Chemometrics and Intelligent Laboratory Systems, 11 (1991) 209–228
Elsevier Science Publishers B.V., Amsterdam

Interpolation and estimation with spatially located data

Donald E. Myers
Department of Mathematics, University of Arizona, Tucson, AZ 85721 (USA)

(Received 10 July 1990; accepted 8 January 1991)

Abstract

Myers, D.E., 1991. Interpolation and estimation with spatially located data. *Chemometrics and Intelligent Laboratory Systems*, 11: 209–228.

Kriging is a regression method used with irregularly spaced data in 1-, 2- or 3-space for the estimation of values at unsampled locations or for the estimation of the spatial average over a length, area or volume. The estimator is linear in the data and the weights are obtained from a system of linear equations in which the coefficients are the values of variograms or covariance functions quantifying the correlation between data at two sample locations or between a sample location and the location to be estimated. The equations are obtained by minimizing the variance of the error of estimation, the variance being computed from a theoretical model for the correlation function rather than from empirical values as in most regression formulations. Estimation and modeling of this structure function is the most important and potentially the most difficult step in the process. While the method is not implemented in standard statistical packages, public domain software for use on an IBM personal computer or clone is available. The theory is briefly reviewed, practical aspects of the application of the method are discussed and available software and extensions are outlined. The US EPA Dallas Lead Study data is used to illustrate the problems and the method.

CONTENTS

0169-7439/91/$03.50

1. INTRODUCTION

Kriging is a regression technique used for the estimation/interpolation of spatially located and spatially correlated data. There are several features which set it apart from related or similar techniques such as trend surface analysis or from more classical parameter estimation problems and techniques. First of all, the location(s) of the samples is presumed to reflect valuable information and there is an assumed or apparent spatial correlation quantifiable in terms of separation distance and direction. Samples taken close together are expected to be more alike than samples far apart. In many cases samples are physically extracted from the earth, for example, in the case of soil or drill core samples. The volume or area of the sample is likewise an important piece of information and reflects the existence of short range nonhomogeneities. The associated volume or area is referred to as the support of the sample.

The technique grew out of problems encountered in mining and hydrology and gave birth to the discipline now known as geostatistics, which has found application in a variety of fields including environmental monitoring and assessment. Much of the early development is due to the group working under the direction of G. Matheron at the Ecole des Mines, France, although similar developments occurred in Sweden for applications in forestry and in the Soviet Union for applications in meteorology. In each case data is collected at a small number of locations in 1-, 2- or 3-dimensional space. The extent to which the number of sample locations is considered small is related to the geographic size of the region to be sampled and also to the scale of spatial dependence. In contrast to the case of more classical statistical techniques where sample size is related to distribution type and parameters such as the variance, sample size for geostatistical analyses is affected by a number of non-statistical characteristics including the dimension of the space. The collection of spatial data usually incorporates costs for physically retrieving the samples, reading the instrument recording the data or the cost of laboratory analysis. The latter may be quite significant. Usually the locations are not on a regular grid and the objective is to estimate values at unsampled locations, to estimate average values over volumes or areas or to estimate the proportion of a region where the concentration level is above a cutoff level. In the latter cases the sample locations may not all be inside the volume or area of interest.

Matheron and others formulated the problem in a random function context and the data are considered to be a non-random sample from one realization of the random function. This is rather different from the usual statistical formulation where the data are considered as multiple realizations of a random variable. In addition to obtaining a 'best' estimate, in a certain sense, the variance of the error of estimation is also obtained. This variance is not constant but rather reflects the spatial correlation and especially the sample location pattern. Kriging can be re-formulated in other equivalent ways, one of which leads to the thin-plate spline. This connection will be discussed below.

The application of kriging to a data set can be broken into several stages: exploratory statistical analysis of the data; estimation and modeling of the function which quantifies the spatial correlation; use of the spatial correlation function to determine the set of linear equations that determine the weights in the kriging estimator; and finally production of the estimated values and the associated (minimized) estimation standard devia-

tions. Often these estimates are obtained for points on a regular grid and the results are then used as the input for a contouring package. Both the variable of interest and the kriging standard deviation could be contoured. Nearly all of these steps would be exceedingly tedious without the aid of the computer, but fortunately software is now readily available. The method is not, however, a black box device. One cannot simply enter the data into a program and obtain the estimates (nor should one). It requires some comprehension and appreciation of the phenomena being analyzed as well as of the strengths and weaknesses of the method. Consequently the analysis is often the joint work of a statistician and a soil physicist, hydrologist, mining engineer or chemist.

2 THE KRIGING ESTIMATOR

We begin by considering the kriging estimator first and then later consider the problem of the estimation and modeling of the spatial correlation function. The kriging estimator has a practical and intuitive basis as well as a theoretical one. Let $x_1, \ldots, x_n$ denote points in 1-, 2- or 3-space. Since we will have few occasions to refer to the coordinates of a sample location, using x to denote a point instead of the first coordinate will not lead to confusion. Let $\mathbf{Z}$ represent the variable of interest and thus $\mathbf{Z}(x_1), \ldots, \mathbf{Z}(x_n)$ represent the data (values). Given an unsampled location x_0 or a volume $\mathbf{V}$, we wish to estimate $\mathbf{Z}(x_0)$ or the (spatial) average value over $\mathbf{V}$. If the sample locations were all inside $\mathbf{V}$ and corresponded to random selection, then estimation of the spatial average over $\mathbf{V}$ would correspond to the usual problem of estimating the mean (of a distribution). Therefore, the sample mean, i.e., the arithmetic mean, would be a natural choice for the estimator. That is, one would form a linear combination of the data with all weights (coefficients) being equal to $1/n$. This analogy would not apply if some locations were outside of $\mathbf{V}$, in which case it would seem reasonable to use unequal weights with the closest locations having the largest weights. In a practical sense a linear combination is the simplest function of the data. This simple form is important for several reasons. If the random function formulation is used, then $\mathbf{Z}(x_1), \ldots, \mathbf{Z}(x_n), \mathbf{Z}(x_0)$ could be considered as jointly distributed random variables. In that case the minimum variance unbiased predictor of $\mathbf{Z}(x_0)$, given the data $\mathbf{Z}(x_1), \ldots, \mathbf{Z}(x_n)$, would be the conditional expectation of $\mathbf{Z}(x_0)$, given the data $\mathbf{Z}(x_1), \ldots, \mathbf{Z}(x_n)$. Moreover, in the case of multivariate normality the conditional expectation is a linear function of the data. Motivated by this special case, the kriging estimator is given by

$$\mathbf{Z}^*(x_0) = \Sigma \lambda_i(x_0)\mathbf{Z}(x_i) \tag{1}$$

Although, as indicated, the weights, $\lambda_i(x_0)$, are a function of x_0, we shall write them simply as λ_i. We now need a method for determining the weights, which should be chosen so as to minimize some measure of the error of estimation. The variance is an obvious choice for this measure, partly because it leads easily to a system of linear equations. This variance and the assignment of the weights is dependent on the spatial correlation function but not on the data values.

Without some model assumptions it is not possible to formulate a solution. We will impose conditions on the random function which, of course, are not (statistically) testable in terms of the available data. These assumptions are usually called the Intrinsic Hypothesis.

