HILGARDIA

A JOURNAL OF AGRICULTURAL SCIENCE PUBLISHED BY THE CALIFORNIA AGRICULTURAL EXPERIMENT STATION

Volume 51 • Number 3 • June 1983

Geostatistical Theory and Application to Variability of Some Agronomical Properties

S. R. Vieira, J. L. Hatfield,
D. R. Nielsen, and J. W. Biggar

UNIVERSITY OF CALIFORNIA DIVISION OF AGRICULTURAL SCIENCES

In agronomic problems the sampling procedure may create some confusion and bias in the analysis. Geostatistics provides a method for the analysis of the spatial and temporal properties in a data set and a method of interpolation between selected points. This paper describes the theory of geostatistics and its application to selected agronomic problems. Geostatistics considers a set of data collected in either space or time at discrete intervals. These samples may be correlated with each other to provide some unique information about the parameters which would not be detected in the classical statistical methods. Through the application of geostatistics to this type of problem, we can estimate the spatial or temporal dependence of samples and from this knowledge arrive at an estimation of the sampling procedures or structure at a field. The application of these techniques is shown for air temperature, surface temperature, yield, clay content, and fertilizer content in various fields and reveals the versatility of the techniques.

Geostatistics also allows for the evaluation of the dependence between two parameters in either time or space. From this information it is possible to develop sampling procedures which would allow the more costly or time consuming variable to be sampled less frequently and estimated from the other variable by the method of kriging. This report summarizes all of these techniques and provides several different examples of their utilization. Examples of the computer code are provided for the reader wishing to apply these techniques.

THE AUTHORS:

S. R. Vieira is a former graduate research assistant; presently Soil Scientist, Institute Agronomico de Campinas, Campinas, Brazil.

J. L. Hatfield is Associate Professor of Meteorology and Land, Air and Water Resources and Associate Biometeorologist in the Experiment Station, University of California, Davis.

D. R. Nielsen is Professor, Soil and Water Science and Land, Air and Water Resources and Water Scientist in the Experiment Station, University of California, Davis.

J. W. Biggar is Professor, Land, Air, and Water Resources, University of California, Davis.

S. R. Vieira, J. L. Hatfield, D. R. Nielsen, and J. W. Biggar

Geostatistical Theory and Application to Variability of Some Agronomical Properties[1]

INTRODUCTION

Researchers have been working on variability problems since the early 1900's (Montgomery, 1913; Robinson and Lloyd, 1915; Pendleton, 1919). Very careful field experiments have been done to determine the effects of soil variability on plot yields (Montgomery, 1913; Smith, 1910), nitrification (Waynick, 1918), and nitrogen and carbon (Waynick and Sharp, 1919); a variety of sampling schemes were used that allowed maximum coverage over the field with the distances between samples known. Further, Harris (1920) presented a collection of analyzed data sets using an equation very similar to what we know as block variance. Sampling soon became a discipline: by the early 1960's many books had been written on sampling theory, and they had such common schemes as random and stratified sampling.

Analysis of data used to be conducted using such classical statistical methods as analyses of variance (Beckett and Webster, 1971), which assume independence between samples. This assumption is commonly accepted when the observation follows a normal distribution. However, a given set of observations may be normally distributed, regardless of the sampling distance, and a good deal of autocorrelation may exist, depending on the process. Autocorrelation requires the use of a relatively new statistical approach called geostatistics.

Matheron (1963, 1971) developed a theory he called "Theory of Regionalized Variables," which describes the fundamentals of geostatistics. In that theory, a regionalized variable is a numerical space function which varies from one place to the next with apparent continuity but which varies in a manner that

[1]Accepted for publication December 14, 1982.

cannot generally be represented by an ordinary workable function (Matheron, 1963). The kriging method of interpolation, which is based on the theory of regionalized variables while using the degree of autocorrelation between adjacent samples, estimates values for any coordinate position within the domain measured without bias and with minimum variance. In this sense, kriging is an optimum interpolator (Burgess and Webster, 1980a). As a result, estimation can be done for as many values as a computer budget will allow. Very precise contour maps can be drawn for space-distributed variables, and short-time intervals can be estimated for time variables, reducing sampling and analysis costs. Kriging has been used successfully in mining (David, 1970), hydrology (Delhomme, 1976), and soil science (Hajrasuliha et al., 1980; Vieira et al., 1981).

Similar to one variable being autocorrelated in either space or time, when two or more variables are measured for the same domain, they may be correlated to each other or cross-correlated two by two. This allows values of one variable to be estimated using the measured values of all the variables. This estimation method, called cokriging, is particularly useful when one variable is more difficult to measure than the other variables and consequently has fewer samples than the other variables with which it is cross-correlated. For example, in most instances, agrometeorological observations such as air and surface temperatures are easily obtained by either ground or remote measurements. Intuitively, we can expect a cross-correlation between these temperatures and such other variables as surface water content of a bare soil, water evaporation, sand, clay, and organic matter content of the soil surface. Values of some of these variables are more difficult to obtain than others, and cross-correlation functions between them would provide ways to estimate values for the unrecorded times or locations using the available data.

Therefore, there are potential applications of geostatistical methods for analysis of variability of agronomical observations. The purpose of this article is primarily to expose the reader to the fundamentals of geostatistics and to show some examples of applications using field data. The expression and applications of the theory have been intentionally kept detailed to provide an easily understandable reference for particular cases; we feel that most of the references on the subject are difficult to read, mainly because the theories are too general. For example, in most geostatistical applications to mining cases, the samples are defined in three-dimensional space, as a bulk volume (Blais and Carlier, 1968; Ugarte, 1972; Journel, 1974), and one of the most complete derivations of the cokriging system is done by Journel and Huijbregts (1978) for a generic number of variables defined in three-dimensional space. However, in agronomic studies, we will mostly be dealing with one- or two-dimensional samples and only two cross-correlated variables, so it is important that these theoretical derivations be available to the interested agricultural user for reference.

THEORY

The Semivariogram

Consider a field of area S, for which we have measured a set of n values $\{z(x_i),\ i = 1,n\}$, in which each x_i identifies a coordinate position. The coordinate position represented by the x_i's can be either space or time; for simplicity of presentation, we will restrict our discussion to space. Each x_i represents a pair of coordinates $(\hat{x}_i, \hat{y}_i)$, $i = 1,n$. Each $z(x_k)$ can be considered a particular realization of a certain random variable, $Z(x_k)$, for a particular fixed point, x_k. The regionalized variable $Z(x_i)$, for all x_i inside S, can be considered a realization of the set of random variables $\{Z(x_i)$, for all x_i inside S$\}$. This set of random variables is called a random function and is written $Z(x_i)$ (Journel and Huijbregts, 1978).

Until the late 1960's, analysis of field data such as that described above has been treated under the assumption of statistical independence or random spatial distribution so as to allow for the use of statistical methods such as analysis of variance, and parameters such as coefficients of variation (Harradine, 1949; Ball and Williams, 1968). However, this assumption cannot simply be made before the correlation of the samples with distance is proven to be nonexistent.

One of the oldest methods of estimating space or time dependency between neighboring observations is through autocorrelation. This method, which has its origins in time-series analyses, has been used intensively in soil science (Webster, 1973; Webster and Cuanalo, 1975; Vieira et al., 1981). Because it is a measure of the dependency between neighboring samples, it has several important applications such as to the location of soil boundaries (Webster, 1973) and to the design of sampling schemes from transect measurements (Vieira et al., 1981). When the observations are distributed two-dimensionally across a field surface, two-dimensional autocovariance functions are used to ascertain the spatial dependency. However, when interpolation between measurements is needed, a more adequate tool is used to measure the correlation between measurements. This is the semivariogram, which is defined as

$$\gamma(h) = \frac{1}{2} E\{[Z(x_i) - Z(x_i + h)]^2\} \qquad [1]$$

which, in turn, can be estimated by

$$\gamma^*(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [z(i) - z(i + h)]^2 \qquad [2]$$

in which E is the expected value and N(h) is the number of pairs of observations [z(i), z(i + h)] separated by a distance or lag vector h (Journel and

Huijbregts, 1978, pg. 12). A plot of $\gamma^*(h)$ versus the corresponding values of h, called a semivariogram, is thus a function of a vector h, which may depend on both the magnitude and the direction of h. In the latter case, when the semivariograms for different directions are different, the physical phenomenon under study is anisotropic, and anisotropic semivariograms must be submitted to some transformation that will make them isotropic. This subject is well described in the literature (Journel and Huijbregts, 1978, pg. 175; Burgess and Webster, 1980).

The graph of the experimental semivariogram computed using equation [2] will display a series of discrete points corresponding to each value of h, to which a continuous function $z^*(h)$ must be fitted. Delhomme (1976) discusses several theoretical models for this function that are applicable to different phenomena. Only the ones with particular importance to this paper will be discussed here.

We should expect that the differences $[z(i) - z(i + h)]$ decrease as h, the magnitude of the separation between them, decreases. Observations located close together are expected to be more alike than observations separated by large distances. For second-order stationary processes, $\gamma(0) = 0$, as shown in equation [2]. However, in practice (Delhomme, 1976; Campbell, 1978; Hajrasuliha et al., 1980; Burgess and Webster, 1980a), as h approaches 0 (zero), $\gamma^*(h)$ approaches a positive value called the nugget effect C_o. This value reveals the discontinuity of the semivariogram near the origin at distances less than the shortest sampling distance. The discontinuity may be due to variability at scales smaller than the sampling distance or to measurement errors (Delhomme, 1976).

Fitting a theoretical model to the experimental semivariogram is one of the important aspects of the applications of the theory of regionalized variables and may be one of the major sources of ambiguity in these applications, since all the calculations depend on the value of the semivariogram for specified distances (Vieira et al., 1981). Only the most commonly used models will be presented in this article; other models may be more appropriate for particular situations. Automatic curve fitting is not recommended, at least in mining geostatistics (Journel and Huijbregts, 1978), probably because these methods (e.g., least squares) do not allow for the flexibility of giving more weight to semivariogram values that result from more pairs of experimental values. However, if the theoretical model is a conditional positive definite function (Journel and Huijbregts, 1978), no restriction should be imposed on the choice of method used to obtain the model.

Depending on the behavior of the semivariogram at large values of h, the model to be used can be classified into two distinct categories: models with a sill and those without a sill.

1. Models with a Sill or Transition Models – As h increases from zero, $\gamma^*(h)$ increases up to a certain corresponding value of h, say, a, after which

it remains basically constant. The value of $\gamma^*(h)$ at this point is approximately equal to the variance of the observations and is called the sill, C, and the distance, a, is called the range (see Fig. 1). The range is a very important parameter, since measurements separated by distances closer than a are correlated to each other, whereas those measurements separated by distances greater than a are not correlated. When the range, a, is smaller than the closest sampling distance, we have a pure nugget effect and the physical phenomenon has a completely random spatial distribution with respect to the sampling space available; classical statistical methods may be applied. However, if more samples are taken at closer spacing, the semivariogram may reveal some structure. In fact, the existence of a pure nugget effect is the only situation that theoretically allows for the use of classical statistical methods.

Basically, four transition theoretical models are in use: (a) linear; (b) spherical; (c) exponential; and (d) Gaussian.

In the following expressions, C_o is the nugget effect, $C_o + C_1$ is the sill, and a is the range of the semivariogram.

(a) Linear

$$\gamma^*(h) = C_o + B \cdot h \qquad 0 \leqq h \leqq a$$

$$\gamma^*(h) = C_o + C_1 \qquad h \geqq a \qquad [3]$$

in which B is the slope for $0 \leqq h \leqq a$.

