribution of the deposition of some chemical pollutant, and to help design an "optimum" network of sampling stations for observing that distribution. Most approaches to optimal network design have aimed to minimize quadratic estimation error at single points or over regions; cf. Formery and Matheron (1963), Delhomme and Delfiner (1973), Cabannes (1979), Hughes and Lettenmaier (1981), Escoufier et al. (1984), and Olea (1984). Caselton and Zidek (1984) consider instead an information theoretic approach based on Shannon's (1948) measure of information. They focus on a hypothetical high-density network of m possible sampling stations. The design problem is to determine a subset of sampling stations of size n<<m so that measurements at these stations will minimize the uncertainty about values at the m-n unobserved sites; or equivalently, the subsample of stations is to be chosen so as to maximize the amount of information transmitted between observed measurements and estimates at the unobserved sites. Caselton and Zidek provide specific results for the case of a multivariate normal model for the observations at the m potential sampling sites when the spatial covariances among all of these sites are known. Clearly some method of estimating covariances is required for their computations to proceed, as estimates of covariance are typically available only among stations in a relatively small pilot monitoring network. For the more commonly addressed problem of adding or deleting
stations from an essentially complete network (on the basis of quadratic estimation error), the cubic spline interpolation of spatial covariances utilized by Chami and Gonzalez (1984) provides a reasonable, although less elegant, alternative to the method proposed here.

The next section summarizes the method. More detailed descriptions of the two component tools, the scaled-metric scaling algorithm and the method of minimum-Laplacian interpolation (and biorthogonal grids), follow in Sections 3 and 4. Section 5 includes two preliminary applications of the method: to solar radiation data exhibiting strong spatial covariance structure (Hay 1984), and to pH-deposition data with very weak spatial covariance structure (Eynon and Switzer 1983). We conclude with a list of theoretical and applied questions to be addressed in further research.

2. OUTLINE OF THE METHOD

Suppose that data are available from each of \( n \) sampling stations at \( t \) points in time: write \( x_{ik} \) for the observation taken at location \( i \) at time \( k \), \( i=1, 2, \ldots, n \), \( k=1, 2, \ldots, t \).

Arriving at this setup is not necessarily trivial. For instance, there is often a question about the relevant timescale for averaging. Some wet-deposition networks, such as MAP3S (1982), provide data on a rain-event basis, while others, such as NADP (1980), provide only weekly totals. The spatial covariation of a process depends on the temporal scale on which it is defined and on modeling of temporal trends which might be "removed" from raw data. Data averaged over periods of time commensurate with the time required for weather systems to travel and influence multiple sampling stations result in larger observed covariances among stations. There also may be a direct dependence of spatial variability on time or season, as Suckling and Hay (1978) observed for solar radiation variability. These factors and others may need to be considered in determining an "appropriate" or relevant temporal scale for analysis.

Supposing that these and other issues of temporal scale have been resolved, we have an \( n \times t \) data matrix \( X \). We view this as \( n \) points in \( \mathbb{R}^t \) since our analysis concerns relationships among,
and perhaps changes in, the set of \( n \) stations. The covariance between values at stations \( i \) and \( j \) is estimated as

\[
s_{ij} = \frac{1}{t} \sum_{k=1}^{t} (x_{ik} - \bar{x}_i)(x_{jk} - \bar{x}_j).
\]

Observed covariances are often associated with temporal processes, such as short-term meteorological processes distinct from the long-term seasonal trends typically removed. For this reason the residual series \( x_{ik} \) at individual stations may display serial (temporal) autocorrelation. There is no reason, in general, to model this residual autocorrelation prior to computing covariances, although a long series (such as the solar radiation data studied in Section 5) might be subsampled in order to obtain nearly uncorrelated observations. This has not been an issue in analyses of pH deposition by Eynon and Switzer (1983; rain event-based data from the EPRI network) and Egbert and Lettenmaier (1986; weekly data from the NADP network) which found no significant temporal autocorrelation after accounting for seasonal variation.

To estimate (or interpolate) values of a random function at unobserved sites from values recorded at certain sampling stations, it is essential to know the variances of differences between sampling stations (Matheron 1971). Let \( X(z) \) represent a random function of geographical coordinates \( z \), and let \( h \) represent a displacement; then most analyses are carried out in terms of the semi-variogram \( \gamma(h) \) defined as

\[
\gamma(h) = \frac{1}{2} \text{Var} (X(z+h)-X(z)).
\]

The dependence of \( \gamma(h) \) only on \( h \) and not \( z \) is a statement of the intrinsic law, or (weak) stationarity. This is a less stringent condition than second order stationarity since the variance of the random function need not exist, only the variances of differences or increments.

We focus not on the variogram per se, but on the variances of differences between observations at stations \( i \) and \( j \) estimated without assuming stationarity:

\[
d_{ij}^2 = \hat{\text{Var}} (X_i - X_j) = s_{ii} + s_{jj} - 2 s_{ij}.
\]
This is occasionally referred to as the sample variogram (Matheron 1971), but more often the term variogram is reserved for the context of the intrinsic hypothesis. We will refer to these variances simply as *spatial dispersions* (the word "dispersion" being chosen because both its statistical and physical connotations are relevant).