(i) $\mathbf{E}\{\mathbf{Z}(x+h) - \mathbf{Z}(x)\} = 0$ for all points x and all vectors h

(ii) $0.5\mathrm{Var}\{\mathbf{Z}(x+h) - \mathbf{Z}(x)\} = \gamma(h)$ exists and depends only on h

It is possible to weaken (i) by assuming that the mean of $\mathbf{Z}(x)$ is representable as a linear combination of known functions usually taken to be monomials in the position coordinates. The function representing the mean is called the drift and hence (i) corresponds to constant drift. When $\gamma(h)$, called the variogram, depends only on the length of the vector h and not on its direction; the variogram and the random function are said to be isotropic. Otherwise they are said to be anisotropic. The most common form of anisotropy is called geometric. Such an anisotropy is removable with an affine transformation applied to the domain of the random function. An affine transfor-

mation combines a rotation with a stretching and/or a shrinking. If $\mathbf{Z}(x)$ is second order stationary, then $\gamma(h)$ is related to the covariance function $\mathbf{C}(h)$ as follows:

$$\gamma(h) = \mathbf{C}(0) - \mathbf{C}(h) \tag{2}$$

Whereas the covariance function is bounded and asymptotically goes to zero as the magnitude of h gets large, the variogram can exist when the covariance does not and it need not be bounded. It is also somewhat simpler to estimate the variogram since it is not necessary to separately estimate the mean of $\mathbf{Z}(x)$. Using the variogram, the variance of the error of estimation can be written as a quadratic form in the unknown weights. A Lagrange multiplier is introduced in order to obtain unbiasedness, i.e. to compensate for the constraint which results in unbiasedness. Minimizing the variance leads to the following systems of linear equations

$$\Sigma\lambda_i\gamma(x_i - x_j) + \mu = \gamma(x_j - x_0); \quad j = 1, \ldots, n$$
$$\Sigma\lambda_i = 1 \tag{3}$$

For the details of the derivation see Journel and Huijbrechts [1].

In the case of non-constant drift, additional Lagrange multipliers are required because of additional constraints needed to ensure unbiasedness. While the estimator does not change, the system of equations does and this is called universal kriging (as contrasted with ordinary kriging). In matrix form the system appears as

$$\begin{bmatrix} \mathbf{K} & \mathbf{E}^T \\ \mathbf{E} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \lambda \\ \mu \end{bmatrix} = \begin{bmatrix} \mathbf{K}_0 \\ 1 \end{bmatrix} \tag{4}$$

where $\mathbf{K}$ is the block consisting of the values of the variogram corresponding to pairs of sample locations and $\mathbf{K}_0$ consists of the values of the variogram for pairs consisting of a sample location and the point where an estimate is desired. $\mathbf{E}$ is a row of 1s, $\mathbf{E}^T$ is a column of 1s and μ is the Lagrange multiplier introduced because of the constraint. The minimized estimation variance is given by

$$\Sigma\lambda_i\gamma(x_i - x_0) + \mu \tag{5}$$

The estimator has a number of useful properties in addition to the unbiasedness and minimum variance conditions imposed in order to derive the kriging equations given in (3). It is easily seen that if the estimator is used to obtain an estimate at a location where there is data and if that data is used; then the estimated value is the observed value, that is, the estimator is exact. It is less obvious that the estimator has the desirable property of assigning larger weights to sample locations that are close to x_0. Note that although the sum of the weights is 1, the weights need not be positive and hence estimated values could be negative. Moreover, an estimated value could be larger than the largest data value although in practice neither of these circumstances occurs often.

The kriging equations given in (3) can also be written in terms of the covariance function (under an assumption of second order stationarity) by using (2). When there is no spatial dependence, i.e., the variogram is a positive constant for all non-zero lags; all the weights in the estimator are $1/n$ and the kriging estimator reduces to the arithmetic average. Certain sample location patterns may have the same effect even with a non-nugget variogram.

3 VARIOGRAM ESTIMATION AND MODELING

While empirical values of the variogram could be used in the coefficient matrix in (4), the variogram entries on the right hand side would have to be computed from a theoretical model or interpolated. In practice the estimated values of the variogram for certain distances is used to fit a theoretical model which is then used to compute the entries in the system of equations. Not all functions can be variograms; in particular the estimation variance should be non-negative and of course the coefficient matrix in (4) should be invertible. A weak form of positive definiteness is sufficient for both of these properties. (See Myers [2].) Unfortunately, it is not simple to test a function for either ordinary or conditional positive definiteness and in practice the theoretical model is constructed as a positive linear combination of known valid models. (This is known as a nested structure.)

Recall from matrix theory that a function $\boldsymbol{g}(\boldsymbol{x})$ defined in p-space is positive definite if for any points $x_1, \ldots, x_n$ and any coefficients $c_1, \ldots, c_n$ the quadratic form

$$\Sigma\Sigma c_j c_k \boldsymbol{g}(x_j - x_k) \tag{6}$$

is greater than or equal to zero. If a covariance function is used instead of a variogram, then the estimation variance is an expression of this form and covariances are known to be positive definite. In this case the matrix **K** in (4) will be positive definite. If it is strictly positive definite, then the coefficient matrix in (4) is invertible and hence the system has a unique solution. When a variogram is used instead of a covariance, then the negative of (6) must be non-negative, but only for those coefficients which add to zero. This is the definition of conditional positive definiteness. While it does not ensure that **K** is invertible, it does ensure that the coefficient matrix is invertible as shown in Myers [2]. Unfortunately, neither positive definiteness nor conditional positive definiteness are easily determined by a plot of a function and as a practical matter we consider only nested structures of certain known valid models.

These models have the advantage that they are characterized by a few parameters which are interpretable from the plots. All variograms are zero at lag zero, but they may have a discontinuity. This discontinuity is known as the nugget effect. The other two characteristics are the sill and the range. In terms of a covariance function the sill would be the value of the covariance at zero less the nugget and the range is the distance at which the covariance becomes zero. In general the plot of a covariance is the same as the plot of a variogram except that it is inverted. Five of the standard (isotropic) models are as follows:

Spherical

$$\gamma(r) = \begin{cases} C_1\{1.5(r/a) - 0.5(r/a)^3\}, & 0 \le r \le a \\ C_1, & r > a \end{cases}$$

Exponential

$$\gamma(r) = C_1\{1 - \exp - (r/a)\}, \quad 0 \le r$$

Gaussian

$$\gamma(r) = C_1\{1 - \exp(-(r/a)^2)\}, \quad 0 \le r$$

Each of these models corresponds to a covariance and for each C_1 is the sill and a is the range. In the case of the Exponential and the Gaussian models there is only an effective range generally taken to be $a' = 3a$.

Power

$$\gamma(r) = C_1 r^a, \quad 0 < a < 2$$

Nugget

$$\gamma(r) = \begin{cases} C_0, & r \ne 0 \\ 0, & r = 0 \end{cases}$$

The power model does not correspond to a covariance and does not have a sill or range. The case of $a = 1$ for the Power model gives the Linear model. For simplicity, however, most computer programs will denote C_1 as the sill and a the range even though these are not the true sill and range. One may be tempted to use a truncated linear model, i.e., linear up to a range and then constant thereafter, but such a model is not valid except in 1-space.

By analogy with estimators used in classical problems in statistics, the natural estimator for the variogram is the experimental or sample variogram given by

$$\gamma^*(\boldsymbol{h}) = \{0.5/\mathrm{N}(\boldsymbol{h})\}\Sigma[\mathbf{Z}(x_i + \boldsymbol{h}) - \mathbf{Z}(x_i)]^2 \tag{7}$$

where $x_i + \boldsymbol{h}$, x_i is a pair of sample locations $\boldsymbol{h}$ apart and $\mathrm{N}(\boldsymbol{h})$ is the number of such pairs. When the sample locations are not on a regular grid, there may not be more than one pair for any particular $\boldsymbol{h}$. In that case pairs are grouped into distance and angle classes. The choice of the classes

will be illustrated in the numerical example to follow. Note that in general we do not simply fit a curve on the plot of the sample variogram for several reasons: (a) the model must satisfy the appropriate positive definiteness conditions; (b) not all plotted points are of equal importance or reliability in estimating the value of the model for that lag; (c) there is some arbitrariness in the choice in the length of the distance classes and hence of the points for which an estimate is plotted. In general the variogram model, i.e., the values of the variogram, is more important for short than for long distances since the kriging estimator has the property that it gives larger weights to data at points close to the location where an estimate is desired and lesser weights to those far away. Consequently the modeling process is more critical for short than for long lags.

As an average of squares the sample variogram is not very robust. A number of other variogram estimators have been considered as well as direct estimation of the parameters in the model(s). These include the use of weighted least squares and maximum likelihood. The latter requires an assumption of normality. For a more extensive discussion see Myers [2].