(b) Spherical

$$\gamma^*(h) = C_o + C_1 \left[\frac{3}{2}\frac{h}{a} - \frac{1}{2}\left(\frac{h}{a}\right)^3\right] \qquad 0 \leqq h \leqq a$$

$$\gamma^*(h) = C_o + C_1 \qquad h \geqq a \qquad [4]$$

The spherical model is obtained by first selecting the nugget effect, C_o, and the sill value, $C_o + C_1$. Then a line intercepting the y-axis at C_o and tangent to the points near the origin will reach the sill at a distance $a' = \frac{2}{3}a$. Thus the range is $a = \frac{3a'}{2}$. The spherical model behaves linearly up to approximately $\frac{1}{3}a$.

(c) Exponential

$$\gamma^*(h) = C_o + C_1[1 - \exp(-h/a_o)] \qquad 0 \leqq h \leqq d \qquad [5]$$

in which d is the maximum distance over which the semivariogram is defined. The parameter, a_o, is obtained by taking a tangent to the points near the origin, intercepting the y-axis at C_o. The distance, a_o, at which the tangent

line reaches the sill is approximately $a_o = a/3$, where a is the range. Unlike the spherical and linear models, the exponential reaches the range only asymptotically.

(d) Gaussian

$$\gamma^*(h) = C_o + C_1[1 - \exp(-h^2/a_o^2)] \qquad [6]$$

The parameter, a_o, is related to the range by $a_o = \frac{1}{\sqrt{3}} a$, in which the range, a, is obtained visually as the distance after which the experimental variogram becomes stable.

2. Models Without a Sill – Models without a sill correspond to phenomena with an infinite capacity of dispersion, for which neither the variances of the data nor the covariances can be defined (Journel and Huijbregts, 1978). In general, these models can be written as

$$\gamma(h) = C_o + Ah^B \qquad 0 < B < 2 \qquad [7]$$

Parameter B must be strictly greater than 0 (zero) and strictly less than 2 in order to guarantee that the function $-\gamma(h)$ is conditional positive definite.

Some phenomena may have a semivariogram that shows nested structures, or more than one structure. In this case, a combination of models may be needed instead of a single one.

Stationarity Assumptions

With a single sampling, all we know about our random function $Z(x_i)$ is one realization. If we want to estimate values for the unrecorded locations, we must introduce the restriction that the regionalized variable must be statistically homogeneous and isotropic, which permits us to make statistical inference (Olea, 1975). Formally, a regionalized variable is stationary if the statistics on the random variables $Z(x_i+h)$ are the same for every vector h. According to the number k of statistical moments that are constants, the variable is called stationary of order k. Second-order stationarity is all that is usually required in geostatistics (Olea, 1975).

Suppose random function $Z(x_i)$ has expected values $E\{Z(x_i)\} = m(x_i)$ and $E\{Z(x_i + h)\} = m(x_i + h)$ and variances var $\{Z(x_i)\}$ and var $\{Z(x_i + h)\}$, respectively, for the locations x_i and $x_i + h$. The covariance $C(x_i, x_i + h)$ between $Z(x_i)$ and $Z(x_i + h)$ is defined by

$$C(x_i, x_i + h) = E\{Z(x_i)\, Z(x_i + h)\} - \{m(x_i)\, m(x_i + h)\} \qquad [8]$$

and the variogram $2\gamma(x_i, x_i + h)$ is defined by

$$2\gamma(x_i, x_i + h) = E\{[Z(x_i) - Z(x_i + h)]^2\} \quad [9]$$

The variance of $Z(x_i)$ is

$$\text{var}\{Z(x_i)\} = E\{Z(x_i)\, Z(x_i + 0) - m(x_i)\, m(x_i + 0)\}$$

$$= E\{Z^2(x_i) - m^2(x_i)\} = C(x_i, x_i) \quad [10]$$

and the variance of $Z(x_i + h)$ is

$$\text{var}\{Z(x_i + h)\} = E\{Z^2(x_i + h) - m^2(x_i + h)\} = C(x_i + h, x_i + h) \quad [11]$$

Then three possible assumptions concerning stationarity of the random function $Z(x_i)$ can be made, and at least one assumption must be met in order to apply geostatistical estimation. These assumptions are as follows.

1. Stationarity of Order 2 – A random function $Z(x_i)$ is stationary of order 2 (Journel and Huijbregts, 1978) when: i) the expected value $E\{Z(x_i)\}$ exists and does not depend on the position x. Mathematically,

$$E\{Z(x_i)\} = m \text{ for all } x_i \text{ inside area S} \quad [12]$$

ii) for each pair of random variables, $\{Z(x_i), Z(x_i + h)\}$, the covariance function, C(h), exists and depends on h.

$$C(h) = E\{Z(x_i)\, Z(x_i + h)\} - m^2 \text{ for all } x_i \text{ inside area S} \quad [13]$$

Equation [13], the stationarity of the covariance, implies the stationarity of the variance and the variogram. Thus, using the linearity of the expected value, E, in equation [10] results in

$$\text{var}\{Z(x_i)\} = E\{Z(x_i)\, Z(x_i + 0)\} - E\{m^2(x_i)\}$$

$$= E\{Z(x_i)\, Z(x_i + 0)\} - m^2(x_i)\} \quad [14]$$

Apply the stationarity conditions [12] and [13] to obtain

$$\text{var}\{Z(x_i)\} = E\{Z^2(x_i)\} - m^2 = C(0) \quad [15]$$

The variogram $2\gamma(x_i, x_i + h)$ in equation [9] can be developed into

$$2\gamma(x_i, x_i + h) = 2\gamma(h) = E\{Z^2(x_i) - 2Z(x_i)\, Z(x_i + h) + Z^2(x_i + h)\} \quad [16]$$

Adding and subtracting $2m^2$ to equation [16] gives

$$2\gamma(h) = E\{Z^2(x_i) - m^2 - 2Z(x_i)\,Z(x_i + h) + 2m^2 + Z^2(x_i + h) - m^2\} \quad [17]$$

Since the expected value, E, is a linear operator and the expected value of a constant is the constant itself, then

$$2\gamma(h) = E\{Z^2(x_i)\} - m^2 - 2[E\{Z(x_i)\,Z(x_i + h)\} - m^2] + E\{Z^2(x_i + h)\} - m^2 \quad [18]$$

Substituting equations [13] and [15] into equation [18] gives

$$2\gamma(h) = C(0) - 2C(h) + C(0) = 2C(0) - 2C(h) \quad [19]$$

or, simplified,

$$\gamma(h) = C(0) - C(h) \quad [20]$$

Therefore, if the assumption of stationarity of order 2 can be made, the covariance C(h) and the variogram $2\gamma(h)$ are two equivalent tools for characterizing the autocorrelation between two variables $Z(x_i)$ separated by a distance h. By assuming stationarity, we can essentially repeat an experiment even though samples must be collected at different points as all samples are assumed to be drawn from populations having the same moments (Olea, 1975).

However, the assumption of stationarity of order 2 implies the existence of a finite variance of the measured values, $\text{var}\{Z(x)\} = C(0)$. This assumption may not be satisfied for some physical phenomena having an infinite capacity of dispersion. Examples include gold values in South Africa gold mines (Krige, 1951), Brownian motion (Journel and Huijbregts, 1978) and some multiplicative Markov chains (Bartlett, 1966). In this situation, a weaker assumption, the intrinsic assumption could be applicable.

2. Intrinsic Assumption – To avoid restricting the existence of a finite variance, required by the assumption of stationarity of order 2, the intrinsic assumption is made, which requires the existence and stationarity of the variogram only. A random function Z(x) is intrinsic when, in addition to the condition in equation [12], the increment $[Z(x_i) - Z(x_i + h)]$ has a finite variance that does not depend on x_i for all vectors h. This is mathematically written as

$$\text{var}\{[Z(x_i) - Z(x_i + h)] = E[Z(x_i) - Z(x_i + h)]^2\}$$

$$\text{for all } x_i \text{ within the areas} \quad [21]$$

substituting equation [2] gives

$$2\gamma(h) = E\{[Z(x_i) - Z(x_i + h)]^2\} \quad [22]$$

The function $\gamma(h)$ is the semivariogram. The reason for the prefix "semi" is that equation [22] can be written as

$$\gamma(h) = \frac{1}{2} E\{[Z(x_i) - Z(x_i + h)]^2\} \quad [23]$$

The factor 2 was simply introduced in the definition of the variogram, $2\gamma(h)$, because the quantity that is used most is $\gamma(h)$ rather than $2\gamma(h)$.

3. The Hypothesis of Universal Kriging – Under this hypothesis the random function $Z(x_i)$ for every location, x_i, consists of two components.

$$Z(x_i) = m(x_i) + \varepsilon(x_i) \quad [24]$$

in which $m(x_i)$ is the drift and $\varepsilon(x_i)$ is the residual error. Therefore, for each position x_i, we need to determine the drift $m(x_i)$ and have an expression for the semivariogram of the residuals from that drift (Webster and Burgess, 1980).

Because of the arbitrary nature involved in the expression for the drift and the amount of trial and error involved in the technique to calculate the two components, $m(x_i)$ and $\varepsilon(x_i)$, we will refrain from presenting any further theoretical development on universal kriging. An excellent work is reported in this respect by Olea (1975, 1977), and a good example of the application of universal kriging is found in Webster and Burgess (1980).

If a random function is stationary of order k ($k > 0$), then it is also stationary of all orders smaller than k. Consequently, if a random function $Z(x_i)$ is stationary of order 2, then it is also intrinsic. However, the converse is not necessarily true.

Kriging

Suppose that we want to estimate values, z^*, for all the locations, x_0, where values have not been measured, and that the estimation is to be a linear combination of measured values. Thus,

$$z^*(x_0) = \sum_{i=1}^{N} \lambda_i \, z(x_i) \quad [25]$$

in which N is the number of measured values $z(x_i)$ involved in the estimation and λ_i are the weights attached to each measured value, $z(x_i)$. By taking $z(x_i)$ as a realization of the random function $Z(x_i)$ and assuming stationarity of order 2, the estimator becomes

$$Z^*(x_0) = \sum_{i=1}^{N} \lambda_i \, Z(x_i) \quad [26]$$

Therefore, we must determine the weights, λ_i, before our problem can be solved. There are numerous ways to distribute the weights, λ_i, and examples include the inverse of the square of the distances, the inverse of the distances, and the inverse of the number of values. The best estimator must be unbiased and must have minimum variance. This is mathematically written as

$$E\{Z^*(x_0) - Z(x_0)\} = 0 \qquad [27]$$

and

$$E\{[Z^*(x_0) - Z(x_0)]^2\} = \text{minimum} \qquad [28]$$

which are the conditions of unbiasedness and of minimum variance of estimation, respectively.