### 2.1 The G-plane and the D-plane

The stations of the data set have geographical coordinates of longitude and latitude (or their equivalent after a map projection). Call the plane of map coordinates the *G-plane* ("G" for geographic). A second planar representation of interest is determined from the spatial dispersions. Spatial dispersions provide the "standard" transformation of the covariance matrix \( S = [s_{ij}] \), considered as a similarity matrix, to a distance matrix \( D \) (Mardia, Kent and Bibby 1979, Greenacre and Underhill 1982). In this plane, distances \( \delta_{ij} \) between stations are approximately given by the root-dispersions \( d_{ij} \) for small values of \( d_{ij} \). That is, squared inter-station distances *locally* reflect spatial dispersion. On a global scale we assume these distances to be monotone nonlinear functions of \( d_{ij} \). Call this representation the *D-plane* ("D" for dispersion). Because the spatial dispersions are functions only of \( \delta_{ij} \) in the D-plane, we may view the D-plane as a transformation of the G-plane determined so that the spatial dispersion structure is both stationary and isotropic.

There are two reasons why the \( d_{ij} \)-metric is required to apply only locally in the D-plane. First, interpolation is essentially a problem of local variations in a random function, and our description of the spatial covariance field will likewise be local, specifying in what direction from a given point the spatial covariances are strongest (or the variogram or \( d_{ij} \) smallest), and in what direction weakest. Second, whenever values at stations are uncorrelated beyond a finite range, this planar representation could not possibly reproduce the unscaled \( d_{ij} \) exactly.

Consider, for example, the G-plane distribution of sampling stations depicted in Figure 1. Suppose positive spatial correlations among stations 1-4 and among stations 5-7, but that any station from the first group is uncorrelated with any station from the second group. If the variance of the random function is constant, then \( (d_{ij})^2 \) must be constant (= 2 x variance of an observation)
for any choice of \( i \) in \( \{1,2,3,4\} \) and \( j \) in \( \{5,6,7\} \). However, almost any planar representation of these stations will be such that the distance from station \( i \) to station \( j \) will vary over \( j \) in \( \{5,6,7\} \), and so we cannot accurately represent the \( d_{ij} \) with stations located in a plane.

-- Figure 1 --

2.2 Computing the D-plane representation

The first step in the covariance interpolation method is the representation of the stations in a D-plane locally reflecting the \( d_{ij} \). This is computed using a variant of metric multidimensional scaling (see, for example, Mardia, Kent, and Bibby 1979, Chap. 14; Gower and Digby 1981; or Greenacre and Underhill 1982). An alternative approach would be to map the stations not into a plane, but an arbitrary two-dimensional Riemannian manifold (Pieszko 1975). The approach proposed here is to compute a smooth, monotone non-decreasing function \( \delta_{ij} = \delta(d_{ij}) \), approximately linear in a neighborhood of the origin, such that the \( \delta_{ij} \) can be accurately represented by the interpoint distances in a two-dimensional metric rescaling. The inverse of this nonparametric function, applied to the D-plane, is analogous to the usual variogram applied to the G-plane under an assumption of stationarity.

Although described with the metric \( d_{ij} \) in mind, this proposed scaling procedure is algorithmically a variant of well-known nonmetric scaling. It is implemented simply by replacing a step common to most nonmetric MDS algorithms, Kruskal's (1977) least squares monotone regression of the D-plane distances \( \delta_{ij} \) on the dispersions \( d_{ij} \), by an ACE-type monotone regression (Breiman and Friedman 1985). In contrast with the usual nonmetric scaling solution, the approximate relationship between the dispersions \( d_{ij} \) and the rescaled distances \( \delta_{ij} \) is assumed to be smooth rather than just monotone. Furthermore, with the ACE algorithm we deal partially with our uncertainty about heterogeneity of variance in the \( d_{ij} \) and a corresponding need to transform or weight the regression relating \( d_{ij} \) and \( \delta_{ij} \) via least squares.
Because this algorithm preserves the metric interpretation of interpoint distances in the D-plane, I call it \textit{scaled-metric scaling} (SMS). It may also be viewed as a nonparametric version of least squares scaling as described by Greenacre and Underhill (1982). In fact, weighted, parametric least squares scaling, using parametric variogram models to transform the $\delta_{ij}$ and weights as in Cressie (1985b), could provide a reasonable alternative approach. However, it would be difficult to justify any of the commonly used variogram transformations for the D-plane representation of the stations.

\subsection*{2.3 Interpolating between the G and D-planes}

Once the stations are located in the D-plane, we need to compute a smooth function $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ that maps (interpolates) the coordinate representation of the sampling stations in the G-plane exactly into their scaled "spatial dispersion coordinates" in the D-plane. The interpolant is a new version (Bookstein et al. in prep.) of the \textit{minimum-Laplacian interpolator} introduced by Bookstein (1978a,b) in the field of morphometrics. The mapping $f$, computed as a composite of two thin-plate splines (Meinguet 1979, 1984), maps \textit{any} locations in the G-plane into points in the D-plane. If $z_a$ and $z_b$ are two G-points, then the distance between their images in the D-plane, $\delta_{ab} = |f(z_a)-f(z_b)|$, yields an estimate of the covariance or spatial dispersion between observations at the two locations

$$\hat{d}_{ab} = \left( \delta^{-1}(\delta_{ab}) \right)$$

where $\delta^{-1}$ denotes the function inverse to $\delta$ so that $\delta^{-1}(\delta_{ij}) = d_{ij}$.