While one may be interested in the variogram or covariance for reasons other than their use in the kriging equations; that is the most common reason. It is reasonable then to link the estimation to the behavior of the kriging estimator. Although it was not imposed when deriving the kriging equations, one of the properties of the kriging estimator is that it is exact (sometimes called perfect). If one estimates at a data location using all data including the data at that location, then the estimated value will be the observed value. This suggests a way of testing the variogram modeling. Namely one systematically deletes each sample location one at a time and estimates a value for that location using only the other data. This is known as cross-validation and it produces, for each location, the data value, an estimated value and the minimized estimation variance. If the variogram has been adequately modeled, then the estimated values should be 'close' to the observed values (in an appropriate sense). There are several possible measures of 'closeness' such as the mean error (which should be close to zero); the normalized mean square error (which should be close to one); the correlation of the estimated and observed values (which should be close to one); and the correlation of the estimated and the error of estimation (which should be close to zero). One can vary the parameters in the variogram model in order to try to optimize all the cross-validation statistics. This is illustrated later in the example.

4 PRACTICAL ASPECTS

The form of the kriging equations given in [3] makes it appear that all data are used to estimate all locations. In the case of large data sets, the coefficient matrix in (4) would be large and would require special methods for inversion. However, except for the nearby locations, the weights will be zero or nearly so. Hence in practice a moving neighborhood is used and only data from nearby locations (those in the search neighborhood) are used in the estimator. When implemented in software, the user must specify the size of the neighborhood and the number of locations to be used (both minimum and maximum). In practice it is seldom useful to set the maximum at more than 25 and the size of the neighborhood is generally taken to be the same as the range of the variogram.

Returning to the question of the choice of the number of distance classes and the width of the classes when computing the sample variogram for irregularly spaced data, there are two desirable features for the sample variogram which work against each other. In order that detailed information about the shape be gleaned from the graph, the classes should be narrow; but in order to enhance the reliability of each plotted point, the number of pairs used to produce that point should be large. Unfortunately for a given sample location pattern one cannot have both. Increasing the number of sample locations or using somewhat unusual sample location patterns as shown by Warrick and Myers [4] is the only way to attain both. Often however the data have already been collected and neither the number of locations nor the pattern can be altered. In this case it is neces-

sary to experiment with the number and the width of the classes to optimize the information gleaned from the plot. Recall that the behavior of the variogram or covariance is most important for short lags and in many cases one ignores the plot for longer lags. In particular the longest lag should not exceed half the maximum distance between sample locations. Note that there is no purpose in using lag distances smaller than the smallest inter-location distance.

As pointed out by Myers [5], stationarity is a property of the random function and not of data. If, as implicit in the geostatistical model, the data are viewed as one sample from one realization of the random function; then it is not possible to test for stationarity of the random function. Nevertheless, various aspects of the analyses of the data can be indicative of non-stationary. For example, the sample variogram estimates half of the mean square difference rather than half of the variance of the difference and these two only coincide in the case of a constant mean (of the random function). If the mean of the random function is not constant, then the plot of the sample variogram will exhibit a growth rate that is greater than is theoretically possible for a valid variogram. A valid variogram does not have to be bounded, but the growth rate as a function of the lag distance must be less than quadratic. Hence, if the plot of the sample variogram exhibits a quadratic or higher growth rate, there are no valid models to fit to the plot.

The solution is to decompose the random function into a stationary random function and a deterministic component representing the non-constant mean. The problem is how to determine this decomposition using only the data. There are several possible approaches. One is to first fit a trend surface to the data, which represents the deterministic component. Using the trend surface, residuals are computed and used to compute the sample variogram. However, trend computed in this fashion is not the same as the drift. In particular, the trend surface as an estimator of the drift is not optimal in the sense of minimizing the variance of estimation. Moreover, when these residuals are used to compute the sample variogram, it is a biased estimator of the variogram as shown by Sabourin [6]. While this technique is defensible as a practical technique, it has theoretical as well as computational disadvantages, as described by Neumann and Jacobsen [7] and Cressie [8]. One must use this technique with some care. The objective of using the residuals is to obtain a sample variogram that does not exhibit the rapid growth rate. The important characteristics to look for is the lack of quadratic growth in the variogram of the residuals when it is present in the variogram of the data.

Non-stationarity and anisotropy are sometimes difficult to distinguish and the non-stationarity often has a directional aspect. This interrelation can sometimes be exploited. For example non-stationarity may occur only in one direction as in the case of a subsurface flow, whereas in a perpendicular direction no non-stationarity is evident. In that case it may be possible to model the variogram using the sample variogram for the direction perpendicular to the direction corresponding to the non-stationarity using an isotropic model.

For economic or analytical reasons one often uses composited samples. This practice can cause problems in the geostatistical analysis in several ways. If there was no spatial dependence and the data were considered as a random sample from a single population; then the statistical effect of compositing corresponds to the use of samples of size greater than one. This is often desirable since there is a reduction in the variance of the estimator. When the data are spatially correlated, the effect is different and less desirable. If the data are used to estimate variograms, then it is necessary to compensate for the support of the samples. Compositing will lead to a reduction in the number of (apparent) sample locations. Hence the number of pairs for each plotted point is reduced and often the minimum distance between sample locations is increased thereby reducing information concerning the variogram at short lags. Finally there is also a theoretical consequence since the wrong variogram is being estimated, i.e., the sample variogram in this case estimates an averaged value of the variogram. This 'regularized' variogram will have a smaller sill and an apparent shorter range of dependence. This problem was recognized early

in mining applications where samples are frequently represented by sectioned drill hole cores or by samples composited from all the blast holes in a mining block. When the compositing corresponds to a reasonably regular geometric shape, there are formulas for inverting the process after the variogram has been modeled. If composited samples are used as data in the kriging estimator, then the kriging equations must also be modified.

To easily see the effect of compositing on variogram computations, consider a rectangular array of pixels, perhaps 100×100. If the values of the pixels were those of a random function with variance 100, the variogram should have a sill of 100. If now the pixels were composited either as 2×2 squares or 5×5 squares, then the sill will be reduced to 25 and 4 respectively. Moreover, for the original data there will be many pairs for a distance of 1 whereas for the 2×2 squares the minimum distances between centers will be 2 and 5 for the 5×5 squares. If the original range had been less than 2 then the sample variograms for the composited data will appear as pure nugget. The compositing effect is an intuitively obvious one since the data is being smoothed and hence local heterogeneities are being removed. The variance is reduced and short range correlations are smoothed out.

5 SOFTWARE

Before the advent of the personal computer equipped with a numeric co-processor, sufficient memory and disk space, it was necessary to use a mainframe computer or at least a minicomputer for geostatistical analysis. Commercial software, while available, tended to be very expensive, except for large users such as mining or oil companies. Consequently, most non-commercial users wrote their own or modified an available program. Geostatistics programs are frequently published in the journal *Computers and Geosciences*. The mathematics of a geostatistical package are not very demanding, since it basically only involves solving a system of linear equations together with appropriate input/output and internal data management. The advent of the microcomputer has not only made sufficient computing power easily available, it has also spawned several public domain packages one of which has become a defacto standard in terms of features, ease of use and the availability of good documentation. The Geo-EAS [9] package was produced under the sponsorship of the U.S. EPA Environmental Monitoring Systems Laboratory at Las Vegas, Nevada, and was officially released by EPA into the public domain. Moreover, EPA continues to provide support for the package. Geo-EAS includes a number of geostatistical components such as PREVAR, VARIO, XVALID, KRIGE as well as utilities such as TRANS, POSTPLOT and CONREC. The use of these will be described later in the example. An earlier package, STATPACK, produced by the U.S. Geological Survey at Denver, Colorado, was in part a downloading of components written for a mainframe computer and reflected a batch mode of operation to a considerable degree, whereas the Geo-EAS package is highly interactive with extensive graphics. The Geo-EAS package has been used in the following example. The Geo-EAS documentation should be consulted for program details. The executable code is available from the Las Vegas office and the source code is available from ACOGS [10] (Arizona Computer Oriented Geological Society).

TABLE 1

Format of Geo-EAS data file

Dallas Lead Study, average of first inch of soil, lead data			
4			
Easting			
Northing			
Lead			
ID #			
13.643	41.303	29.000	7
14.643	41.513	22.000	8
14.793	11.023	28.000	363
.....			
.....			