Substituting equation [26] into [27] gives

$$E\{Z^*(x_0) - Z(x_0)\} = E\left\{\sum_{i=1}^{N} \lambda_i Z(x_i) - Z(x_0)\right\} = 0 \qquad [29]$$

Applying linearity of the E operator gives

$$E\{Z^*(x_0) - Z(x_0)\} = \sum_{i=1}^{N} \lambda_i E\{Z(x_i)\} - E\{Z(x_0)\} = 0 \qquad [30]$$

Substituting equation [12] into [30] and factoring like terms gives

$$E\{Z^*(x_0) - Z(x_0)\} = m \sum_{i=1}^{N} \lambda_i - 1 = 0 \qquad [31]$$

Therefore, the estimation will be unbiased if

$$\sum_{i=1}^{N} \lambda_i = 1 \qquad [32]$$

Developing equation [28] gives

$$E\{[Z^*(x_0) - Z(x_0)]^2\} = E\{[Z^{*2}(x_0) + Z^2(x_0) - 2Z^*(x_0)\, Z(x_0)]\} \qquad [33]$$

By linearity of the E operator,

$$E\{[Z^*(x_0) - Z(x_0)]^2\} = E\{Z^{*2}(x_0)\} + E\{Z^2(x_0)\} - 2E\{Z^*(x_0)\, Z(x_0)\} \qquad [34]$$

By developing each term on the right-hand side of equation [34] individually and using equation [26] successively, we get

$$E\{Z^{*2}(x_0)\} = E\{[\sum_i \lambda_i Z(x_i)]^2\} = E\{\sum_i \lambda_i Z(x_i) \sum_j \lambda_j Z(x_j)\} \quad [35]$$

$$= E\{\sum_i \lambda_i \sum_j \lambda_j Z(x_i) Z(x_j)\} = \sum_i \sum_j \lambda_i \lambda_j E\{Z(x_i) Z(x_j)\} \quad [36]$$

Substituting equation [13] into equation [36] results in

$$E\{Z^{*2}(x_0)\} = \sum_{i=1}^{N} \sum_{j=1}^{N} \lambda_i \lambda_j C(x_i, x_j) + m^2 \quad [37]$$

in which $C(x_i, x_j)$ refers to the covariance function corresponding to a vector with origin at x_i and extremity at x_j. The second term on the right-hand side of equation [34] is

$$E\{Z^2(x_0)\} = E\{Z(x_0) Z(x_0 + 0)\} \quad [38]$$

Substituting equation [13] into equation [38] results in

$$E\{Z^2(x_0)\} = C(0) + m^2 \quad [39]$$

The third term on the right-hand side of equation [34] is

$$E\{Z^*(x_0)Z(x_0)\} = E\{\sum_i \lambda_i Z(x_i) Z(x_0)\} = \sum_i \lambda_i E\{Z(x_i) Z(x_0)\} \quad [40]$$

Substituting equation [13] into equation [40] results in

$$E\{Z^*(x_0) Z(x_0)\} = \sum_{i=1}^{N} \lambda_i C(x_i, x_0) + m^2 \quad [41]$$

in which $C(x_i, x_0)$ refers to the covariance function corresponding to a vector with origin at x_i and extremity at x_0. Substituting [37], [39], and [41] into equation [34] results in

$$E\{[Z^*(x_0) - Z(x_0)]^2\} = \sum_i \sum_j \lambda_i \lambda_j C(x_i, x_j) + m^2 + C(0) + m^2$$

$$- 2 \sum_i \lambda_i C(x_i, x_0) - 2m^2 \quad [42]$$

or, simplified,

$$E\{[Z^*(x_0) - Z(x_0)]^2\} = \sum_i \sum_j \lambda_i \lambda_j C(x_i, x_j) + C(0) - 2 \sum_i \lambda_i C(x_i, x_0) \quad [43]$$

Equation [43] is, therefore, to be minimized under the constraint that $\sum_i \lambda_i = 1$ This minimization process can be done using Lagrangian multiplier techniques thoroughly described in advanced calculus textbooks. In order to satisfy equation [28], the N partial derivatives,

$$\partial[E\{[Z^*(x_0) - Z(x_0)]^2\} - 2 \mu \sum_i \lambda_i]/\partial\lambda_i \quad [44]$$

are set equal to 0 (zero); μ is a Lagrangian multiplier. Thus,

$$2 \sum_j \lambda_j C(x_i, x_j) - 2C(x_i, x_0) - 2\mu = 0 \quad [45]$$

By canceling the factor 2 in equation [45], rearranging, and combining with equation [32], we have the kriging system:

$$\sum_{j=1}^{N} \lambda_j C(x_i, x_j) - \mu = C(x_i, x_0), \quad i=1 \text{ to } N$$

$$\sum_{j=1}^{N} \lambda_j = 1 \quad [46]$$

The first N equations in system [46] can be rearranged into

$$\sum_{j=1}^{N} \lambda_j C(x_i, x_j) = \mu + C(x_i, x_0) \quad [47]$$

By substituting equation [47] into equation [43], we have the minimum estimation variance, $\sigma_k^2 (x_0)$:

$$\sigma_k^2 (x_0) = E\{[Z^*(x_0) - Z(x_0)]\} = \sum_i \lambda_i[\mu + C(x_i, x_0)] + C(0)$$

$$- 2 \sum_i \lambda_i C(x_i, x_0) \quad [48]$$

Rearranging and canceling like terms yields

$$\sigma_k^2 (x_0) = C(0) + \mu - \sum_{i=1}^{N} \lambda_i C(x_i, x_0) \quad [49]$$

The kriging system [46] can be written in matrix notation as

$$[C][\lambda] = [b] \qquad [50]$$

whose solution is of the form

$$[\lambda] = [C]^{-1}[b] \qquad [51]$$

in which [C] is the covariance matrix, or the kriging matrix in terms of covariance, $[C]^{-1}$ is the inverse of [C], $[\lambda]$ is the matrix of unknown weighting factors λ_i and [b] is the right-hand side of equation [46].

In matrix notation, equation [49] becomes

$$\sigma_k^2 (x_0) = C(0) - [\lambda]^t[b] \qquad [52]$$

in which $[\lambda]^t$ is the transpose of the matrix of lambdas.

Suppose that N = 4. The matrix [C] is then a 5 x 5 matrix and can be explicitly written as

$$[C] = \begin{bmatrix} C(x_1, x_1) & C(x_1, x_2) & C(x_1, x_3) & C(x_1, x_4) & 1 \\ C(x_2, x_1) & C(x_2, x_2) & C(x_2, x_3) & C(x_2, x_4) & 1 \\ C(x_3, x_1) & C(x_3, x_2) & C(x_3, x_3) & C(x_3, x_4) & 1 \\ C(x_4, x_1) & C(x_4, x_2) & C(x_4, x_3) & C(x_4, x_4) & 1 \\ 1 & 1 & 1 & 1 & 0 \end{bmatrix} \qquad [53]$$

The matrix of lambdas can be written as

$$[\lambda] = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \\ -\mu \end{bmatrix} \qquad [54]$$

The right-hand side of equation [46] can be written as

$$[b] = \begin{bmatrix} C(x_1, x_0) \\ C(x_2, x_0) \\ C(x_3, x_0) \\ C(x_4, x_0) \\ 1 \end{bmatrix} \qquad [55]$$

If stationarity of order 2 can be assumed, then the kriging system [46] can be written in terms of either covariance C(h) or semivariogram $\gamma(h)$, and shifting from one to the other is done using the relationship between $\gamma(h)$ and C(h), expressed in equation [20]. There is, however, one strong advantage in using covariance C(h), which is related to the numerical method used to obtain the solution of system [46]. Since the covariance function is a decreasing function, the main diagonal contains the largest values in the matrix. This is not true when a semivariogram $\gamma(h)$ is used, since all the elements of the main diagonal in this case are zeros (0). When the diagonal elements are the largest matrix coefficients, efficient Gaussian elimination procedures may be used.

However, if the intrinsic assumption expressed in equations [12] and [21] is applicable, we can assume then the only alternative is to use the semivariogram function $\gamma(h)$. The kriging system in terms of $\gamma(h)$ is, then, easily obtained by replacing C(h) in system [46] by C(0) - $\gamma(h)$. Thus

$$\sum_{j=1}^{N} \lambda_j \gamma(x_i, x_j) + \mu = \gamma(x_i, x_0), \quad i = 1 \text{ to } N$$

$$\sum_{j=1}^{N} \lambda_j = 1 \qquad [56]$$

and the estimation variance $\sigma_k^2(x_0)$

$$\sigma_k^2(x_0) = \mu + \sum_{i=1}^{N} \lambda_i \gamma(x_i, x_0) \qquad [57]$$

Cokriging

In soil science, agrometeorology, and remote sensing, very often some variables are cross-correlated with others. In addition, some of these variables are easier to measure than others. Examples include the saturated hydraulic conductivity and the percentage of silt and clay in soil science, where silt and clay content is easier to measure than saturated hydraulic conductivity, where surface temperature is easier to measure than surface water content and surface water content (0-5 cm) and surface temperature of a bare soil in agrometeorology. In such situations, estimation of one variable using information about both itself and another cross-correlated and easier-to-measure variable ought to be more useful than the kriging of that variable by itself. This estimation is easily done using cokriging.

Consider a field for which two variables, z_1 and z_2, have been measured with numbers of samples n_1 and n_2, respectively. Consider also for didactical purposes in this article only, that variable z_2 has been undersampled with respect to the variable z_1 ($n_1 \leq n_2$): say z_2 has been sampled on a 50-m square

grid whereas z_1 has been sampled on a 10-m square grid, and for all the locations where there is a sample of z_2, there is also one of z_1. Now let z_1 and z_2 be defined by $z_1(x_{1i})$, i = 1 to n_1, and $z_2(x_{2i})$, j = 1 to n_2, such that the arguments x_{1i} and x_{2j} correspond to x-y coordinates as $x_{1i} = [X_{1i}, Y_{1i}]$ and $x_{2j} = [X_{2j}, Y_{2j}]$, respectively. In addition, let $z_1(x_{1i})$ and $z_2(x_{2j})$ be particular realizations of the random functions $Z_1(x_{1i})$ and $Z_2(x_{2j})$, respectively.

With the above in mind and assuming stationarity, the first- and second-order moments of the random functions $Z_1(x_{1i})$ and $Z_2(x_{2j})$ are, respectively,

$$E\{Z_1(x_{1i})\} = m_1 \text{ for all } x_{1i} \text{ within the field} \quad [58]$$

$$E\{Z_2(x_{2j})\} = m_2 \text{ for all } x_{2j} \text{ within the field} \quad [59]$$

The covariance of $Z_1(x_{1i})$ is

$$C_{11}(h) = E\{Z_1(x_{1i} + h)\, Z_1(x_{1i})\} - m_1^2 \quad [60]$$

The cross-covariance between $Z_1(x_{1i})$ and $Z_2(x_{2j})$ i

$$C_{12}(h) = E\{Z_1(x_{2j} + h)\, Z_2(x_{2j})\} - m_1 m_2 \quad [61]$$

The cross-covariance between $Z_2(x_{2j})$ and $Z_1(x_{1i})$ is

$$C_{21}(h) = E\{Z_2(x_{2j} + h)\, Z_1(x_{2j})\} - m_2 m_1 \quad [62]$$

The covariance of $Z_2(x_{2j})$

$$C_{22}(h) = E\{Z_2(x_{2j} + h)\, Z_2(x_{2j})\} - m_2^2 \quad [63]$$

The semivariogram of $Z_1(x_{1i})$

$$\gamma_{11}(h) = \tfrac{1}{2} E\{[Z_1(x_{1i} + h) - Z_1(x_{1i})]^2\} \quad [64]$$

The cross-semivariogram between $Z_1(x_{1i})$ and $Z_2(x_{2j})$, which is also equal to the cross-semivariogram between $Z_2(x_{2j})$ and $Z_1(x_{1j})$, is

$$\gamma_{12}(h) = \tfrac{1}{2} E\{[Z_1(x_{2j} + h) - Z_1(x_{2j})][Z_2(x_{2j} + h) - Z_2(x_{2j})]\} = \gamma_{21}(h) \quad [65]$$

The semivariogram of $Z_2(x_{2j})$ is

$$\gamma_{22}(h) = \tfrac{1}{2} E\{[Z_2(x_{2j} + h) - Z_2(x_{2j})]^2\} \quad [66]$$

The estimator $Z_2^*(x_0)$ is

$$Z_2^*(x_0) = \sum_{i=1}^{N_1} \lambda_{1i} Z_1(x_{1i}) + \sum_{j=1}^{N_2} \lambda_{2j} Z_2(x_{2j}) \quad [67]$$

in which N_1 and N_2 are the number of neighbors of z_1 and z_2, respectively, used in the estimation of one cokriged value, $Z_2^*(x_0)$.