Along with the mapping function $f$ we also generate a graphical depiction of the nature of the spatial covariance field. This depiction consists of two families of curves drawn on the G-plane. Curves within families do not intersect, while every intersection of curves between families is orthogonal. Through every point of the G-plane passes one curve from each of the two families, and their orthogonal directions indicate the directions of strongest and weakest spatial dispersion. Examples of these families of curves, called \textit{biorthogonal grids} (Bookstein 1978a,b), are provided in Sections 4 and 5.
It should be emphasized that the approximation of observed (estimated) spatial dispersions by distances in the D-plane is a smoothing of the original values. Smoothing noisy estimates of spatial covariance is desirable; fitting parametric functions of geographic distance to nonparametrically estimated variograms is probably the most common approach. (See also the nearest-neighbor kernel smoothing of Egbert and Lettenmaier 1986.) Stein (1984) notes that when a plot of a nonparametric estimator of the semi-variogram $\gamma(|h|)$ is smoothed, the resulting function is not, in general, a permissible semi-variogram as it is not conditionally positive definite. However, the SMS dispersion estimates are smoothed by the constraints of Euclidean geometry. Because the estimates are computed from a metric representation of the stations, the covariance function they represent will be positive-definite in "well-behaved" cases.

With the mapping $f$ one may estimate the covariances for any hypothetical network of sampling stations. These estimates, free of stationarity assumptions, are exactly those required by the network design methodology of Caselton and Zidek (1984). Furthermore, the biorthogonal grids provide an insightful graphical image of the spatial covariance field with their indication of the geographical directions along which station-to-station covariances are strongest and weakest. I know of no other single graphical display for global description of a spatial covariance field. In the case of a stationary (or homogeneous), but anisotropic spatial covariance structure, a single plot of iso-correlation lines suffices (see, for example, Egbert and Lettenmaier 1986), but for the non-stationary case a large number of plots are necessary to approximately represent the field (see Chami and Gonzalez 1984).

3. SCALED-METRIC SCALING OF SPATIAL DISPERSIONS

The SMS algorithm for mapping the sampling stations into a D-plane is best described in terms of the usual Kruskal-Shepard nonmetric scaling algorithm (Kruskal 1977). Underlying the Kruskal-Shepard algorithm is an hypothesized "model" relating the observed "distances" (or
"dissimilarities") $d_{ij}$ to hypothesized distances $\delta_{ij}$ defined among points in a two-dimensional Euclidean space,

$$d_{ij} = h(\delta_{ij} + \epsilon_{ij}) ,$$

where the $\epsilon_{ij}$ are errors and $h$ is an unknown monotone-increasing function. The basic elements of the algorithm in this application are as follows.

(a) Order the off-diagonal elements of the dissimilarity matrix $D=(d_{ij})$ so that

$$d_{i,j_1} \leq d_{i,j_2} \leq \cdots \leq d_{i,j_m} , \ m = n(n-1)/2 .$$

(b) Let $\hat{Y}$ be an $(n \times 2)$ matrix of coordinates defining a configuration of sampling stations in $\mathbb{R}^2$, and denote the interpoint distances by $\hat{\delta}_{ij}$. Compute a least squares monotone regression of the $\hat{\delta}_{ij}$ on the $d_{ij}$ and denote the fitted values by $\hat{d}_{ij}$. Define the squared stress of $\hat{Y}$ by

$$S^2(\hat{Y}) = \frac{\sum_{i<j} (\hat{d}_{ij} - \hat{\delta}_{ij})^2}{\sum_{i<j} \hat{\delta}_{ij}^2}.$$ 

$S^2(\hat{Y})$ provides a measure of the degree to which the rank order of the configuration distances $\hat{\delta}_{ij}$ agree with the rank order of the dissimilarities $d_{ij}$ to be mapped. It is invariant with respect to translation, rotation, or rescaling of the configuration.

(c) Determine a configuration $\hat{Y}$ that minimizes the squared stress $S^2(\hat{Y})$ using a steepest descent algorithm (Kruskal 1977).

For the application to spatial dispersion, the fitted configuration distances $\hat{\delta}_{ij}$ should relate smoothly to the $d_{ij}$ via a monotone increasing function that is approximately linear in a neighborhood of the origin. To guarantee this we should replace the simple monotone regression in part (b). As long as there is no "nugget effect" in our spatial covariance structure (Matheron 1971), smooth functions will provide the requisite approximate linearity in a neighborhood of the origin.
One approach would be to use a smooth, monotone nonparametric regression (e.g. Friedman 1984). However, in this application there is considerable noise in the regressor variable $d_{ij}$. Assuming Gaussian data one might estimate $\text{var}(d_{ij})$ and carry out a weighted analysis analogous to Cressie's (1985b) weighted least squares estimation of variograms, however, this would be a challenging task and the results may be sensitive to the assumptions and estimated weighting.