6 AN EXAMPLE

Several years ago the U.S. EPA was concerned with lead contamination in the air and soil of certain parts of Dallas, Texas and subsequent accumulation at significant levels in humans. It was asserted that the contamination came from a smelter which had been shut down, but it was still necessary to determine the extent of the pollution. To this end, data of various kinds were collected, together with soil samples. The data used in this example pertain only to the soil samples, for which sample position coordinates were given as well as the concentration of lead. Table 1 shows the header required on the Geo-EAS file together with the first few records of the data set. There were 361 data locations in the study. One of the columns in the data file is used only for sample identification.

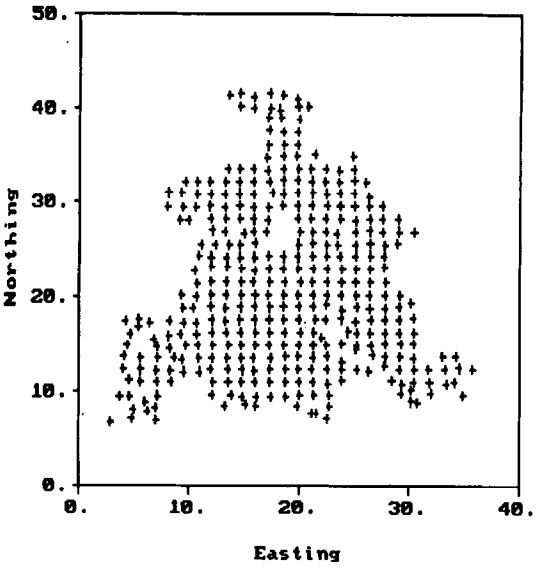

Fig. 1. Sample locations.

6.1 Exploratory analysis

Since the sample locations were spatially located, it is important to consider not only the empirical distribution of the data values but also the pattern of sample locations when considering the effect of possible outliers. It is useful to produce a plot of the locations either on screen or in hard copy. Some programs will code each location by the value of the variable at that location. Fig. 1 is such a plot for the Dallas Lead data, which was produced by POSTPLOT. The POSTPLOT incorporated into VARIO uses color to code the data values, but the separate POSTPLOT utility allows the use of symbols, symbol size, color and numeric

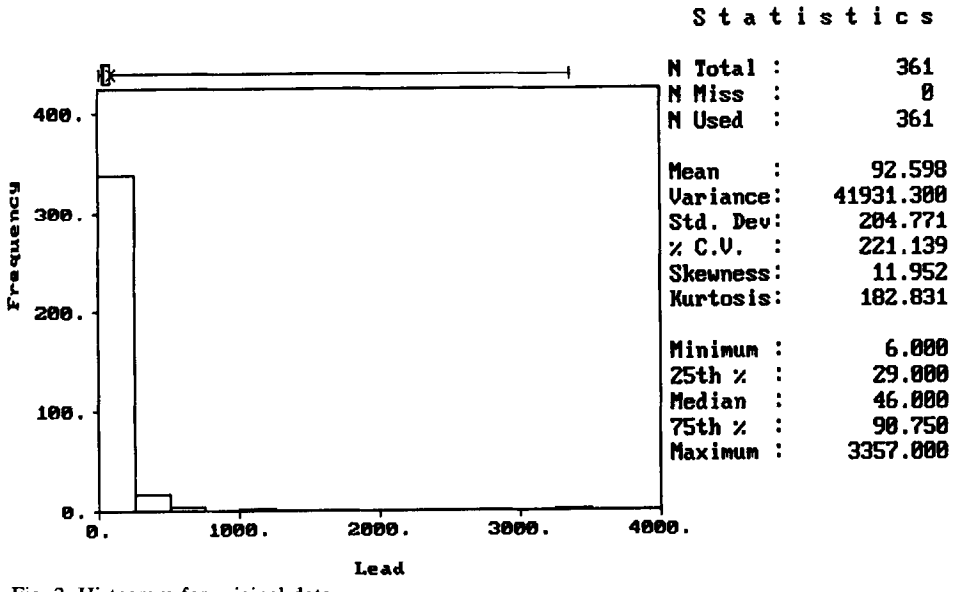

Fig. 2. Histogram for original data.

values. This also allows coding sample locations by the values of several variables at the same time. Perusal of this plot indicates the spread of the sample locations both in terms of distance and direction. This information can be useful in setting the parameters for the variogram estimation program. Based on the plot and possibly also on the histogram of the data values, it is sometimes useful to partition the data set spatially into one or more subsets. The histogram is important for providing evidence of bi-modality or skewness of the distribution. While the derivation of the kriging equations did not depend on any distributional assumptions, both the kriging estimator and the sample variogram are weighted averages and somewhat sensitive to skewed distributions. Chemical data are often found to be lognormally distributed and the histogram is a quick way to examine the data for such characteristics. Figs. 2 and 3, respectively, show the histograms of the original data and log transformed data; these figures were produced by STAT1. In the case of multivariate data it may be useful to consider principal components analysis, correspondence analysis or multiple regression in the exploratory analysis stage.

6.2 Variogram estimation

In Geo-EAS this is done in two stages. PREVAR sorts the data set into pairs of locations identified by separation distance and separation direction. This file is then the input for VARIO. PREVAR is the most restrictive of all the Geo-EAS programs with respect to the size of the data set since it will only use the first 150 records in the data file because of dimensioning and memory limitations. VARIO will compute the sample variogram given by (7) above as well as several other estimators. The user must specify the principal direction and the angle tolerance but the program will compute default distance classes although these may be changed by the user. The program will default to an initial direction of 0° and a tolerance of 90° which has the effect of ignoring the direction, i.e. the all-directional or isotropic sample variogram is computed. Unless information is available to indicate a particular direction of interest, the usual practice is to compute the all-directional variogram as well as directional variograms for 0°, 45°, 90°, 135°, each with a tolerance of 22.5°. If these appear to be the same and hence well represented by the all-direc-

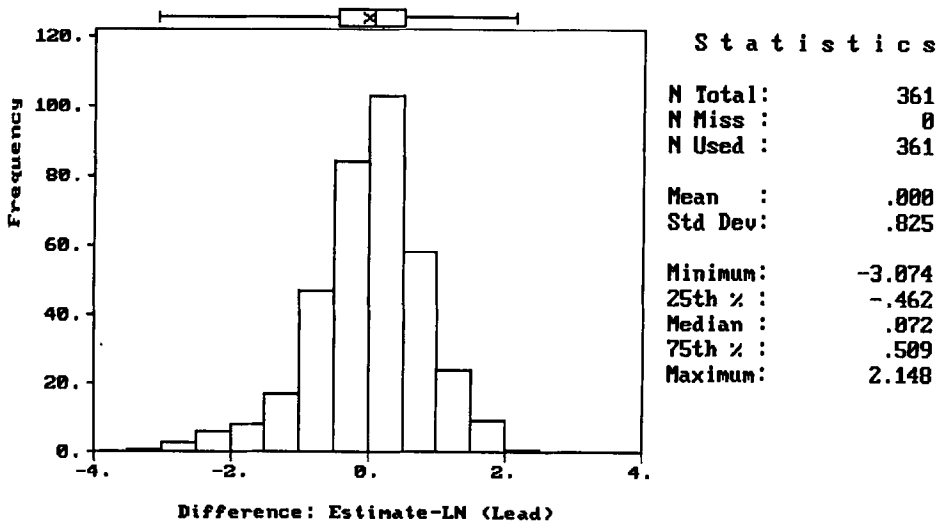

Fig. 3. Histogram for log transformed data.

tional variogram, then an isotropic model is used. Otherwise, one must attempt to determine the direction with the maximum range for the variogram as well as the ratio of the maximum range to the range in the direction orthogonal to it. These are then used to specify the (geometric) anisotropy. Note that in general the directional variograms will have fewer pairs for each plotted point than the all-directional variogram. Therefore for some directions the plot may be so erratic as to be difficult to determine a model and an isotropic model may be used for lack of adequate plots. The sample location pattern may mask or induce an (apparent) anisotropy. In general consideration of the location plot is essential for interpretation of the variogram plots.