The unbiasedness condition is

$$E\{Z_2^*(x_0) - Z_2(x_0)\} = 0 \quad [68]$$

and the minimum variance condition is

$$\sigma_e^2(x_0) = E\{[Z_2^*(x_0) - Z_2(x_0)]^2\} = \text{minimum} \quad [69]$$

Substituting equation [67] into equation [68] gives

$$E\{Z_2^*(x_0) - Z_2(x_0)\} = E\left\{\sum_{i=1}^{N_1} \lambda_{1i} Z_{1(x_{1i})} + \sum_{j=1}^{N_2} \lambda_{2j} Z_2(x_{2j}) - Z_2(x_0)\right\} = 0 \quad [70]$$

Using the linearity of the E operator gives

$$E\{Z_2^*(x_0) - Z_2(x_0)\} = \sum_i \lambda_{1i} E\{Z_1(x_{1i})\} + \sum_j \lambda_{2j} E\{Z_2(x_{2j})\} - E\{Z_2(x_0)\} = 0 \quad [71]$$

Substituting [58] and [59] into [71] gives

$$m_1 \sum_i \lambda_{1i} + m_2 \sum_j \lambda_{2j} - m_2 = 0 \quad [72]$$

Factoring m_2 results in

$$m_1 \sum_i \lambda_{1i} + m_2\left[\sum_j \lambda_{2j} - 1\right] = 0 \quad [73]$$

Therefore, the estimator $Z_2^*(x_0)$ will be unbiased if

$$\sum_{j=1}^{N_2} \lambda_{2j} = 1 \quad [74]$$

and

$$\sum_{i=1}^{N_1} \lambda_{1i} = 0 \quad [75]$$

Now, by developing the square in equation [69] we get

$$E\{[Z_2^*(x_0) - Z_2(x_0)]^2\} = E\{Z_2^{*2}(x_0) + Z_2^2(x_0) - Z_2^*(x_0)\, Z_2(x_0) - Z_2(x_0)\, Z_2^*(x_0)\} \quad [76]$$

By applying linearity of the E operator we get

$$E\{[Z_2^*(x_0) - Z_2(x_0)]^2\} = E\{Z_2^{*2}(x_0)\} + E\{Z_2^2(x_0)\} - E\{Z_2^*(x_0)\, Z_2(x_0)\} - E\{Z_2(x_0)\, Z_2^*(x_0)\} \quad [77]$$

Developing each term on the right-hand side of equation [77] one by one and substituting equation [67] in the first term results in

$$E\{Z_2^{*2}(x_0)\} = E\{Z_2^*(x_0)\, Z_2^*(x_0)\} = E\{[\sum_i \lambda_{1i}\, Z_1(x_{1i}) + \sum_j \lambda_{2j}\, Z_2(x_{2j})]\, [\sum_k \lambda_{1k}\, Z_1(x_{1k}) + \sum_\ell \lambda_{2\ell}\, Z_2(x_{2\ell})]\} \quad [78]$$

Expanding inside the brackets on the right-hand side gives us

$$E\{Z_2^{*2}(x_0)\} = E\{\sum_i \lambda_{1i}\, Z_1(x_{1i}) \sum_k \lambda_{1k}\, Z_1(x_{1k}) + \sum_i \lambda_{1i}\, Z_1(x_{1i}) \sum_\ell \lambda_{2\ell}\, Z_2(x_{2\ell}) + \sum_j \lambda_{2j}\, Z_2(x_{2j}) \sum_k \lambda_{1k}\, Z_1(x_{1k}) + \sum_j \lambda_{2j}\, Z_2(x_{2j}) \sum_\ell \lambda_{2\ell}\, Z_2(x_{2\ell})\} \quad [79]$$

Rearranging the summations and using linearity of the E operator gives us

$$E\{Z_2^{*2}(x_0)\} = \sum_i \sum_k \lambda_{1i}\, \lambda_{1k}\, E\{Z_1(x_{1i})\, Z_1(x_{1k})\} + \sum_i \sum_\ell \lambda_{1i}\, \lambda_{2\ell}\, E\{Z_1(x_{1i})\, Z_2(x_{2\ell})\} + \sum_j \sum_k \lambda_{2j}\, \lambda_{1k}\, E\{Z_2(x_{2j})\, Z_1(x_{1k})\} + \sum_j \sum_\ell \lambda_{2j}\, \lambda_{2\ell}\, E\{Z_2(x_{2j})\, Z_2(x_{2\ell})\} \quad [80]$$

Substituting equations [60], [61], [62], and [63] into [80] gives us

$$E\{Z_2^{*2}(x_0)\} = \sum_i \sum_k \lambda_{1i}\, \lambda_{1k}\, [C_{11}(x_{1i}, x_{1k}) + m_1^2] + \sum_i \sum_\ell \lambda_{1i}\, \lambda_{2\ell}\, [C_{12}(x_{1i}, x_{2\ell}) + m_1 m_2] + \sum_j \sum_k \lambda_{2j}\, \lambda_{1k}\, [C_{21}(x_{2j}, x_{1k}) + m_2 m_1] + \sum_j \sum_\ell \lambda_{2j}\, \lambda_{2\ell}\, [C_{22}(x_{2j}, x_{2\ell}) + m_2^2] \quad [81]$$

Multiplying inside the brackets gives us

$$E\{Z_2^{*2}(x_0)\} = \sum_i \sum_k \lambda_{1i}\,\lambda_{1k}\,C_{11}(x_{1i}, x_{1k}) + m_1^2 \sum_i \sum_k \lambda_{1i}\,\lambda_{1k} +$$

$$\sum_i \sum_\ell \lambda_{1i}\lambda_{2\ell}C_{12}(x_{1i}, x_{2\ell}) + m_1m_2 \sum_i \sum_\ell \lambda_{1i}\,\lambda_{2\ell} + \sum_j \sum_k \lambda_{ij}\,\lambda_{1k}$$

$$C_{21}(x_{2j}, x_{1k}) + m_2m_1 \sum_j \sum_k \lambda_{2j}\,\lambda_{1k} + \sum_j \sum_\ell \lambda_{2j}\,\lambda_{2\ell}$$

$$C_{22}(x_{2j}, x_{2\ell}) + m_2^2 \sum_j \sum_\ell \lambda_{2j}\,\lambda_{2\ell} \qquad [82]$$

Substituting equations [74] and [75] into the constant terms in which m_1, m_2, or both appear in [82] gives us

$$E\{Z_2^{*2}(x_0)\} = \sum_i \sum_k \lambda_{1i}\,\lambda_{1k}\,C_{11}(x_{1i}, x_{1k}) + \sum_i \sum_\ell \lambda_{1i}\,\lambda_{2\ell}\,C_{12}(x_{1i}, x_{2\ell})$$

$$+ \sum_j \sum_k \lambda_{2j}\,\lambda_{1k}\,C_{21}(x_{2j}, x_{1k}) + \sum_j \sum_\ell \lambda_{2j}\,\lambda_{2\ell}\,C_{22}(x_{2j}, x_{2\ell}) + m_2^2 \qquad [83]$$

The second term on the right-hand side of equation [77] can now be expressed as

$$E\{Z_2^2(x_0)\} = E\{Z_2(x_0)\,Z_2(x_0)\} = E\{Z_2(x_0 + 0)\,Z_2(x_0)\} \qquad [84]$$

Substituting equation [63] into [84] for h = 0 gives us

$$E\{Z_2^2(x_0)\} = C_{22}(0) + m_2^2 \qquad [85]$$

Substituting equation [67] for the third term on the right-hand side of equation [77] results in

$$E\{Z_2^*(x_0)\,Z_2(x_0)\} = E\{[\sum_k \lambda_{1k}\,Z_1(x_{1k}) + \sum_\ell \lambda_{2\ell}\,Z_2(x_{2\ell})]\,Z_2(x_0)\} \qquad [86]$$

Multiplying inside the bracket and using linearity of the E operator results in

$$E\{Z_2^*(x_0)Z_2(x_0)\} = \sum_k \lambda_{1k}\,E\{Z_1(x_{1k})\,Z_2(x_0)\} + \sum_\ell \lambda_{2\ell}\,E\{Z_2(x_{2\ell})\,Z_2(x_0)\} \qquad [87]$$

Substituting equations [61] and [63] into [87] gives us

$$E\{Z_2^*(x_0)\,Z_2(x_0)\} = \sum_k \lambda_{1k}\,[C_{12}(x_{1k}, x_0) + m_1m_2] + \sum_\ell \lambda_{1\ell}\,[C_{22}(x_{2\ell}, x_0) + m_2^2] \qquad [88]$$

Multiplying inside the brackets gives us

$$E\{Z_2^*(x_0)\, Z_2(x_0)\} = \sum_k \lambda_{1k}\, C_{12}(x_{1k}, x_0) + m_1 m_2 \sum_k \lambda_{1k} + \sum_\ell \lambda_{2\ell}\, C_{22}(x_{2\ell}, x_0) + m_2^2 \sum_\ell \lambda_{2\ell} \quad [89]$$

By applying conditions [74] and [75] to the constant terms of equation [89], we get

$$E\{Z_2^*(x_0)\, Z_2(x_0)\} = \sum_k \lambda_{1k}\, C_{12}(x_{1k}, x_0) + \sum_\ell \lambda_{2\ell}\, C_{22}(x_{2\ell}, x_0) + m_2^2 \quad [90]$$

By substituting equation [67] for the fourth term on the right-hand side of equation [77], we get

$$E\{Z_2(x_0)\, Z_2^*(x_0)\} = E\{Z_2(x_0)[\sum_k \lambda_{1k}\, Z_1(x_{1k}) + \sum_\ell \lambda_{2\ell}\, Z_2(x_{2\ell})]\} \quad [91]$$

Multiplying inside the bracket and using linearity of the E operator yields

$$E\{Z_2(x_0)\, Z_2^*(x_0)\} = \sum_k \lambda_{1k}\, E\{Z_2(x_0)\, Z_1(x_{1k})\} + \sum_\ell \lambda_{2\ell}\, E\{Z_2(x_0)\, Z_2(x_{2\ell})\} \quad [92]$$

Substituting equations [62] and [63] into [92] yields

$$E\{Z_2(x_0)\, Z_2^*(x_0)\} = \sum_\ell \lambda_{1k}[C_{21}(x_0, x_{1k}) + m_2 m_1] + \sum_\ell \lambda_{2\ell}[C_{22}(x_0, x_{2\ell}) + m_2^2] \quad [93]$$

By multiplying inside the brackets, we get

$$E\{Z_2(x_0)\, Z_2^*(x_0)\} = \sum_k \lambda_{1k}\, C_{21}(x_0, x_{1k}) + m_2 m_1 \sum_k \lambda_{1k} + \sum_\ell \lambda_{1\ell}\, C_{22}(x_0, x_{2\ell}) + m_2^2 \sum_\ell \lambda_{2\ell} \quad [94]$$

Applying conditions [74] and [75] to the constant terms of equation [94] gives us

$$E\{Z_2(x_0)\, Z_2^*(x_0)\} = \sum_k \lambda_{1k}\, C_{21}(x_0, x_{1k}) + \sum_\ell \lambda_{2\ell}\, C_{22}(x_0, x_{2\ell}) + m_2^2\} \quad [95]$$

Substituting equations [83], [85], [90], and [95] into [77] gives us

$$\sigma_E^2(x_0) = E\{[Z_2^*(x_0) - Z_2(x_0)]^2\} = \sum_i \sum_k \lambda_{1i}\, \lambda_{1k}\, C_{11}(x_{1i}, x_{1k}) + \sum_i \sum_\ell \lambda_{1i}\, \lambda_{2\ell}\, C_{12}(x_{1i}, x_{2\ell}) + \sum_j \sum_k \lambda_{2j}\, \lambda_{1k}\, C_{21}(x_{2j}, x_{1k}) + \sum_j \sum_\ell \lambda_{2j}\, \lambda_{2\ell}\, C_{22}(x_{2j}, x_{2\ell}) + m_2^2 + C_{22}(0) + m_2^2$$

$$- \sum_k \lambda_{1k} C_{12}(x_{1k}, x_0) - \sum_\ell \lambda_{2\ell} C_{22}(x_{2\ell}, x_0) - m_2^2$$

$$- \sum_k \lambda_{1k} C_{21}(x_0, x_{1k}) - \sum_\ell \lambda_{2\ell} C_{22}(x_0, x_{2\ell}) - m_2^2 \qquad [96]$$