The "ACE algorithm" of Breiman and Friedman (1985) provides, perhaps, the most general approach to this problem. ACE can determine monotone functions $\theta(\cdot)$ and $\phi(\cdot)$ that maximize the correlation between $\theta(\hat{\delta}_{ij})$ and $\phi(d_{ij})$. The iteration indicated in step (c) above then applies to minimize the squared stress defined in terms of the ACE transforms:

$$S_{\text{ACE}}^2 = \frac{\sum_{i<j} \left( \theta(\hat{\delta}_{ij}) - \phi(d_{ij}) \right)^2}{\sum_{i<j} \left( \theta(\hat{\delta}_{ij}) \right)^2}.$$

The ACE algorithm proceeds by a series of alternating conditional expectations (or smooth, monotone nonparametric regressions): given a trial $\theta(\cdot)$ compute $\phi(d_{ij}) = E[\theta(\hat{\delta}_{ij}) | d_{ij}]$; then use this new estimate of the transformation $\phi(\cdot)$ to compute $\theta(\hat{\delta}_{ij}) = E[\phi(d_{ij}) | \hat{\delta}_{ij}]$; and iterate until the estimates of the functions $\theta(\cdot)$ and $\phi(\cdot)$ cease to change. The existence of optimal transformations is proved in Breiman and Friedman (1985).

This procedure would appear to be computationally intensive, as the iterative ACE algorithm is nested within iterations of the scaling search. However, the ACE solutions should converge relatively quickly for this simple bivariate problem.

An important issue in multi-dimensional scaling generally, and this scaled-metric scaling in particular, is the choice of an initial configuration. Different initial configurations will often result in convergence to different final configurations. Fortunately, we have good prior information about what preferred solutions should look like, and this information suggests at least three different starting points. Because the mapping $f$ to deform the G-plane into the D-plane is required to be smooth (see Section 4), one logical choice for the initial configuration is the
geographic configuration itself. With this starting point it would be especially interesting to view kinematically the search for the optimal configuration (Buja 1982). We would then see the spatial covariance pattern "deforming" the geographic configuration.

An expression of the smoothness of the mapping is that it is locally affine (or locally linear); that is, over small geographic regions it behaves as a mapping \(z \mapsto Az\) where \(A\) is the \(2 \times 2\) matrix of derivatives of \(f\). In the case of just three sampling stations, the computed mapping would be exactly the affine mapping that transforms the triangle of stations in the G-plane into a triangle of stations representing the \(d_{ij}\) in the D-plane as illustrated in Figure 2. This suggests choosing a triangulation of the geographic stations (with triangle sides chosen as small as possible) and then piecing together triangles in the D-plane which exactly represent the corresponding \(d_{ij}\) in groups of three. A piecewise affine transformation would then describe the relationship between the G-plane and this initial D-plane configuration.

--- Figure 2 ---

Another interesting initial configuration corresponds to a stationary covariance structure. In the stationary case, for any pair of stations \(i\) and \(j\) with geographic coordinates \(z_i\) and \(z_j\) such that \(|z_i - z_j|\) is "small", we would have

\[
d_{ij}^2 = (z_i - z_j)^T A^T A (z_i - z_j)
\]

for some single matrix \(A\). Some form of weighted least squares estimation of \(A\) (with weights a decreasing function of \(|z_i - z_j|\)) would supply an estimate of the local affine structure \(A\) that would then be applied globally to transform all stations.

This particular initialization suggests a version of the SMS approach for semiparametric estimation of a covariance structure taken to be stationary but anisotropic. Generalizing the estimation of the affine structure from local to global, one would model the \((d_{ij})^2\) as a monotone function of distances among affine-transformed geographic coordinates:
\[
d_i^2 = h \left( (z_i - z_j)^T A^T A (z_i - z_j) \right) = h \left( \hat{\delta}_{ij}^2 \right)
\]
for all station pairs \(i, j\). Estimation would proceed, again, by an alternating least squares scheme: given an estimate of \(A\), \(h\) is estimated by nonparametric regression of the \((d_{ij})^2\) on the values of the quadratic form \((\hat{\delta}_{ij})^2\); then the elements of \(A\) are estimated by least squares using this \(h\), resulting in another estimate of \(A\), and so on.

Lacking an implementation of this ACE-variant of multidimensional scaling, I have obtained apparently reasonable results utilizing nonmetric scaling to explore the form of appropriate monotone transformations. I then carry out classical (Householder-Young) metric scaling on simple transformations of the dispersion estimates \((d_{ij})^2\). These approximate results are demonstrated in Section 5.

4. MAPPING THE G-PLANE INTO THE D-PLANE

The presentation in this section is drawn primarily from Bookstein (1978a,b) and Bookstein et al. (in prep.). Our aim is to demonstrate the idea of plane mappings as it applies to the interpolation of spatial dispersions. Mathematical details are left to the references just cited.

4.1 Mapping Three Sampling Stations

Consider first just three sampling stations with estimated covariances \(s_{ik}, i,k=1,2,3\). Provided that the root spatial dispersions \(d_{12}, d_{13},\) and \(d_{23}\) satisfy the triangle inequality, we can represent the stations exactly by three points in \(\mathbb{R}^2\) with interpoint distances given by the \(d_{ik}\) (up to an arbitrary translation and rotation). This is the D-plane representation. Because these three stations have an \(\mathbb{R}^2\) representation given by their geographical coordinates (the G-plane), the \(d_{ik}\) should satisfy the triangle inequality if there is a reasonably smooth correspondence between the \(d_{ik}\) and geographical distance. However, in the case of three stations which are nearly collinear in
the geographic plane, it is easy to see that modest noise could result in estimates $d_{ik}$ that do not satisfy the triangle inequality.