As was noted earlier if the variogram plot exhibits a growth rate that is quadratic or larger, then it is necessary to use some method for removing the drift before modeling the variogram. VARIO includes an estimator known as the 'InCov' (in a previous version it was called 'Nonergodic'). The algorithm locally estimates the drift at each sample location, computes a residual and then computes an estimate of the variogram. It is called 'InCov' in part because it computes an estimate of the covariance which is internally converted to an estimate of the variogram. Some programs include the use of a trend surface to estimate the drift and residuals are then computed and used to compute the sample variogram.

This process of variogram estimation and modeling should be viewed as an interactive one since it may be necessary to try different variogram estimators, change the distance classes, limit the geographical extent of the data set, change the direction tolerances and then compare the plots. Note that it is not necessary to fit a model to the entire plot of the sample variogram since the short lag portion of the plot is the most important.

6.3 Choosing a model

In practice this means selecting a linear combination of the models provided in the program (almost certainly all of the models given above will be included) and for each term in the combination choosing the parameters. For example, we might choose a nugget with a sill of 5, a spherical with a sill of 10 and a range of 50 and a spherical with a sill of 15 and a range of 100. VARIO will allow interactive selection of the variogram types used in the linear combination as well as the parameters and then will plot that model against the sample variogram. It may be necessary to try different combinations to determine the (apparent) best fit. If some form of least squares is used to fit a model to the plot of the sample variogram, it should use some form of weighting. Even in that case the fitted model should not be considered to be optimal in terms of its use to determine the weights in the kriging estimator. For example, the plot can be altered, perhaps significantly, by changing the distance classes and hence the least squares fitted model may be different for these different plots. There are three initial pieces of information to look for in the plot: the magnitude of the nugget (or the absence of a nugget), the distance (called the range) at which the plot appears to level off (if it does) and if it does, the magnitude of that constant value. In general kriging is only useful when the nugget is moderately small compared to the total sill. The nugget may represent noise in the data such as measurement errors or it may reflect insufficient information at short lags to adequately determine the shape of the variogram. Figs. 4, 5 and 6 show the all-directional variogram for the transformed lead data, the 90° directional variogram and a model plotted against the all-directional variogram. The comparison of these is used to justify the use of an isotropic variogram. Table 2 tabulates the computed values of the variogram for the log transformed data (all-directional variogram).

6.4 Cross-validation

While one may be interested in estimating and modeling the variogram for purposes other than kriging, that is the most common use. Hence it is important to consider the relationship between the estimation/modeling process and the subsequent use of the variogram in the kriging equations. XVALID will produce the various cross-validation statistics described above. To use the program the user must specify the variogram model and the

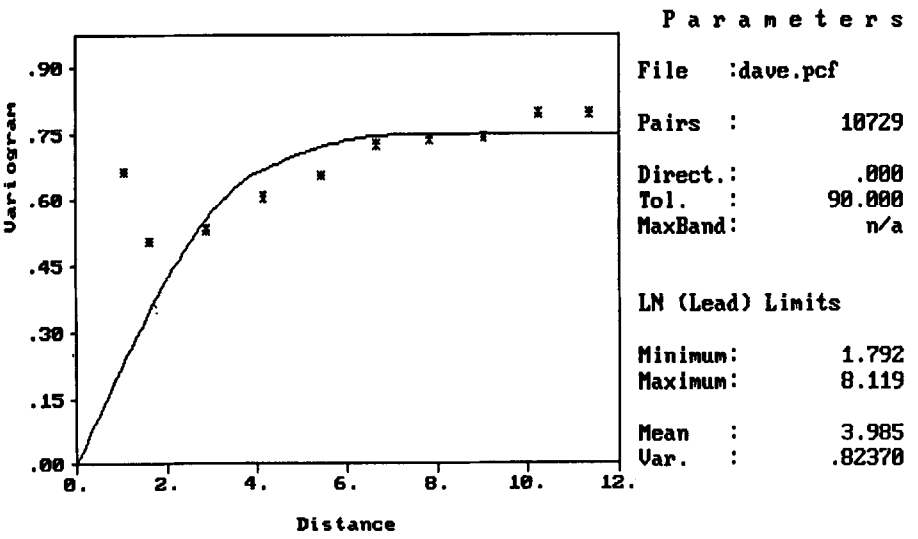

Fig. 4. Sample variogram, all directional and model.

search neighborhood parameters. The latter include indicating whether the neighborhood is circular or elliptical, the radius or lengths of the major and minor axes, the minimum number of sample points to be used for each estimation, and the maximum number of points for each estimation.

Ordinarily the search neighborhood is set to match the variogram model, i.e. as a default value the radius of search neighborhood is set equal to the range of the variogram. If an insufficient number of data points are found within the search neighborhood for a particular location, then that estimate will be skipped. If more than the speci-

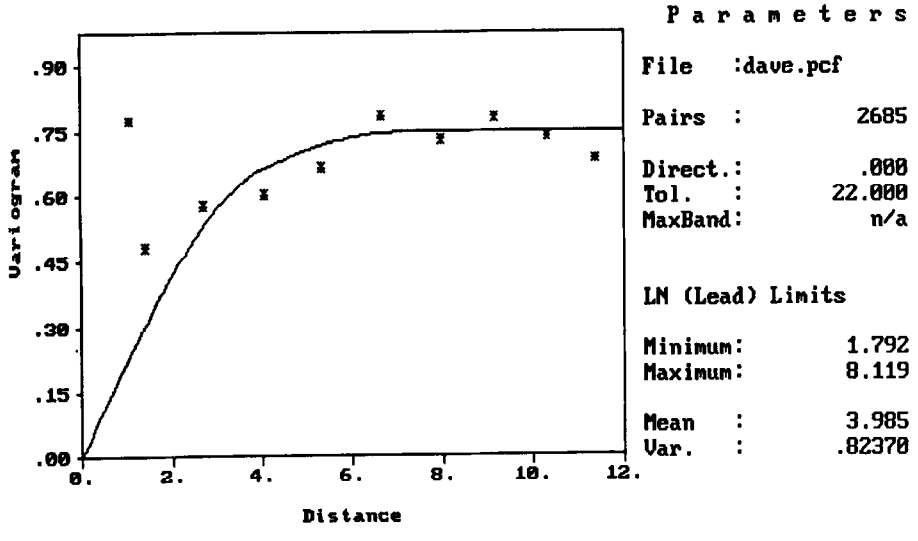

Fig. 5. Sample variogram, 0° direction and 22° tolerance.

TABLE 2

Computed values of variogram, all directional

Lag No.	Number of pairs	Average distance	γ^*
1	42	1.082	0.664
2	576	1.628	0.505
3	827	2.874	0.531
4	1173	4.159	0.608
5	1290	5.437	0.655
6	1294	6.662	0.725
7	1391	7.840	0.738
8	1503	9.038	0.740
9	1370	10.231	0.795
10	1263	11.350	0.797

TABLE 3

Nested structure for variogram model

Term No.	Type	Sill	Range
1	nugget	0.0	0.0
2	spherical	0.4	4.0
3	spherical	0.3	7.0
4	spherical	0.05	9.0

fied maximum number of points are found in a neighborhood, then only the closest ones are used. In XVALID this process can be seen graphically as it occurs by choosing the Debug option. Table 4 shows the Results screen giving the numerical

TABLE 4

Cross-validation results

	Data	Estimate	Krg. std. dev.	Normalized
Minimum	1.792	2.057	0.362	−6.024
25th percentile	3.367	3.556	0.484	−0.842
Median	3.829	3.934	0.494	0.169
75th percentile	4.500	4.348	0.506	1.045
Maximum	8.119	6.083	0.811	4.728
Mean	3.985	3.982	0.501	−0.003
Std. dev.	0.908	0.636	0.044	1.641

statistics resulting from cross-validation. Figs. 7 and 8 show scatterplots of 'estimated vs. observed' and of 'estimated vs. normalized error of estimation'. These results cannot be judged against absolutes but rather they provide a tool for ranking the choices for the variogram model. Therefore it is usually necessary to try several possibilities. Note that the cross-validation statistics can be affected by the choice of the search neighborhood parameters. In some instances these statistics are relatively insensitive to changes in the variogram or the search neighborhood.