By canceling the m_2^2 and remembering that $C(h) = C(-h)$ and $C_{k'k}(h) = C_{kk'}(-h)$, equation [96] can be rewritten as

$$\sigma_E^2(x_0) = \sum_i \sum_k \lambda_{1i} \lambda_{1k} C_{11}(x_{1i}, x_{1k}) + \sum_i \sum_\ell \lambda_{1i} \lambda_{2\ell} C_{12}(x_{1i}, x_{2\ell})$$

$$+ \sum_j \sum_k \lambda_{2j} \lambda_{1k} C_{21}(x_{2j}, x_{1k}) + \sum_j \sum_\ell \lambda_{2j} \lambda_{2\ell} C_{22}(x_{2j}, x_{2\ell})$$

$$- 2 \sum_k \lambda_{1k} C_{12}(x_{1k}, x_0) - 2 \sum_\ell \lambda_{2\ell} C_{22}(x_{2\ell}, x_0) + C_{22}(0) \qquad [97]$$

Minimization of the estimation variance on equation [97] may be done by the application of Lagrangian multiplier techniques with

$$\partial[\sigma_E^2(x_0) - 2\mu_2 \sum_\ell \lambda_{2\ell}]/\partial\lambda_{2\ell} = 0 \qquad [98]$$

and

$$\partial[\sigma_E^2(x_0) - 2\mu_1 \sum_k \lambda_{1k}]/\partial\lambda_{1k} = 0 \qquad [99]$$

in which μ_1 and μ_2 are Lagrangian multipliers. Equation [98] yields

$$2 \sum_i \lambda_{1i} C_{11}(x_{1i}, x_{1k}) + \sum_\ell \lambda_{2\ell} C_{12}(x_{1i}, x_{2\ell}) + \sum_j \lambda_{2j} C_{21}(x_{2j}, x_{1k})$$

$$- 2 C_{12}(x_{1k}, x_0) - 2\mu_1 = 0 \qquad [100]$$

By remembering the fact that $C_{k'k}(h) = C_{kk'}(-h)$ and ℓ is a dummy index, we can rewrite equation [100] as

$$2 \sum_i \lambda_{1i} C_{11}(x_{1i}, x_{1k}) + 2 \sum_j \lambda_{2j} C_{12}(x_{1k}, x_{2j}) - 2C_{12}(x_{1k}, x_0) - 2\mu_1 = 0 \qquad [101]$$

Canceling and rearranging gives us

$$\sum_i \lambda_{1i} C_{11}(x_{1i}, x_{1k}) + \sum_j \lambda_{2j} C_{12}(x_{1k}, x_{2j}) - \mu_1 = C_{12}(x_{1k}, x_0) \qquad [102]$$

Equation [99] yields

$$\sum_i \lambda_{1i} C_{12}(x_{1i}, x_{2\ell}) + \sum_k \lambda_{1k} C_{21}(x_{2j}, x_{1k}) + 2 \sum_j \lambda_{2j} C_{22}(x_{2j}, x_{2\ell})$$

$$- 2 C_{22}(x_{2\ell}, x_0) - 2\mu_2 = 0 \qquad [103]$$

or

$$\sum_i \lambda_{1i}\, C_{12}(x_{1i}, x_{2\ell}) + \sum_j \lambda_{2j}\, C_{22}(x_{2j}, x_{2\ell}) - \mu_2 = C_{22}(x_{2\ell}, x_0) \qquad [104]$$

Combining equations [74], [75], [102], and [104] yields the cokriging system

$$\sum_{i=1}^{N_1} \lambda_{1i}\, C_{11}(x_{1i}, x_{1k}) + \sum_{j=1}^{N_2} \lambda_{2j}\, C_{12}(x_{1k}, x_{2j}) - \mu_1 = C_{12}(x_{1k}, x_0),$$

$$k = 1,\ldots,N_1$$

$$\sum_{i=1}^{N_1} \lambda_{1i}\, C_{12}(x_{1i}, x_{2\ell}) + \sum_{j=1}^{N_2} \lambda_{2j}\, C_{22}(x_{2j}, x_{2\ell}) - \mu_2 = C_{22}(x_{2\ell}, x_0),$$

$$\ell = 1,\ldots,N_2$$

$$\sum_{i=1}^{N_1} \lambda_{1i} = 0$$

$$\sum_{j=1}^{N_2} \lambda_{2j} = 1 \qquad [105]$$

Rearranging the first and second sets of equations in system [105] and substituting them into equation [97] yields the cokriging estimation variance, $\sigma^2_{k2}(x_0)$.

$$\sigma^2_{k2}(x_0) = C_{22}(0) + \mu_2 - \sum_{i=1}^{N_1} \lambda_{1i}\, C_{12}(x_{1i}, x_0) - \sum_{j=1}^{N_2} \lambda_{2j}\, C_{22}(x_{2j}, x_0) \qquad [106]$$

Solution of system [105] will yield N_1 weights λ_{1i}, N_2 weights λ_{2j}, and the Lagrangian multipliers, μ_1 and μ_2.

The cokriging system can be written in matrix notation as in equation [50]. Suppose that $N_2 = 2$ and $N_1 = 4$. Then, matrix [C] will be an 8 x 8 matrix and can be explicitly written as

$$[C] = \begin{bmatrix} C_{11}(x_{11},x_{11}) & C_{11}(x_{12},x_{11}) & C_{11}(x_{13},x_{11}) & C_{11}(x_{14},x_{11}) & C_{12}(x_{11},x_{21}) & C_{12}(x_{11},x_{22}) & 1 & 0 \\ C_{11}(x_{11},x_{12}) & C_{11}(x_{12},x_{12}) & C_{11}(x_{13},x_{12}) & C_{11}(x_{14},x_{12}) & C_{12}(x_{12},x_{21}) & C_{12}(x_{12},x_{22}) & 1 & 0 \\ C_{11}(x_{11},x_{13}) & C_{11}(x_{12},x_{13}) & C_{11}(x_{13},x_{13}) & C_{11}(x_{14},x_{13}) & C_{12}(x_{13},x_{21}) & C_{12}(x_{13},x_{22}) & 1 & 0 \\ C_{11}(x_{11},x_{14}) & C_{11}(x_{12},x_{14}) & C_{11}(x_{13},x_{14}) & C_{11}(x_{14},x_{14}) & C_{12}(x_{14},x_{21}) & C_{12}(x_{14},x_{22}) & 1 & 0 \\ C_{12}(x_{11},x_{21}) & C_{12}(x_{12},x_{21}) & C_{12}(x_{13},x_{21}) & C_{12}(x_{14},x_{21}) & C_{22}(x_{21},x_{21}) & C_{22}(x_{21},x_{22}) & 0 & 1 \\ C_{12}(x_{11},x_{22}) & C_{12}(x_{12},x_{22}) & C_{12}(x_{13},x_{22}) & C_{12}(x_{14},x_{22}) & C_{22}(x_{22},x_{21}) & C_{22}(x_{22},x_{22}) & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \end{bmatrix} \quad [107]$$

The matrix [λ] of unknowns can be written as

$$[\lambda] = \begin{bmatrix} \lambda_{11} \\ \lambda_{12} \\ \lambda_{13} \\ \lambda_{14} \\ \lambda_{21} \\ \lambda_{22} \\ -\mu_1 \\ -\mu_2 \end{bmatrix} \quad [108]$$

The right-hand side [b] can be written as

$$[b] = \begin{bmatrix} C_{12}(x_{11}, x_0) \\ C_{12}(x_{12}, x_0) \\ C_{12}(x_{13}, x_0) \\ C_{12}(x_{14}, x_0) \\ C_{22}(x_{21}, x_0) \\ C_{22}(x_{22}, x_0) \\ 0 \\ 1 \end{bmatrix} \qquad [109]$$

Unlike the kriging system, the cokriging system can be written simply in terms of cross-semivariogram $\gamma_{k'k}(h)$ only if $C_{k'k}(h) = C_{kk'}(h)$ (Journel and Huijbregts, 1978, p. 326). This can be verified in practice by calculating the two cross-covariograms $C_{12}(h)$ and $C_{21}(h)$ and determining whether or not they are equal. This is advisable, since the calculation of the cross-semivariogram does not require knowledge of the means of the two variables, whereas the cross-covariogram does require it. The decision as to how to calculate the mean depends on the distribution function, which describes each one of the variables and is an inconvenience, as it is one additional assumption to be made. When $C_{12}(h) = C_{21}(-h)$ the cokriging system can be written in terms of cross-semivariograms using the relation

$$\gamma_{k'k}(h) = C_{k'k}(0) - C_{k'k}(h) \qquad [110]$$

Notice that the cokriging system [105] was derived using $C_{12}(x_{1i}, x_{2j})$ using the relation $C_{12}(x_{1i}, x_{2j}) = C_{21}(x_{2j}, x_{1i})$. Therefore, it should not make any difference whether $C_{12}(x_{1i}, x_{2j})$ or $C_{21}(x_{2j}, x_{1i})$ is used. Notice also that matrix [C] (equation [107]) is symmetric.

Unique Neighborhood and Neighborhood of Estimation

When the size of the data set is reasonable relative to the computer facilities available, the entire data set may be used with the kriging system [46] constituting (N + 1) equations, in which N is the number of values in the data set. If matrix inversion is used on matrix [C] to obtain the lambdas [λ] (see equation [51]), then only one inversion is required for any number of kriging estimations, $Z^*(x_0)$, whatever x_0 within the field. This is true because matrix [C] will be invariant with x_0, and only the right-hand side [b] changes with x_0. However, semivariograms and covariograms (or cross semivario grams and cross covariograms) are not defined for distances $h > L$, in

which L is one-half the extension of the field. Therefore, unique neighborhood is not advised since exceeding the distance, L, is unavoidable when this technique is made. This is the first limitation of the estimation neighborhood.

Since the weights, λ_i, decrease as the distance between x_0 and x_i increases, points far from x_0 may be omitted without resulting in serious consequences (Burgess and Webster, 1980a). Journel and Huijbregts (1978) recommend a size of estimation neighborhood $A(x_0)$ that provides sufficient screening of the data exterior to $A(x_0)$. However, the meaning of "sufficient" remains subjective until the desired precision of estimation is defined. The precision of the estimation or the estimation variance and also the screen effect are inversely related to the nugget effect, C_0. Burgess and Webster (1980b) found that the nugget effect makes the largest contribution to the estimation variances even at points midway between sampling points. Vieira et al. (1981) used the decrease of the estimation variance with the increase of neighborhood of estimation $A(x_0)$ as a basis for selecting the ideal value of $A(x_0)$, estimating four arbitrarily selected points; the results of this technique may change with different choices of points estimated and may be meaningless without tests. Other suggestions can be found in the literature, such as the two nearest points in each octant around x_0 for irregularly spaced data (Olea, 1975) or the nearest 16 or 25 points for data sampled on a square grid (Burgess and Webster, 1980a).

In conclusion, it appears to us that there is not a definite rule applicable to all situations, and common sense added to a great deal of arbitrariness is usually involved in selecting the neighborhood of estimation. The analysis of the estimation error presented in the next section may be of use in supporting the value $A(x_0)$ selected. The discussion above is also applicable to cokriging, although some further and obvious complications can be expected owing to the introduction of a second variable.