We compare the G and D-plane representations of the sampling stations by imagining a "deformation" that maps the G configuration into the D configuration. Let inter-station distances in the G-plane be denoted by $m_{ik}$. An example is given in Figure 2 where we note that $m_{12} = m_{13} = m_{23}$. Lacking any other information we must assume that the deformation is homogeneous between stations. The deformation that takes one triangle into another is then simply an affine (or linear) transformation. It is best pictured by observing how the inscribed circle of the triangle in the G-plane is transformed into an ellipse in the D-plane, the "strain ellipse" of the deformation. The transformation is described by corresponding pairs of orthogonal principal axes in the two planes. With respect to these axes the deformation is described by a pair of extensions (stretch or shrink), the lengths of the axes in the D-plane.

The orientations of the two triangles are irrelevant. The information of interest in this picture is contained primarily in the position of the principal axes with respect to the stations and the ratio of the extensions along these axes. We might scale the D-plane figure so that the arithmetic or geometric mean of the principal extensions is unity. Then the first principal axis, with extension $> 1$, indicates the geographical direction along which inter-station dispersions are weakest, and hence, inter-station predictability is greatest. The ratio of the extensions measures the degree of anisotropy in the spatial covariance structure.

It is, perhaps, easiest to interpret this picture in the case where within-station variances are constant, $s_{ii} = s$. Then $d_{ik} = (2(s-s_{ik}))^{1/2}$. Figure 2 indicates that the covariance between observations at stations 1 and 2 is nearly the same as it is for stations 1 and 3; however, stations 2 and 3 are separated further in the D-plane ($d_{23} > d_{12} = d_{13}$) and hence the spatial covariance is weaker in this direction, which is roughly aligned with the first principal axis of the ellipse.
4.2 The General Case and its Depiction by Biorthogonal Grids

Treating the D-plane representation of the sampling stations as an affine transformation of the G-plane locations, while simple and appealing, is not reasonable in the case of more than three sampling stations over a large area. However, the presumption of smoothness for the mapping between the planes implies deformation models that are affine locally.

For transformations other than the affine we cannot find two straight lines at 90° which transform into two other straight lines at 90° in the new image. We might instead look for families of curved lines which satisfy this property. In fact, through almost every point of a differentiable transformation pass just two differentials (i.e. local principal axes) which are at 90° both before and after transformation. The integral curves of these differentials, sampled finitely, form a 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. (See Bookstein 1978a, for an existence proof.)

An example of the depiction of such biorthogonal grids is given in Figure 3. In this example one of the integral curves of the differentials, corresponding all along its length to the local principal axis of greater extension, follows a direction that nearly connects stations 1 and 3. The estimated spatial covariance is weakest along these curves, while curves of the orthogonal family lie along directions of strongest spatial covariance point by point.

-- Figure 3 --

4.3 Interpolation of a Differentiable \( \mathbb{R}^2 \to \mathbb{R}^2 \) Mapping

Let Cartesian coordinates in the G-plane be denoted by \((x,y)\), those in the D-plane by \((u,v)\). Given \(n\) pairs of corresponding locations, \((x_i, y_i)\) and \((u_i, v_i)\), \(i=1, 2, \ldots, n\), we must determine a bivariate function \(f\) such that
\[ f \left( \begin{array}{c} x_i \\ y_i \end{array} \right) = \begin{array}{c} u_i \\ v_i \end{array} \], \quad i = 1, 2, \ldots, n. \]

At this point the problem is not well-posed. Interpolation algorithms generally select a function from a specified vector space of functions (e.g. polynomials), or choose a function to minimize a certain "roughness" criterion. For example, in \( \mathbb{R}^1 \) the familiar cubic spline minimizes \( \int (f'(x))^2 \, dx \) subject to constraints at the knots and at the boundary. The measure of roughness for the bivariate problem can be pictured as follows. Consider a square and its image under a mapping \( f \) as shown in Figure 4. "Roughness" is defined to be the square of the distance between the image of the centroid and the centroid of the image.

-- Figure 4 --

Expressing the mapping \( (x,y) \Rightarrow (u,v) \) in a second order Taylor series one finds that this distance is proportional to \( \left( (\nabla^2 u)^2 + (\nabla^2 v)^2 \right)^{1/2} \) where \( \nabla^2 \) is the Laplacian operator \( (\partial^2/\partial x^2) + (\partial^2/\partial y^2) \). It is therefore reasonable to require that the interpolant \( f \), the distortion function, have minimum net roughness \( \int |\nabla^2 f|^2 \) subject to the constraints of mapping \( (x_i,y_i) \) onto \( (u_i,v_i) \), \( i=1,2,\ldots,n \). This is an explicit generalization of what cubic splines accomplish: the minimization of \( \int (f'(x))^2 \, dx \) yields a cubic spline which satisfies \( f^{(4)} \equiv 0 \), except at knots. In the bivariate problem, \( \int (f'(x))^2 \, dx \) is generalized to \( \int |\nabla^2 f|^2 \). It can be shown that any function minimizing this over a region satisfies the biharmonic equation \( \nabla^2 \nabla^2 f \equiv 0 \), except at the points \( (x_i,y_i) \).