6.5 Kriging

This is perhaps the easiest step since all the parameter choices will have been made at previous

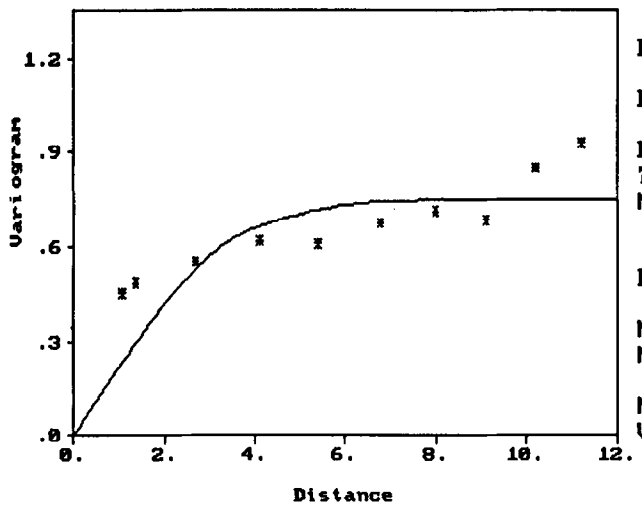

Fig. 6. Sample variogram, 90° direction and 22° tolerance.

```
P a r a m e t e r s

File   :dave.pcf

Pairs  :          2638

Direct.:         90.000
Tol.   :         22.000
MaxBand:            n/a

LN (Lead) Limits

Minimum:          1.792
Maximum:          8.119

Mean   :          3.985
Var.   :         .82370
```

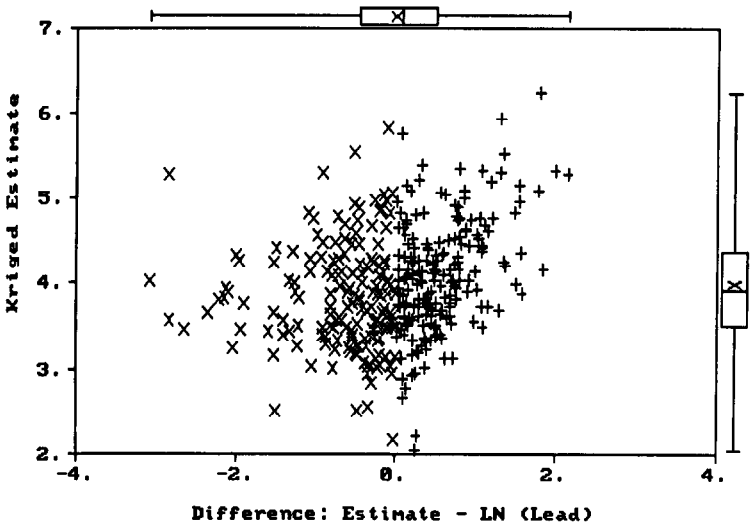

Fig. 7. Scatter plot of Z* vs. Z.

steps, namely the choice of the variogram model (Table 3) and search neighborhood parameters. KRIGE provides for two forms of kriging, point and block (block kriging corresponds to estimation of spatial averages). In each case the estimates are produced for a grid whose position and mesh are chosen by the user. If block kriging is chosen then the size of the block must also be chosen. The output from either form of kriging

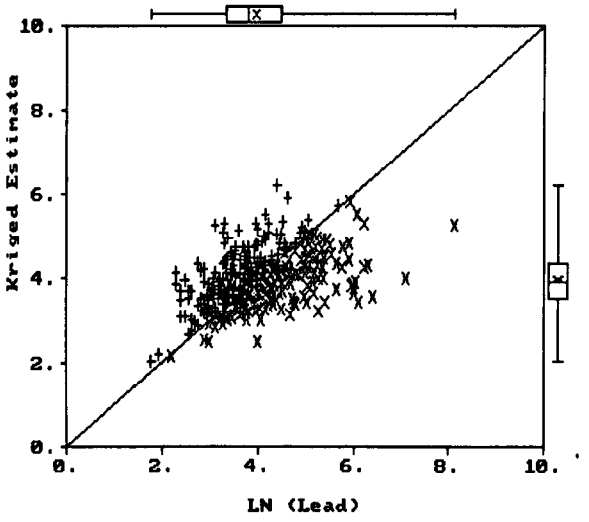

Fig. 8. Scatter plot of Z* vs. Z*−Z.

may be used with a contouring package to produce a map of the estimated values and also a map of the kriging standard deviations. The latter is useful in interpreting the former since the contour map of the kriged variable is not of the same degree of reliability over the entire region. Geo-EAS includes a component called CONREC which will produce an on-screen contour plot. It does this by first producing a metacode file which can be VIEWed or it can be transformed into a Hewlett-Packard plotter file. With the exception of PREVAR all of the computation programs in Geo-EAS produce Geo-EAS files and hence can be used as input files for other components. Kriging is a smoothing process and hence the variance of the kriged values is less than that of the original data. The use of the log transform does cause some problems since a simple exponentiation of the estimate will result in a bias. The necessary corrections can be found in Journel [11]. The plots and variogram model used above do not necessarily represent an optimal modeling of this data set but are given here for illustration only.

One of the best ways to grasp the usefulness of geostatistics is to apply the various components of Geo-EAS to a particular data set and to take advantage of the interactive capabilities of the programs. This allows the user to quickly see the effect of choosing different variogram models and

different variogram and search neighborhood parameters.

7 EXTENSIONS

There are a number of extensions of kriging that have been developed. Some are linear and others are non-linear. Each has been developed in response to some particular application problem or perceived lack in kriging. In many applications there may be several variables of interest. For example in many copper mines molybdenum is also extracted from the ore; and the two metal grades are found to be not only separately spatially correlated but they are also correlated with each other. Cokriging is a linear estimation method which incorporates both the spatial and inter-variable correlation and hence data on both variables. In hydrology there are often many variables of interest, such as hydraulic conductivity, permeability, and porosity. Early applications emphasized the use of data from one variable to enhance the estimation on another more important variable. Often the proxy variable was one that was easier or cheaper to sample. A complete presentation is found in Myers [12,13].

Because the estimator is a linear combination of the data and the variogram estimator is an average of squared differences; both of these are sensitive to 'outliers'. In some cases this effect is reduced by a transformation such as a logarithm, this being motivated by an assumption of or an apparent indication of log normality. Because in many applications one is interested in a particular cutoff value and the associated proportion of an area or volume above (or below) the cutoff, the use of an indicator transform is indicated. Treated as a stationary random function the indicator is closely related to the marginal probability distribution; therefore a linear estimator for indicators can be used leading to a set of kriging equations. This work is largely due to Journel [14,15] and his students at Stanford.

One of the advantages of viewing spatial data as a sample from a realization of a random function is that it leads rather naturally to the idea of generating other realizations perhaps conditioned on the data. Simulation of a random function could be viewed in several contexts, i.e. if the new realizations are to be from the 'same' random function, then certain characteristics must be preserved. In geostatistics this is usually taken to mean that the first and second order moments are preserved as well as the marginal distribution. Other characteristics could be considered. The emphasis on the moments is partially a consequence of the fact that kriging only depends on the first two moments. Simulation of random functions in 1-space, i.e. time series, was well-known but there remained the question of producing simulations in 2- or 3-space and preserving the spatial dependence. The Turning Bands method of Matheron [16] produces a simulation in higher dimensional space by producing many independent simulations in 1-space. Simulations are useful for a variety of applications such as mine planning, modeling the spatial variability in a pollutant dispersal, evaluating potential sites for waste disposal to name a few. Kriging provides a method for conditioning the simulation to the data.

For most distribution based statistical techniques it is possible to determine a sample size that is sufficient with respect to some criterion such as the power of a test or the width of a confidence interval. In the case of spatially located data however it is not sufficient to consider only the number of locations. The sample location pattern is also important. Since the kriging variance does not depend on the data as such (although in general the data is used to estimate and model the variogram); this variance can be computed a priori for a given model and sample location pattern. Hence the sample location patterns can be ranked by the kriging variances. Determination of an optimal sampling pattern may arise in several different contexts. One may have a large number of locations such as wells and wish to monitor only a smaller number but maintain nearly the same level of information. One may wish to add additional locations to an existing pattern (adding only one is easy, simply find the place where the kriging standard deviation is largest on the contour plot) or a set of locations is to be chosen with an assumed variogram model. This problem has received considerable attention in the geostatisti-

cal literature and a brief history is given by Barnes [17].