Estimation Error

Both kriging and cokriging, like any other method of estimation, involve an error. This error is due to the fact that the variable to be estimated is generally somewhat different from the estimated value (Journel and Huijbregts, 1978). The estimation error, $e(x_k)$, is thus defined as the difference between the measured $z(x_k)$ and estimated $z^*(x_k)$ values for the same location x_k. Thus, if we have a set of n measured values, $z(x_i)$, we can obtain a set of n estimated values for identical locations, $z^*(x_i)$, and then calculate a set of n errors, $e(x_i) = [z^*(x_i) - z(x_i)]$ (Vieira et al., 1981) or n reduced errors, $r(x_i) = [z^*(x_i) - z(x_i)]/\sigma_k(x_i)$, in which $\sigma_k(x_i)$ is the standard deviation of estimation. It seems logical to prefer the reduced errors, $r(x_i)$, over the estimation errors, $e(x_i)$, because the first are dimensionless and, therefore, independent of the unit in which the measurements are expressed.

The reduced errors, $r(x_i)$, can be considered a particular realization of the random variable $R(x_i)$ for the point x_i. The set of all random variables $\{R(x_i) = [Z^*(x_i) - Z(x_i)]/S(x_i)\}$ is a random function, $R(x_i)$, in which $S(x_i)$ is a random variable corresponding to the standard deviation of estimation, $\sigma_k(x_i)$. We can now make statistical inference on $R(x_i)$ and calculate the expectation $m_R = E\{R(x_i)\}$ and variance $\sigma_R^2 = \text{var}\{R(x_i)\}$ as the two parameters for the distribution function of the reduced errors $R(x_i)$. $R(x_i)$ can be considered stationary when $Z(x_i)$ itself is stationary.

The quality of the estimation method can then be verified through two conditions of the reduced errors (Delhomme, 1976):

1. The mean error, m_R, must be close to zero.
2. Variance σ_R^2 must be equal to 1.

To calculate confidence intervals, some kind of two-parameter (m_R and σ_R^2) distribution function must be found to describe the distribution of the reduced errors. The normal distribution has been the one observed most often in practice (Journel and Huijbregts, 1978).

Under the assumption of normal distribution of the reduced errors, $R(x_i)$, a *t*-test can be performed on the mean error, m_R, to test the significance of the assumption $m_R = 0$. For the variance of the reduced errors, it is expected (Delhomme, 1976) to be equal to 1, and again a *t*-test can be performed. Further, 95% of the reduced errors are expected to be within ±2, which is equivalent to expecting the reduced errors to follow a normal distribution function with parameters $m_R = 0$ and $\sigma_R = 1$ (Snedecor and Cochran, 1967).

The requirement of closeness to normal distribution of the reduced errors with parameters $m_R = 0$ and $\sigma_R = 1$ makes sense, since any deviation from it would mean either systematic underestimation or overestimation.

Another option, less quantitative than the above, is to examine the plot of paired measured versus estimated values for the same location and measure the deviation of a regression line from a one-to-one relationship.

Some practical explanation as to how to calculate the above reduced errors is necessary owing to the importance of the effect of the errors on the evaluation of the quality of the estimation and the assumptions under which the estimation is made. The technique is commonly called cross-validation or jackknifing.

Suppose we have a set of n measured values, $z(x_i)$, distributed over a field of area S, and for which an experimental semivariogram, $\gamma^*(h)$, has been found to be adequately fitted by some model. Suppose further that for every position, x_i, we estimate a value $z^*(x_i)$ and the estimation variance $\sigma_E^2(x_i)$ using the n nearest neighbors of x_i but not x_i itself. In other words, for every position x_i, we will pretend that the value $z(x_i)$ has not been measured, and the remaining n-1 measured values still have the same variogram $\gamma^*(h)$. As a result, we will have the sets of n measured values $z(x_i)$, estimated values

$z^*(x_i)$, and estimation variances $\sigma_E^2(x_i)$, which make it possible to calculate the reduced errors, $r(x_i)$.

A similar approach can be used for cokriging calculating the reduced errors, $R_2(x_{2i}) = [Z_2^*(x_{2i}) - Z_2(x_{2i})]/\sigma_{k2}(x_{2i})$. Similar interpretations can be obtained, such as the quality of the cokriging estimation, the quality of the structural models used (cross-semivariograms), and validity of the assumptions made.

We shall discuss the main aspects later using field data as examples.

DESCRIPTION OF THE DATA SETS

The number and variety of data sets used as examples are purposely large to illustrate the different agronomic areas in which the theoretical concepts presented above may be useful. Each variable will be referred to by a short-hand notation for simplification of understanding and identification. Table 1 contains a list of all the variables with shorthand names in alphabetical order along with their corresponding statistical moments of interest in this paper. In the paragraph headings that follow, reference is given to the source of the corresponding data sets. The reader is encouraged to consult the sources of the data for details not mentioned here.

Remote Sensing--Vieira and Hatfield (1982)

During the winter growing seasons of 1976-77, 1977-78, and 1978-79, air temperature (TA77, TA78, TA79) and surface temperatures (TS77, TS78, TS79) were measured daily on the bare soil of a field at the University of California, Davis. Other variables measured are analyzed elsewhere (Vieira and Hatfield, 1982). Air temperature (TA) was measured using a standard meteorological thermometer, and surface temperature (TS) was measured using hand-held infrared thermometers. All measurements were made within 15 minutes of 1330 PST. The main objective of the experiment was to quantify the temporal variability of these measurements for remote sensing purposes and to use the autocorrelation for individual measurements with time and the cross-correlation between them to define frequency of sampling.

Variables TA and TS are one dimensional, where the coordinate is time in days, beginning in late November and ending in late April for 1976-77 and 1978-79 and in June for 1977-78. Climatologically, temperatures in 1976-77 and 1978-79 were close to the normal average and those in 1977-78 were higher than normal.

Soil Classification--R. B. Grossman (personal communication)

In an intensive and detailed soil survey at the Agronomy Research Center, Boone County, Missouri, several variables normally researched in soil surveys were measured from September through December 1977. The data used in this paper come from field H, blocks 3a, 3b, 4a, and 4b. The variables used were the depth (inches) to an estimated 40% clay (GRDEPTH) and the percentage of clay in the plow layer as estimated by finger examination (GRCLAY). The sampling grid was irregular. Basic distance was 35 ft: some samples were separated by only 25 ft, some were farther away, and at some points samples were missing. The measurements for clay percent were made on only some stations to provide evaluation for delineations, and no measurements of this kind were made on blocks 3a and 4b. Block 4a had more samples than block 3b

Table 1. List of variables with their respective statistical moments.

Variable name	Number of samples	Mean	Variance	Coefficient of variation	Skewness	Kurtosis
GRCLAY	74	23.97	5.810	0.1005	0.1704	2.331
GRDEPTH	484	13.91	5.430	0.1675	0.5428	4.368
MONIT	224	1.89	0.96×10^{-2}	0.057	0.4270	2.428
MOYLD	224	680.70	7729.000	0.1291	-0.1566	4.413
TA77	150	16.99	29.950	0.3220	-0.0385	2.515
TA78	224	19.03	37.640	0.3225	0.7092	3.001
TA79	142	15.89	24.450	0.3110	0.4838	2.971
TS77	150	26.28	89.290	0.3595	0.2149	2.170
TS78	224	25.36	195.300	0.5510	0.7529	2.612
TS79	142	19.11	121.700	0.5774	1.0890	3.466
WAAMSU1	81	5.03	2.220	0.2962	-0.4562	2.850
WAAMSU2	81	2.81	1.449	0.4292	0.5763	3.422
WABLOD1	81	20.93	32.980	0.2744	0.3637	3.508
WABLOD2	81	10.71	30.720	0.5173	0.6977	2.912
WACADA	100	1.12	0.019	0.1216	2.7570	18.090
WACAOK	100	0.43	0.013	0.2648	1.4310	6.567
WAIN281	81	3.54	1.196	0.3174	-0.3833	2.576
WAIN282	81	1.42	0.292	0.3810	1.3170	6.736
WANIDA	100	1.00×10^{-1}	8.87×10^{-5}	0.0937	0.0142	2.951
WANIOK	100	3.26×10^{-1}	5.27×10^{-5}	0.2225	1.689	7.173
WASOIL1	81	2.75	0.812	0.3279	0.900	5.587
WASOIL2	81	0.74	0.137	0.4977	1.414	5.344

over approximately the same areas. For these reasons, we used only block 4a for clay percent and the entire blocks 3 and 4 for depth to an estimated 40% clay.

Agronomy--Montgomery (1913)

The data used from the agronomy experiment constitute nitrogen percentage in the grain (MONIT) and yield of grain (MOYLD) from a field experiment involving turkey wheat. The author (Montgomery, 1913) was concerned with experimental error and variation and objectively suggested questions such as "Why should one plant, growing under practically the same environment as another, collect from the soil two or three times as much nitrogen?" or "Why should two plants yielding different quantities of grain collect the same quantities of nitrogen?" In an attempt to answer these questions, a field of 77 x 88 ft with a 5-ft margin outside was sown with turkey winter wheat. After a secured uniform stand and growth, the field was divided into 224 blocks of 5.5 x 5.5 ft each at harvest. The harvest of each block was recorded and a composite sample from each block was analyzed for total nitrogen.

Nitrification in Soil--Waynick (1918)

On October 20, 1917, soil samples were collected from a selected field on the University Farm at Davis (now the University of California, Davis) at locations shown in Fig. 2 (Waynick's Fig. 1). During the three preceding years, corn, Sudan grass, and grain sorghum had been grown (in the order named), and the field was kept bare during 1917 and free of any vegetation when the samples were taken. The soil surface was practically air-dry, since no rain had fallen since April, but the subsoil was quite moist. The area chosen was apparently uniform in texture and color level, and free of small depressions. Samples were taken from the surface (1-6 in) and subsoil (6-24 in). The surface samples were taken with a trowel, and an area approximately 6 in. in diameter was included in each sample. Subsoil samples were taken with a 3 in. auger. Samples were placed in sterile soil bags and shipped to the laboratory at the end of the sampling. After air-drying and sieving, four 100-g portions of each sample were weighed and placed in tumblers. One tumbler from each sample was reserved for the determination of the residual nitrate (WASOIL1 surface and WASOIL2 subsoil). To a second, no nitrogen compounds were added (WAIN281 and WAIN282); to a third 0.2 g of ammonium sulfate (WAAMSU1 and WAAMSU2) was added; and to the fourth, 1 g of dried blood (WABLOD1 and WABLOD2) was added. The last three tumblers were brought to an optimum moisture and incubated for 28 days at 28°C, after which the nitrate was determined colorimetrically.

Soil Nitrogen and Carbon--Waynick and Sharp (1919)

Figure 3 shows the diagram of the areas sampled showing the locations from which the samples were taken and the sample numbers (Waynick and Sharp's Fig. 1). Two fields in California were sampled: one on the University Farm at Davis from a silty clay loam soil and the other near the town of Oakley from a blow sand. The fields were selected for their apparent uniformity and were free of vegetation at the time of the sampling. Samples were taken with a 3-in. auger and analyzed in the laboratory for nitrogen (WANIDA-Davis and WANIOK-Oakley) and carbon (WACADA and WACAOK). Results are given as percent.

The coefficients of skewness and kurtosis are presented in Table 1 for comparison with a normal distribution, in which these coefficients are 0 and 3, respectively. We have no intention of finding the exact distribution function for the variables studied, nor do we present a test of significance between calculated and theoretical coefficients, which are beyond our objectives. The majority of the variables in Table 1 approximate a normal distribution as shown by the coefficients of skewness and kurtosis. Some (WACADA, WANIOK, and WACAOK) have a very high coefficient of kurtosis and, therefore, an excess of values near the mean and far from it, resembling the manner in which the t-distribution departs from the normal (Snedecor and Cochran, 1967, p. 86). Our intent is to call attention to the fact that most of the variables approximate a normal distribution, and yet all except three, have adequate structure and therefore are not independent. As was stated in the Introduction, normal distribution of observations is commonly confused with independence between them.