For locations \( (x_i,y_i) \) in the G-plane and corresponding \( (u_i,v_i) \) in the D-plane, the bivariate function \( f \) satisfying the biharmonic equation is actually computed as two thin-plate splines, one for \( u \) and one for \( v \). Further details are contained in Meinguet (1979, 1984) and Bookstein et al. (in prep.). These interpolators are simple linear parametric (but not polynomial) functions defined using basis functions centered at the \( n \) observations in the G-plane. They bear a total of
2n+3 parameters. From this parametric representation it is easy to compute the local 2 x 2 matrices of affine derivatives which are used for computing the principal axes, and then their integral curves, the biorthogonal grids.

4.4 Interpolating observed spatial dispersions

Let \( \delta_{ij} \) denote the Euclidean distances for the optimal configuration \( Y \) at convergence of the scaling search. To estimate spatial dispersions between points other than the sampling stations, the geographic coordinates \( Z \) (n \times 2) of the stations in the G-plane are mapped into the coordinates \( Y \) (n \times 2) in the D-plane using the method of minimum-Laplacian interpolation. Then, if stations \( a \) and \( b \) have coordinates \( z_a \) and \( z_b \), respectively, we will obtain estimated D-plane coordinates \( y_a = f(z_a) \) and \( y = f(z_b) \) with corresponding interpoint distance \( \delta_{ab} \). To estimate the spatial dispersion we would seem to require the function \( \phi^{-1}(\Theta(\cdot)) \). However, as suggested in Breiman and Friedman (1985), the optimal predictor would be determined as \( \text{E}[d_{ab} | \theta(\delta_{ab})] \), and this is computed in practice by a monotone smoothing of the \( d_{ij} \) on the \( \theta(\delta_{ij}) \).

5. EXAMPLES

Example 1. Solar Radiation Data. Figures 5-7 present a preliminary analysis of a solar radiation monitoring network in southwestern British Columbia (Hay 1984). Figure 5, taken from Hay (1983), displays the locations of the 12 monitoring stations. We selected the daily solar radiation total (MJ m\(^2\) day\(^{-1}\)) every fifth day for the years 1980-83 and used the ACE algorithm (Breiman and Friedman 1985) to model a transformation of the radiation total by the sum of a smooth (periodic) seasonal trend and individual station mean effects. Covariance matrices and spatial dispersions were computed from the residuals, separately for each of the four seasons. The (root) spatial dispersions \( d_{ij} \) for autumn (Sept-Oct-Nov) among these twelve stations are given in Table 1.
Inter-station correlations are very high (.80-.97) for these data, and a casual inspection of this matrix shows that these dispersions are closely related to the geographic distances among the stations.

Figure 6a shows the monitoring stations in the D-plane. The locations were computed by a classic metric multidimensional scaling algorithm (as implemented in the S system for data analysis and graphics; Becker and Chambers 1984), an algorithm not as flexible as the semi-metric scaling algorithm proposed here. The scaling was based on the dispersions \((d_{ij})^2\); i.e. we specified \(\delta(d_{ij})=(d_{ij})^2\). The solution resulted in 98% of the variation in the observed spatial dispersion estimates being explained by the distances among the stations in this D-plane representation. The most obvious deviation between the two planar representations is in the relative location of station 1, Grouse Mountain. In fact, this station has an elevation of 1128 meters, while all the others lie below 130 meters. This orographic feature explains the relatively high dispersions (low covariances) between this station and all others, as reflected in the scaling of Figure 6a.

The scaling of station 9 nearly on top of station 8 is a consequence of the simple metric scaling algorithm which does not give sufficient weight to the low dispersions observed among the geographically neighboring stations 7, 8 and 9.

The minimum-Laplacian mapping computed between these two representations of the stations is summarized by biorthogonal grids in Figure 6b. The curves running generally north-south lie along the geographic directions in which spatial covariance is weakest (or spatial dispersion greatest). The orthogonal family of curves lie along the directions of strongest spatial covariance.
The curves themselves are coded to represent the relative magnitude of the spatial dispersion as indicated on the figure. The departure of these curves from a uniform rectangular grid is an expression of non-stationarity in the estimated spatial dispersions.

The clearest feature is the relatively weaker spatial covariance in a north-south direction along the coast, graphically indicated by the solid vertical curves and the denser spacing of the curves running generally east-west. Reanalysis without station 1 shows that this observed effect is largely explained by station 1.

The structure of the curves around station 10 indicates a singularity where the transformation is locally an isotropic change of scale, and, hence, the inferred covariance structure is locally isotropic. Such singularities are discussed by Bookstein (1978a).