8 APPLICATIONS

As indicated by previous citations geostatistics has been and is being applied to a wide variety of problems in a number of disciplines. These range from the characterization and contouring of the likelihood of sudden soil collapse in a semi-arid environment as detailed by Ali et al. [18–21] to the assessment of the movement of a multi-pollutant plume by Myers [22]. Myers [23] and Myers et al. [24] describe the application of kriging to hydrogeochemical data collected as a part of the NURE project. Morkoc et al. [25] use kriging with generalized covariances in a soil physics application. Yates et al. [26] use a non-linear form of kriging together with field measured reflectance and soil temperature data to evaluate field crop heterogeneities. Tabor et al. [27,28] compare nitrogen content in soil with nitrogen content in cotton plant petioles as an indicator of need for fertilizing. Myers and Carr [29] re-examine multivariate data from a Wyoming Bentonite clay deposit and compare multivariate geostatistics with dimension reduction using principal components analysis.

9 COMPARISON WITH OTHER METHODS

One of the motivations for the introduction and application of kriging was that it incorporated the spatial nature of the data and it has the property that data from nearby locations are weighted more heavily than those far away when estimating at a particular location. There were and are other methods that have similar characteristics. One of these is Trend Surface Analysis wherein the data are fitted to a polynomial in the position coordinates by least squares. In general this method estimates at an unsampled location by evaluating a regional value as represented by the polynomial but does not incorporate the local fluctuations. Moreover, as shown explicitly in Marcotte and David [30], Trend Surface Analysis produces the same estimated values as Universal Kriging with a pure Nugget variogram except at the data locations (Trend Surfaces are not exact interpolators). This is a consequence of treating the local fluctuations as 'errors' or noise to be removed. The estimated variance of the error of estimation is constant and reflects an estimate of the magnitude of the errors.

One of the simplest ways to assign weights in a linear estimator is to make them inversely proportional to the distance of the sample location to the point to be estimated, i.e. inverse distance weighting (IWD). This technique has some of the same characteristics and advantages as kriging but not all. In particular the weights in the kriging estimator are also affected by the intersample location distances and directions whereas for IWD only the relationship of sample locations to the point to be estimated are used. The process of estimating the variogram, although not perfect, does have the capability of adapting to different variates and also incorporating anisotropies. It is possible to empirically optimize the exponent on the distance function in IWD by the use of a training set as shown by Kane et al. [31], but one must then assume that the modeling is transferable to the remainder of the data set. While no assumptions are necessary, nor is there even a derivation for the IWD estimator, it might be best thought of as (nearly) a special case of kriging.

The kriging estimator can be rewritten as follows by substituting the solution for the weight vector into the estimator.

$$\mathbf{Z}^*(x_0) = \Sigma b_i \gamma(x_i - x_0) + a \tag{8}$$

where

$$\Sigma b_i \gamma(x_i - x_j) + a = \mathbf{Z}(x_j); \quad j = 1, \ldots, n$$

$$\Sigma b_i = 0 \tag{9}$$

In this form the estimator is analogous to a thin plate spline and in fact the thin plate spline can be obtained as a special case of kriging. For an elementary demonstration see Watson [32] and also Myers [2]. The smoothing spline can be obtained as a special case of cokriging.

10 GLOSSARY

Anisotropic: This term is applied both to a random function and to its variogram (or covariance) when the values of the variogram depend on both the distance and the direction. Also see Isotropic.

Cross-validation: A method for comparing two or more conjectured variogram (or covariance) models. The technique depends on Jackknifing the data and on the exactness of the kriging estimator.

Drift: The expected value of a random function, it may be constant or it may depend on the coordinates of the location. In order for a random function to be stationary, second-order stationary or to satisfy the Intrinsic Hypothesis; the drift must be a constant. The drift is a characteristic of a random function and not of data.

Exact(ness): A property of an estimator/interpolator, namely that if estimating a value at a data location and if that data value is used in the estimation; then the estimated value will coincide with the data value. In some literature this is called Perfect.

Intrinsic Hypothesis: A weak form of stationarity for a random function sufficient for deriving the kriging equations corresponding to the (Ordinary) kriging estimator. See (i) and (ii).

Isotropic: A term applied both to a random function and to its variogram. See anisotropic which is the complementary property.

Kriging Equations: A set of linear equations whose solution includes the values of the weights in the kriging estimator.

Kriging Estimator: While the estimator may be a linear or a non-linear function of the data, in both instances the weights in the estimator are determined by requiring the estimator to be unbiased and have minimum error variance.

Kriging Variance: The minimized value of the estimation variance, i.e., the variance of the error of estimation. This variance is not data dependent but rather is determined by the variogram and the sample location pattern as well as the location of the point to be estimated relative to the sample locations.

Nugget: The variogram may exhibit an apparent discontinuity at the origin. The magnitude of the discontinuity is called the nugget.

Positive Definite: A term applied both to matrices and to functions, (Auto) covariance functions must be positive definite whereas the negative of variograms must be conditionally positive definite. Conditional positive definiteness is a weaker condition.

Random Function: A random function may be seen in two different forms; it may be thought of as a collection of dependent random variables with one for each possible sample location. Alternatively it may be thought of as a 'random variable' whose values are functions rather than numbers.

Range (of a variogram): The distance at which the variogram becomes a constant. The Power model does not have a (finite) range. The Exponential and Gaussian models have only an apparent range.

Sill (of a variogram): The value of the variogram for distances beyond the range of the variogram. The Power model does not have a sill.

Spatial Correlation: Used both as a generic term to denote that data at two locations are correlated in some sense as a function of their locations and also to denote the value of a spatial structure function such as a variogram or (auto)covariance for a pair of locations.

Stationarity (of a random function): Several different forms of stationarity are utilized in geostatistics. Stationarity, in one of its forms, is a property of a random function rather than of a data set. It expresses the property that certain joint distributions are translation invariant or that certain moments of the random function are translation invariant. See second order stationarity and the Intrinsic Hypothesis.

Support: The term is used in both a mathematical and in a physical sense. Many, if not most variables of interest in geostatistics, such as the concentrations of chemical elements or compounds, only have values at 'points' in an idealized sense although the random function treats them in this manner. The data values are usually associated with a physical sample having a length, area or volume; the concentration then represents an average concentration over this length, area or volume. This length, area or volume is called the support. Although it is common to report laboratory analyses in such a way as to not reflect the

original support, non-point support has a significant effect on the variogram modeling process and there is a significant difference in estimating the average value over a large volume and in estimating the average value over a small volume. The kriging estimator and equations allow this to be incorporated.

Trend: While sometimes used interchangeably with the term 'drift', in geostatistics the two are considered separate. The term is usually reserved to denote the deterministic representation obtained by the use of Trend Surface Analysis, i.e., a functional representation for spatially located data (usually taken as a polynomial in the position coordinates). The 'trend' is obtained by a least squares fit to the data. As an estimator to the mean of a random function it is sub-optimal. If the residuals from trend surface analysis are used to model the variogram, a biased variogram estimator results.

Variogram (originally called semi-variogram): This function quantifies the spatial correlation and in the case of second order stationarity it is expressible in terms of the (auto)covariance function. See part (ii) in the Intrinsic Hypothesis and eq. (2). In order to apply kriging to a data set it is necessary to model the variogram. The variogram must satisfy certain positive definiteness conditions.

11 ACKNOWLEDGEMENTS

Although the research described in this article has been funded wholly or in part by the U.S. Environmental Protection Agency through a Cooperative Research Agreement with the University of Arizona, it has not been subjected to Agency review and therefore does not reflect the views of the Agency and no official endorsement should be inferred.

BIBLIOGRAPHY

Books and journals

Because geostatistics is primarily a product of certain areas of application, references are somewhat more difficult to find since one may have to consult literature foreign to one's principal interest or to utilize books concentrating on a particular area of application. There are essentially no books on geostatistics in the statistical literature although a few touch on the subject. The first two comprehensive volumes illustrate this very well in that both focus on applications in mining, these are

A.G. Journel and Ch. Huijbrechts, *Mining Geostatistics*, Academic Press, London, 1978.