RESULTS AND DISCUSSION

Semivariograms

Semivariograms can be calculated easily using hand calculators or a computer program capable of computing Eq. [12]. The semivariogram programs GAMA1D and AVARIO, listed in Appendices A1 and A2, respectively, constitute a very simple algorithm for performing Eq. [12]. The program GAMA1D can be used for only one-dimensional variables, such as transects or time-variable data. The program AVARIO can be used for either one- or two-dimensional variables; when that program is used for two-dimensional variables, the average semivariogram for all directions is computed. Semivariograms for specific directions can be calculated using the program MAREC2 (David, 1977, p. 149) or program GAMA3 (Journel and Huijbregts, 1978, p. 224). The parameters for the models fitted to each semivariogram are shown in Table 2, where the variables are in alphabetical order.

Table 2. Parameters of semivariogram models. Gamma is the kind of model that best fit the experimental semivariogram.

Variable name	Gamma	C_0	C_1	Range	Sill
GRCLAY	spherical	1.80	6.40	550.0 ft	8.20
GRDEPTH	spherical	3.00	2.70	300.0 ft	5.70
MONIT	linear	8.8×10^{-3}	5.57×10^{-5}	21.0 ft	0.01
MOYLD	spherical	4000.00	4000.00	20.0 ft	8000.00
RTA77	spherical	0.00	12.00	8.0 d	12.00
RTA78	spherical	0.00	8.00	4.0 d	8.00
RTA79	exponential	0.00	8.50	9.0 d	8.50
RTS77	spherical	10.00	25.00	6.0 d	35.00
RTS78	spherical	20.00	24.00	3.0 d	44.00
RTS79	spherical	10.00	25.00	8.0 d	35.00
WAAMSU1	linear	1.80	0.012	33.0 ft	2.20
WAAMSU2	spherical	0.00	0.35	27.0 ft	0.35
WABLOD1	----------N U G G E T E F F E C T----------				
WABLOD2	spherical	11.00	20.00	27.0 ft	31.00
WACADA	----------N U G G E T E F F E C T----------				
WACAOK	spherical	0	9.5×10^{-3}	75.0 ft	9.5×10^{-3}
WAIN281	----------N U G G E T E F F E C T----------				
WAIN282	linear	0.125	5.8×10^{-3}	30.0 ft	0.30
WANIDA	linear	1.11×10^{-5}	5.4×10^{-7}	143.0 ft	8.9×10^{-5}
WANIOK	spherical	0	4.1×10^{-5}	90.0 ft	4.1×10^{-5}
WASOIL1	spherical	0	0.85	10.0 ft	0.85
WASOIL2	spherical	0.013	0.127	30.0 ft	0.14

1. One-Dimensional Problem

The air and surface temperatures (TA and TS) measured over the winter growing season are an example of a one-dimensional problem in which the coordinate is time in days. The semivariograms for TA and TS for the three years are plotted in Figs. 4a and 4b, respectively. One common point about these semivariograms is the linear increase but not leveling to a sill. The only one that shows an apparent sill is that for TA77 (Fig. 4a, unbroken line): the sill occurs around h = 30 and extends to about h = 60 days. However, another abrupt increase takes place after that. The experimental semivariogram should become level or fluctuate at the sill value up to at least one-half the length of the series of measurements; otherwise drift or nested structure could be present in the data. Now, one-half the length of TA77 is 75 days, which

suggests the presence of drift. Similar to the semivariogram estimation, the drift can be estimated by

$$D^*(h) = \frac{1}{N(h)} \sum_{i=1}^{N(h)} [z(i) - z(i+h)] \qquad [111]$$

The drifts for TA and TS for the three years are plotted in Figs. 5a and 5b, respectively. The negative sign of the drifts shows that all these measurements increase with time. This is not surprising since the measurements start in late November, generally with low temperatures in Davis, and end late in May with higher temperatures. This is commonly called seasonal trend, but we prefer to refer to it as drift to avoid ambiguity. The main reason drifts are undesirable in a data set like this is that the seasonal variation is quite well defined and can be inferred from net radiation measurements. However, the variation at short scale is not known and is important in designing sampling frequency for these measurements. The plot of the spectral density function, the Fourier transform of the autocorrelation function, showed that the periodicity in this data exists only at low frequency (Vieira and Hatfield, 1982). For this reason, we decided to use a moving average with a large number of points to follow the seasonal variation of these measurements, and a 29-point moving average appeared to be the shortest able to do that. A 29-point moving average avoided the shorter intervals which would be affected by frontal passage and daily weather. Once we had the moving average, we calculated the residuals by subtraction (RTA and RTS). The statistical moments of the RTA and RTS for the three years, shown in Table 3, again approach normal distribution. The high coefficients of variation come from the fact that the mean values are very low, since the residuals are monthly fluctuations from zero.

The semivariograms for RTA and RTS for the three years are shown in Figs. 6a and 6b, and the corresponding drifts are shown in Figs. 7a and 7b, respectively. Now, the semivariograms have a sill although with a pronounced fluctuation from the sill value, mainly for RTA77 and RTS78. The drifts shown in Figs. 7a and 7b are no longer as pronounced and steady an increase as that for TA and TS (see Figs. 5a and 5b). This is true mainly for 1978 for both temperatures, meaning that the 29-point moving average removed the drift more completely for TA78 and TS78 than for the other years.

Table 2 shows the parameters for the models fitted for these semivariograms. A spherical model describes all semivariograms except that for RTA79, which required an exponential model. The ranges of these variograms are an indication of the ideal frequency for sampling. Although more research is needed to support these results over more years, it seems that generally both air and surface temperatures could be sampled every two days and estimated for every day by kriging. This represents a decrease in the number of samples by one-half, which is encouraging from our point of view.

2. Two-Dimensional Problem

Except for the air and surface temperatures, all variables used in this paper are in two dimensions: the measurements come from an area in a given field. That means that for each data point, x and y coordinates are attached, and magnitudes depend on an arbitrarily fixed point.

The semivariograms for GRDEPTH and GRCLAY are shown in Figs. 8a and 8b, respectively. A spherical model fits almost exactly to the one for GRDEPTH (Fig. 8a), with a nugget effect of 3.0, a range of 300 ft, and a sill of 5.7. The semivariogram for GRCLAY shows a sill that is higher than the variance of the observations. For this reason, it may be safer to restrict the approach to the intrinsic hypothesis, since a finite variance may not exist.

Figures 9a and 9b show the semivariograms for MONIT and MOYLD, respectively. The nitrogen in the grains of turkey wheat have a weak and unclear structure (Fig. 9a), whereas the one for the yield is much better described with a spherical model. A tentative linear model was fit to MONIT, and the parameters can be found in Table 3.

The semivariograms for WASOIL2 and WABLOD2 shown in Figs. 10a and 10b, respectively, have a low nugget effect and clearly defined sill, with ranges of approximately 30 feet; spherical models describe them reasonably well. Some periodicity may exist in the data, as shown in Figs. 10a and 10b by a peak in the semivariograms at about 20 feet.

Nitrogen and carbon measured on identical grids in Davis and Oakley have very different semivariograms, as shown in Figs. 11a, 11b, 12a and 12b, respectively for WANIDA, WACADA, WANIOK, and WACAOK. In Fig. 11a, the semivariogram for WANIDA shows a linear increase with distance without a sill, resembling the semivariograms for surface temperature shown in Fig. 4b, in which a strong drift was found, as shown in Fig. 5b. However, it is more difficult to deter-

Table 3. Residuals of air temperature (RTA) and surface temperature (RTS) and their statistical moments.

Variable name	Number of samples	Mean	Variance	Coefficient of variation	Skewness	Kurtosis
RTA77	122	0.0457	11.47	74.14	0.1820	2.252
RTA78	196	0.1253	8.65	23.47	0.3457	2.710
RTA79	114	0.0582	8.38	49.35	0.0726	3.062
RTS77	122	0.2264	31.77	24.90	0.3867	3.696
RTS78	196	0.2313	43.50	28.51	0.5324	3.708
RTS79	114	0.1379	32.51	41.34	-0.0015	3.889

mine the significance of a drift in two dimensions than in one dimension, mainly because the calculation of an average drift for all directions is meaningless. Examination of the drift calculated for a fixed direction is useful for only that direction. We calculated semivariograms and drifts for the four main directions: zero, 90, 45 and -45 degrees, were zero degrees is parallel to the x-axis. These calculations revealed isotropy up to approximately 50 feet and no drift for any of the four directions calculated. Therefore, we can conclude that those measurements of nitrogen in Davis do not have a finite variance and are not second-order stationary. In Fig. 11b, the semivariogram for WACADA shows an unclear structure, which we consider a pure nugget effect. The semivariograms for WANIOK and WACAOK shown in Figs. 12a and 12b, respectively, have a sill that is smaller than the variance. Spherical models were fitted to them, and the parameters are shown in Table 2.

The next six semivariograms are examples of difficult situations in which it is not clear whether a structure exists. Figs. 13a, 13b, and 13c show the semivariograms for WASOIL1, WABLOD1, and WAIN281 respectively, and Figs. 14a, 14b and 14c show those for WAAMSU1, WAAMSU2, and WAIN282, respectively. The semivariograms in Figs. 13b and 13c are probably pure nugget effect. The one in Fig. 14c is not clear enough to allow for any assumptions. In short, the only one that has a more or less clear spherical structure is the one for WAAMSU2, shown in Fig. 14b.

All of the above two-dimensional semivariograms have been calculated using the computer program AVARIO listed in the Appendix.

Cross-Semivariograms

All the cross-semivariograms have been calculated using the computer program XGAMA listed in the Appendix, without distinguishing between one- or two-dimensional problems, since calculation of a drift when mixing two variables is meaningless.

1. One-Dimensional

The cross-semivariogram calculated for the raw values of TA and TS for the three years 1977-79 are shown in Fig. 15, and the corresponding ones for the daily residuals from a 29-point moving average are shown in Fig. 16. The cross-variograms in Fig. 15 are apparently linear without manifesting a sill. For all practical purposes, the cross-semivariograms for the residuals shown in Fig. 16 for the three years should fit one spherical model. This kind of result we consider especially desirable for future research.

2. Two-Dimensional

A linear model describes the cross-semivariogram for GRCLAY vs. GRDEPTH, shown in Fig. 17, although a very low correlation coefficient was found between these variables. Figure 18 shows the nitrogen content of the turkey wheat

grains to be negatively correlated to yield, although the negative correlation was low, and the cross-semivariogram was not very well defined. For variables having an unclear cross-semivariogram such as these, not much gain can be expected using cokriging instead of kriging. This will be illustrated by the contour maps of measured-plus-kriged and measured-plus-cokriged values.

The amount of nitrate produced from the addition of 0.2 g of dried blood to 100 g of subsoil (WABLOD2) seems to be well correlated to the amount of nitrate existing in the subsoil itself, judging from the spherical cross-semivariogram in Fig. 19. The peak at around 20 feet in Fig. 19 is more pronounced than that found for the semivariograms of each variable shown in Figs. 10a and 10b. We do not have a clear explanation for this fact.

Nitrogen and carbon appear to be well correlated for both the Davis and Oakley soils, as shown by the correlation coefficients in Table 4 and Figs. 20 and 21. The two cross-semivariograms are almost identical in shape, even though carbon in Davis had no clear structure. This is a very important result, since for the variable WACADA, the only choice of geostatistical estimation we have is cokriging.

One point in common to all the cross-semivariograms above is the low nugget effect, which reflects the continuity of one variable with respect to the other on a small scale.

Kriging and Cokriging

Five variables were used as examples for the estimation methods used in this article: GRCLAY, WABLOD2, MOYLD, WACAOK, and WACADA. The contour maps for the estimation variance of cokriging are shown as an illustration of the different sampling schemes. Contour maps for the original values are shown in Figs. 22, 23, 24, 25, and 26, respectively.

Table 4. Parameters of the cross-semivariograms. Gamma is the model that best fit the experimental cross-semivariogram.