**Example 2.** Rainfall acidity. Eynon and Switzer (1983) analyzed the variability in rainfall acidity as determined from a rainfall-monitoring network established by the Electric Power Research Institute (EPRI) of Palo Alto, California. After removing effects of seasonality, rainfall volume, and measurement error, Eynon and Switzer fitted a stationary, anisotropic semi-variogram model to the residual pH values. Their model is

\[
\frac{1}{2} \text{Var} (X(z+h) - X(z)) = \frac{g_1}{1 + (h^T A h)^{-1}}
\]

To confirm the validity of our method we applied it first to spatial dispersions computed from the fitted parametric model. In this case the "optimal" nonlinear transformation of the dispersions \((d_{ij})^2\) in the D-plane must be given by the equation for the distances \((h^T A h)^{1/2}\) in terms of the left-hand side spatial dispersions. (Usually this has to be determined from data.) Classic metric multidimensional scaling applied to these transformed dispersions produces the D-plane shown in Figure 7a; the biorthogonal grid for this map is the rectangular grid shown in Figure 7b. In this case the distances \(d_{ij}\) in the D-plane represent the fitted variogram values exactly. The rectangular grid is characteristic of spatial dispersion patterns that are stationary. The transformation from the G-plane to the D-plane is just the affine transformation specified by the matrix A. The grids are
oriented parallel to the principal axes (ENE-WSW and SSE-NNW) determined by $A$, and the (constant) extensions along the grids are given by the eigenvalues of $A$. The greatest spatial covariance is ENE-WSW.

-- Figure 7 --

The correlation between the observed spatial dispersions (of the pH residuals) and the fitted variogram values is only .47. Furthermore, the structure in the spatial dispersions is not related very closely to geographic distances. No multidimensional scaling solution at all resembled the geographic locations of the stations. Likewise, the spatial dispersions for the raw pH data (uncorrected for seasonal, temporal, and measurement effects) shows no well-defined correspondence with geographic location. The matrix of dispersion values is given in Table 2.

-- Table 2 --

The result of a metric scaling of the $(d_{ij}^2)^2$ for the raw pH data is shown in Figure 10. The G-plane is in fact folded (rather than simply "deformed") in the mapping onto the D-plane. Nonetheless, the distances among points in the D-plane account for 70% of the variance in these squared spatial dispersions. Biorthogonal grids can be computed for this mapping; however, in light of the folding of the G-plane this is surely not a useful analysis of the spatial covariance structure of pH. Rather, it illustrates what may happen when this type of scaling approach is applied to data exhibiting no well-defined spatial covariance structure.

-- Figure 8 --
6. DISCUSSION

The results demonstrated in Section 5 establish the promise of the combination of scaled-metric scaling with minimum-Laplacian interpolation and biorthogonal grids for interpolating observed spatial covariances. Further research is necessary before it can be exploited as a routine tool for spatial analysis.

Many questions commonly raised in connection with variograms must now be raised about this more general estimation method. We have explicitly assumed away any nugget effect in the data, but we must know how a nugget effect would manifest itself. Behavior of a variogram in a neighborhood of the origin determines the nature (continuity and differentiability) of the random function. The thinplate spline models yielding our estimates are everywhere continuous, but they are not differentiable at the observed station locations.

Another important characteristic of variogram models is their behavior at infinity. A "sill" parameter (the level of the variogram beyond the range of spatial correlation) appears as a multiplicative factor in prediction variances. The use of monotone transformations in the SMS procedure precludes an exact sill from being estimated. Furthermore, nonstationary spatial dispersions imply that the distance at which a sill is reached varies spatially.

It is not clear what family of (nonstationary) covariance functions is consistent with the SMS estimation method; that is, what covariance functions can be exactly represented by a D-plane representation. The conditions on the scaling solution under which these nonparametric estimates of spatial covariance are guaranteed to be positive-definite are not yet known.

Following this investigation, statistical properties of the spatial dispersion estimates must be assessed. Two important sources of variation are (a) sampling variation at observed monitoring stations, and (b) the sensitivity of estimates (and the biorthogonal grid picture) to variations in the list of monitoring stations. Finally, in view of the principal applications of this method to interpolation and monitoring network design, we need to know how helpful this method is under realistic conditions of nonstationarity.
Preliminary software has been implemented in the S system (Becker and Chambers 1984, 1985) for the computation of minimum-Laplacian interpolations and their display by biorthogonal grids as illustrated here. Further software development remains to be done, including implementation of the proposed semi-metric scaling. In addition, any technique as graphical as this needs to exploit the human visual system as effectively as possible. Experiments are required in the display of grids and anisotropy with different combinations of line and color. These improvements are relevant as well for "retrofitting" the biorthogonal grid graphic into morphometrics, which was their original context.

REFERENCES

Armstrong, M. (1984), "Improving the Estimation and Modelling of the Variogram," in, 
Geostatistics for Natural Resources Characterization (Part 1), NATO ASI Series, eds. G. 


USGS New York Acid Precipitation Data," in Geostatistics for Natural Resources 
Characterization (Part 2), NATO ASI Series, eds. G. Verly, M. David, A.G. Journel, and A. 

Bookstein, F.L. (1978a), The Measurement of Biological Shape and Shape Change, Lecture 

——— (1978b), "Linear Machinery for Morphological Distortion," Computers in Biomedical 
Research, 11, 435-458.

of Biorthogonal Grids."