M. David, *Geostatistical Ore Reserve Estimation*, Elsevier, New York, 1977.

There are two more elementary books also concentrating on mining applications:

I. Clark, *Practical Geostatistics*, Applied Science Publishers, London, 1979.

J.M. Rendu, *An Introduction to Geostatistical Methods of Mineral Evaluation*, S. African Inst. Mining and Metallurgy, Johannesburg, 1978.

Although geostatistics has become a recognized area of interest it has not spawned any journals. Based on the discussion at the Third International Geostatistical Congress in Avignon, France, in September 1988, it is not likely that there will be a separate journal. The subject does have a newsletter, which is sponsored by the North American Council on Geostatistics, a non-membership group which organizes an informal meeting each summer. To obtain information concerning the newsletter and to ask to be put on the mailing list, contact the Editor, Mohan Srivastava, NACOG Newsletter Editor, FSS International, 3900 Quebec Street, Vancouver, British Columbia, Canada V5V 3K8.

There is an international counterpart, The International Geostatistical Society, which publishes De Geostatisticis and the Editor is Hans Wackernagel, Centre de Geostatistique, 35, rue St. Honoré, 77305 Fontainebleau, France.

There are several principal journals which regularly publish papers pertaining to geostatistics. These include:

Mathematical Geology (International Association of Mathematical Geologists);
Water Resources Research (American Geophysical Union);
Journal of the Soil Science Society of America;
Sciences de la Terre (appears on an irregular basis and issues may be purchased individually; a French journal but with nearly all articles in English);
Computers and Geosciences (International Association of Mathematical Geologists), a good source for computer programs.

This list is changing rather rapidly and the above should not be considered complete or exclusive. In addition there are the proceedings of at least three international conferences that are of signal importance:

M. Guarascio, M. David and C. Huijbregts (Editors), *Advanced Geostatistics in the Mining Industry* (NATO ASI 1975), Reidel, Dordrecht, 1976.
G. Verly et al. (Editors), *Geostatistics for the Characterization of Natural Resources* (NATO ASI 1983), Reidel, Dordrecht, 1984.
M. Armstrong (Editor), *Geostatistics* (Third International Geostatistical Congress 1988), Reidel, Dordrecht, 1989.

Selected papers

1 A.G. Journel and Ch. Huijbrechts, *Mining Geostatistics*, Academic Press, London, 1978.
2 D.E. Myers, Interpolation with positive definite functions, *Sciences de la Terre*, 28 (1988) 251–266.
3 D.E. Myers, On variogram estimation, *Proceedings of the First International Conference on Statistical Computing, Cesme, Turkey, March 29 – April 2, 1987* (Mathematical and Management Series, Vol. 26), American Sciences Press, Syracuse, NY, in press.
4 A. Warrick and D.E. Myers, Optimization of sampling locations for variogram calculations, *Water Resources Research*, 23 (1987) 496–500.
5 D.E. Myers, To be or not to be stationary, *Mathematical Geology*, 21 (1989) 347–362.
6 R. Sabourin, Application of two methods for interpretation of the underlying variogram, in M. Guarascio, M. David and C. Huijbregts (Editors), *Advanced Geostatistics in the Mining Industry*, Reidel, Dordrecht, 1976, pp. 101–112.
7 S. Neumann and E. Jacobsen, Analysis of nonintrinsic spatial variability by residual kriging with application to regional ground water levels, *Mathematical Geology*, 16 (1984) 499–521.
8 N. Cressie, Fitting variograms by weighted least squares, *Mathematical Geology*, 17 (1985) 563–586.
9 U.S. Environmental Protection Agency, *Geo-EAS, Version 1.2.1*, the executable code is available from: Evan Englund, EPA EMSL-LV (EAD), P.O. Box 93478, Las Vegas, NV 89193-3478, U.S.A.
10 ACOGS, the Geo-EAS source code together with a programmer's maintenance manual is available at a cost of $91.50 from: ACOGS, P.O. Box 44247, Tucson, AZ 85733–4247, U.S.A.
11 A.G. Journel, The log-normal approach to predicting local distributions of selective mining unit grades, *Mathematical Geology*, 12 (1980) 285–301.
12 D.E. Myers, Cokriging: the matrix form, *Mathematical Geology*, 14 (1982) 249–257.
13 D.E. Myers, Multivariate geostatistics for environmental monitoring, *Sciences de la Terre*, 27 (1988) 411–428.
14 A.G. Journel, Non-parametric estimation of spatial distributions, *Mathematical Geology*, 15 (1983) 445–468.
15 A.G. Journel, The deterministic side of Geostatistics, *Mathematical Geology*, 17 (1985) 1–16.
16 G. Matheron, The intrinsic random functions and their applications, *Advances in Applied Probability*, 5 (1973) 239–265.
17 R.A. Barnes, Partial history of spatial sampling design, *Geostatistics Newsletter*, 3 (2) (1989) 10–13.
18 M. Ali, E. Nowatzki and D.E. Myers, Probabilistic analysis of collapsing soil by indicator kriging, *Mathematical Geology*, 22 (1990) 15–38.
19 M. Ali, E. Nowatzki and D.E. Myers, Geostatistical methods to predict collapsing soils, in *Proceedings of the XIIth International Conference on Soil Mechanics and Foundation Engineering, August 1989, Rio de Janeiro, Brazil*, Balkema, Rotterdam, 1989, pp. 567–570.
20 M. Ali, E. Nowatzki and D.E. Myers, Use of geostatistics to estimate the probability of occurrence of collapse susceptible soils in Tucson, AZ, in F.H. Kulhawy (Editor), *Proceedings of the 1989 Foundation Engineering Congress, 25–29 June 1989, Evanston, Illinois*, American Society of Civil Engineers, New York, 1989, pp. 176–190.
21 M. Ali, E. Nowatzki and D.E. Myers, Geostatistical techniques to estimate collapse-related soil parameters, in R.J. Watters (Editor), *Engineering Geology and Geotechnical Engineering*, Balkema, Rotterdam, 1989, pp. 289–296.
22 D.E. Myers, Borden field data and multivariate geostatistics, in M.A. Ports (Editor), *Hydraulic Engineering*, American Society of Civil Engineers, New York, 1989, pp. 795–800.
23 D.E. Myers, Kriging hydrogeochemical data, in D.F. Merriam (Editor), *Current Trends in Geomathematics*, Plenum Press, New York, 1989, pp. 117–142.
24 D.E. Myers, C.L. Begovich, T.L. Butz and V. Kane, Variogram models for regional groundwater geochemical data, *Mathematical Geology*, 14 (1982) 628–644.
25 A. Morkoc, J. Biggar, D. Neilsen and D.E. Myers, Kriging with generalized covariances, *Journal of the Soil Science Society*, 51 (1987) 1126–1130.
26 S. Yates, A. Warrick and D.E. Myers, Disjunctive kriging II: Examples, *Water Resources Research*, 22 (1986) 623–630.
27 J. Tabor, A. Warrick, D. Pennington and D.E. Myers, Spatial variability of nitrate in irrigated cotton II: Soil nitrate and correlated variances, *Journal of the Soil Science Society America*, 49 (1985) 390–394.

28 J. Tabor, A. Warrick, D. Pennington and D.E. Myers, Spatial variability of nitrate in irrigated cotton I: Petioles, *Journal of the Soil Science Society of America*, 48 (1984) 602–607.
29 D.E. Myers and J. Carr, Cokriging and principal component analysis: Bentonite data revisited, *Sciences de la Terre*, 21 (1985) 65–77.
30 D. Marcotte and M. David, Trend surface analysis as a special case of IRF-k kriging, *Mathematical Geology*, 20 (1988) 821–824.
31 V. Kane, C. Begovich, T. Butz and D.E. Myers, Interpolation of regional geochemistry using optimal interpolation parameters, *Computers and Geosciences*, 8 (1982) 117–136.
32 G. Watson, Smoothing and interpolation by kriging and with splines, *Mathematical Geology*, 16 (1984) 601–616.