Table 1. Spatial Dispersions $(d_{ij})^2$ for 12 Solar Radiation Monitoring Stations in B.C.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>2.19</td>
<td>0.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>2.40</td>
<td>0.77</td>
<td>0.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>2.72</td>
<td>1.30</td>
<td>1.10</td>
<td>0.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>3.16</td>
<td>2.09</td>
<td>2.03</td>
<td>1.22</td>
<td>0.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>2.40</td>
<td>1.29</td>
<td>1.19</td>
<td>1.34</td>
<td>1.81</td>
<td>0.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>2.76</td>
<td>2.08</td>
<td>1.98</td>
<td>2.12</td>
<td>2.21</td>
<td>1.33</td>
<td>0.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>2.91</td>
<td>2.04</td>
<td>1.92</td>
<td>2.02</td>
<td>2.16</td>
<td>1.32</td>
<td>0.71</td>
<td>0.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>2.94</td>
<td>1.98</td>
<td>1.93</td>
<td>2.07</td>
<td>2.20</td>
<td>1.28</td>
<td>0.91</td>
<td>0.63</td>
<td>0.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>2.63</td>
<td>1.56</td>
<td>1.51</td>
<td>1.47</td>
<td>1.67</td>
<td>0.86</td>
<td>1.08</td>
<td>1.09</td>
<td>1.06</td>
<td>0.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>2.56</td>
<td>0.80</td>
<td>0.62</td>
<td>0.73</td>
<td>1.62</td>
<td>1.17</td>
<td>2.02</td>
<td>1.92</td>
<td>1.91</td>
<td>1.39</td>
<td>0.</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>2.63</td>
<td>1.13</td>
<td>0.98</td>
<td>0.70</td>
<td>1.69</td>
<td>1.46</td>
<td>2.32</td>
<td>2.24</td>
<td>2.27</td>
<td>1.66</td>
<td>0.78</td>
<td>0.</td>
</tr>
</tbody>
</table>
Table 2. Spatial Dispersions $(d_{ij})^2$ for pH at 9 Monitoring Stations in the EPRI Network

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.109</td>
<td>0.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.224</td>
<td>0.106</td>
<td>0.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.152</td>
<td>0.085</td>
<td>0.152</td>
<td>0.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.160</td>
<td>0.087</td>
<td>0.153</td>
<td>0.099</td>
<td>0.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.148</td>
<td>0.154</td>
<td>0.259</td>
<td>0.145</td>
<td>0.109</td>
<td>0.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0.206</td>
<td>0.214</td>
<td>0.251</td>
<td>0.156</td>
<td>0.164</td>
<td>0.231</td>
<td>0.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>0.093</td>
<td>0.069</td>
<td>0.082</td>
<td>0.048</td>
<td>0.013</td>
<td>0.045</td>
<td>0.159</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>0.163</td>
<td>0.125</td>
<td>0.165</td>
<td>0.113</td>
<td>0.133</td>
<td>0.135</td>
<td>0.204</td>
<td>0.032</td>
<td>0.0</td>
</tr>
</tbody>
</table>
FIGURE LEGENDS

Figure 1. Two groups of sampling stations separated so that hypothetical spatial covariances between members of the first (1-4) and second (5-7) groups are zero. Euclidean distances cannot, in general, exactly corroborate this situation.

Figure 2. The affine mapping taking three equidistant sampling stations in the (geographic) G-plane (a) into locations in the (spatial dispersion) D-plane (b). Squared interpoint distances in the D-plane are given by \( \text{Var}(x_i - x_j) \).

Figure 3. Biorthogonal analysis of an artificial example, a bilinear map (from Bookstein 1978a). (a) A square defined by four sampling stations and a mesh representing the optimally smooth interpolation between the interiors. (b) The local principal axes, homologous directions at 90° in both images, computed on the mesh of (a). (c) A grid of integral curves for the tensor field of principal directions sampled in (b). (d) Two integral curves, and the extensions from square to quadrilateral of corresponding segments along them. The scale of the right-hand form was reduced for drawing.

Figure 4. (a) A small square and its image under the mapping from the G-plane to the D-plane (from Bookstein 1978a). (b) Enlargement of the right-hand side of (a). Roughness is the squared distance between the image of the centroid and the centroid of the image.

Figure 5. The 12-station solar radiation monitoring network, Vancouver, B.C. (from Hay 1983).

Figure 6. (a) Metric scaling (D-plane) representation of the 12-station solar radiation monitoring network derived from the spatial dispersion matrix in Table 1. (b) Biorthogonal grid depiction of the computed minimum-Laplacian interpolation between the G-plane and D-plane representations. The curves running generally north-south indicate local directions of greatest spatial dispersion (weakest spatial covariance).
Figure 7. (a) Metric scaling (D-plane) representation of the 9-station EPRI monitoring network derived from the fitted (stationary) variogram model of Eynon and Switzer (1983). Scaling was based on the dissimilarities \( (d_{ij})^2/(2g-(d_{ij})^2) \) where \( (d_{ij})^2 = 2g/(1+((z_i-z_j)^TA(z_i-z_j))^{-1}) \). (b) Biorthogonal grid depiction of the computed minimum-Laplacian interpolation between the G-plane and D-plane representations. Stationary variogram models are described by regular rectangular grids.

Figure 8. Metric scaling (D-plane) representation of the 9-station EPRI monitoring network derived from the observed (estimated) spatial dispersions in Table 2. Scaling was based on the squared dispersions \( (d_{ij}^2)^2 \). The result is an unsatisfactory folding of the G-plane representation.
Figure 1
Figure 2
Figure 6(b)
Figure 7(a)
Figure 7(b)