Z_1	Z_2	Gamma	C_0	C_1	Range	Sill	r
GRDEPTH	GRCLAY	linear	0.25	3.3×10^{-3}	600.0 ft	2.25	0.1406
MONIT	MOYLD	spherical	0	-1.20	25.0 ft	-1.20	-0.1200
RTA77	RTS77	spherical	2.5	10.00	6.0 d	12.50	0.5453
RTA78	RTS78	spherical	4.0	10.00	3.0 d	14.00	0.6739
RTA79	RTS79	spherical	2.5	10.50	7.0 d	13.00	0.7815
WASOIL2	WABLOD2	spherical	0	0.55	30.0 ft	0.55	0.3832
WANIDA	WACADA	Gaussian	0	5×10^{-3}	57.7 ft	5×10^{-3}	0.5531
WANIOK	WACAOK	Gaussian	0	4×10^{-4}	46.2 ft	4×10^{-4}	0.7410

Since GRCLAY was originally measured on an irregular grid, we estimated values to regularize the grid to 35 x 35 feet. The contour map for the measured-plus-kriged values is shown in Fig. 22, and the one for measured-plus-cokriged values is shown in Fig. 22. These maps show that the addition of kriged or cokriged values produced identical maps. Thus there is no gain in applying cokriging instead of kriging, so kriging should be preferred because it is easier. This is no surprise, since the correlation between GRDEPTH and GRCLAY is only 0.14 (Table 4), and the cross-semivariogram between these two variables is not very well defined (Fig. 17).

The contour map for the estimation variances of cokriging is shown in Fig. 22. This map does not add much to the interpretation of the ones for kriged or cokriged values, except for the fact that the locations were samples far from each other and had higher estimation variances (Fig. 22). Since the sampling was irregular, the map for the estimation variances also has little pattern. It will be interesting to compare this map with the one for other sampling schemes. Comparison between contour maps for kriged (Fig. 22) or cokriged (Fig. 22) does not reveal significant differences, except for the left top corner, where very few measurements were taken. This is exactly the area that shows the highest estimation variances. Obviously, if an increase in precision of estimation is sought, more samples should be taken in this area, decreasing the spacing between samples. Addition of other contour levels did not improve the appearance of the above maps, which is why they are shown as they are.

The contour maps for measured, kriged-plus-measured, and cokriged-plus-measured values and cokriging estimation variances for MOYLD are shown in Figs. 23a, 23b, 23c, and 23d, respectively. MOYLD contained only part of the original data set, with measurements considered only on an 11-foot-square grid. This was done to examine the effect of cokriging using the correlated variable MONIT, with the complete data set on a 5.5-foot-square grid. The contour maps for kriged and cokriged values, however, are practically identical, again reflecting the poor correlation between the variables. Both kriged and cokriged maps, however, are different from the one for the measured values alone. The high estimation variances shown in Fig. 23 reflect the high nugget effect of the semivariogram for MOYLD (see Table 2 and Fig. 9b). Because MOYLD had a high variance of 7729, estimation variances should be expected to reflect this value. The regions on the map in Fig. 23d with longer contours with label 6900 correspond to the regions in which MOYLD changes the fastest with distance, as shown by the posted numbers in Fig. 23a.

Figures 24a, 24b, 24c, and 24d show the contour maps for measured, measured-plus-kriged, measured-plus-cokriged, and cokriging estimation variances, respectively, for WABLOD2. Little difference is found between kriged and cokriged maps, except that the cokriged maps seem to be slightly smoother than the kriged. However, both estimated maps are very different from the map

for measured values alone. Both kriging and cokriging estimations were made every 5 feet for the entire area covered in the maps, and since the original sampling was made in 8 radii, the areas between the extremes of the radii away from the center have very few measurements. Again, those are the areas in which the estimation variances are the largest, as shown in Fig. 24c for the cokriging estimation variances. Except for its curious appearance, this map does not show much information, since this pattern should always be expected for this kind of sampling pattern.

Figures 25a, 25b, 25c, and 25d show the contour maps for the measured, measured-plus-kriged, measured-plus-cokriged, and cokriging estimation variance values, respectively, of WABLOD2. Both kriging and cokriging estimations were made for every 15-foot interval. Little difference is noticed between the map for estimated values and either of those kriged or cokriged and the map for the measured values. The reason for that is probably the large number of measurements. In Fig. 35d, only three contour levels were used: 0.0030, 0.0035, and 0.0040. The five areas with contour level 0.0030 correspond to areas more intensively sampled, with a sampling distance between 10 and 15 feet. The oval-shaped contours are the 0.0035 levels and correspond to the areas exactly around the measured values on a 30-foot-square grid. The diamond-shaped contours are the 0.0040 levels and correspond to areas in between measured values, where the distance between measured and estimated values was the greatest.

The semivariogram model used for the kriging estimation to obtain the contour map in Fig. 25b had a pure nugget effect, with the value equal to the calculated variance, 0.019. With this model the solution of the kriging system produces identical weights for any neighboring distance. The similarity of the kriging to the cokriging map is attributed to the low correlation between WANIDA and WACADA. We believe that if the number of samples of WACADA were smaller, with larger sampling distances compared with WANIDA, these maps would be significantly different. We can cokrige WACADA using its nugget semivariogram, the semivariogram for WANIDA, and the cross-semivariogram between them, even though WACADA itself was randomly distributed in space, which means that we can decrease the number of samples for a variable difficult to measure (even if the variable has no structure) and still cokrige, provided the correlated variable has adequate structure and cross-structure.

The contour maps for the measured, kriged-plus-measured, and cokriged-plus-measured values of WACAOK, shown in Figs. 26a, 26b, and 26c, respectively, are quite different from each other, which reflects the correlation between WANIOK and WACAOK of 0.74 and their structures and cross-structures, shown in Figs. 12a, 12b, and 21. The discussion for the estimation variance contour map in Fig. 25d is also applicable to the one in Fig. 26d for WACAOK.

The kriging estimations were done using the computer program KRIGE and the required subroutines, listed in Appendix A4. Similarly, the cokriging estimations were made using the computer program COKRI listed in Appendix A5.

Estimation Error and Neighborhood of Estimation

The "jackknifing" procedure described earlier was used to aid in selecting size of the neighborhood of estimation and validation of the assumptions under which the estimations were made. As an example, the results from "jackknifing" WACAOK with increasing neighborhood size are shown in Table 5. Ideally, when this kind of calculation is done, the mean reduced error is zero and the variance of the reduced error is one (Delhomme, 1976). However, when the estimation variances are less than one, the variance of the reduced errors may become large and different from 1. An alternative is to use the absolute error: the mean should be close to zero and variance should be small enough to allow for 95% of the error within $m_\varepsilon \pm 2\sigma_\varepsilon$ (Journel and Huijbregts, 1978, p. 49). This is the same as requiring the estimation errors to be normally distributed. A third alternative is to examine the linear regression line between measured and estimated values, and as the slope approaches one and the intercept approaches zero, the ideal situation is being achieved. However, the third alternative really has no basis, since all that either kriging or cokriging is required to do is to be unbiased and to have a minimum variance. That only means that the estimated values need to be equal to the measured values on the mean and the variance of estimation must be the minimum, but point-to-point agreement is not guaranteed.

In Table 5, except where indicated, the neighborhood used was based on a fixed number of neighbors. An overall judgment of the results in Table 5 using the concepts above reveals that the results seem to be best when four neighbors are used for each estimation, although not very much difference is found when other numbers of neighbors are used. When all neighbors within 30 feet were used, the results were a little better than those with four neighbors. We have not always found this result when using a completely regular square grid, since, for the regions away from the edges of the field, the two neighborhood choices produce identical results. At the edges, however, fewer neighbors are used in each estimation when a fixed neighborhood distance is used, as compared with the number of neighbors. In this sense, for a completely regular square grid, a number of neighbors that is a multiple of four (4, 8, 16, 32, etc.) seems to produce better results than the corresponding distance neighborhoods. If WACAOK were sampled on a completely regular square grid, then all neighbors within 30 feet, or four neighbors would be identical neighborhoods, since the basic sampling distance was 30 feet.

The fact that there are five clusters of measurements in the field at distances closer than 30 feet may have given the distance neighborhood some advantage over number of neighbors. For irregular grids of sampling, such as for GRCLAY, or even the radial sampling for WABLOD2, a neighborhood based on a fixed distance is quite impractical for programming purposes. The reason for that is the difficulty to predict the number of neighbors for a given point, causing difficulties in dimensioning the matrix system. However, it is equally

Table 5. Jackknifing for WACAOK. Mean and variance for measured values are 0.43 and 0.013, respectively.

Number of	Jackknifed			Absolute Error		Reduced Error		Regression Parameters		
neighbors	Mean	Variance	COV	Mean	Variance	Mean	Variance	Intercept	slope	r
2	0.4254	0.0074	0.0056	-0.0076	0.0093	-0.094	1.815	0.1104	0.7582	0.5697
4	0.4277	0.0068	0.0056	-0.0052	0.0087	-0.053	1.978	0.0790	0.8275	0.5938
6	0.4288	0.0071	0.0058	-0.0041	0.0087	-0.034	1.995	0.0840	0.8138	0.5995
8	0.4292	0.0073	0.0058	-0.0038	0.0089	-0.027	2.089	0.0932	0.7915	0.5888
10	0.4269	0.0070	0.0054	-0.0061	0.0092	-0.057	2.156	0.1002	0.7795	0.5682
15	0.4271	0.0070	0.0054	-0.0059	0.0093	-0.060	2.200	0.1010	0.7774	0.5659
20	0.4275	0.0070	0.0054	-0.0055	0.0094	-0.056	2.249	0.1049	0.7674	0.5614
25	0.4261	0.0069	0.0052	-0.0069	0.0097	-0.079	2.328	0.1145	0.7474	0.5428
30[1]	0.4289	0.0068	0.0057	-0.0040	0.0085	-0.030	1.986	0.0731	0.8389	0.6025
60[1]	0.4278	0.0073	0.0057	-0.0051	0.0090	-0.047	2.139	0.0979	0.7832	0.5842
4[2]	0.4303	0.0077	0.0069	-0.0027	0.0069	-0.023	1.367	0.0442	0.9037	0.6909

[1] All neighbors within distance specified were used.

[2] Cokriging was used.

difficult to predict the distances over which neighbors are located when fixed numbers of neighbors are used. Moreover, a neighborhood based on a fixed number of neighbors is not easily interpreted from the physical sense, because the kriging and cokriging systems constitute the semivariogram and cross-semivariogram values computed for specified distances, not number of neighbors. In short, what is easily programmable does not make much physical sense, and what makes physical sense is not easily programmable unless for idealized situations such as a completely regular square grid.

The calculations contained in Table 5 show the effect of the number of neighbors on the estimations and serve as an example of the neighbor effect. Also shown in Table 5, that the results produced from cojackknifing WACAOK with four neighbors of each WANIOK and WACAOK are better than any jackknifing results.

All variables estimated or used in estimation in this article were subject to the procedure above until the best and smallest neighborhood could be selected.

SUMMARY AND CONCLUSIONS

A long and detailed theoretical development was necessary for the reader to understand what kriging and cokriging equations are and how they can be used effectively. To illustrate the theory, a wide variety of data sets was used as practical examples using field data from many agronomic disciplines such as remote sensing, soil science, and agronomy.

Semivariograms calculated for a variety of data sets such as surface temperature, yield of turkey wheat grains, and carbon percentage of the soil, obtained from a variety of sampling schemes, such as transect, regular square grid, and radial direction, had ranges from a few meters to more than 100 meters, even though many of these variables approach a normal distribution.

Removing the regional drift from air and surface temperature data with a 29-point moving average rendered the residual correlated to around 5-6 days and enabled the design of frequency for sampling to allow for estimation at any desired finer interval. The semivariograms of the residuals had a reasonably defined sill and the drifts computed for the residuals were practically negligible. Semivariograms for the residuals for different years were somewhat

different, having different ranges, nugget effects, and sills, most likely owing to extreme climatic changes for the years analyzed. The 29-point moving average removed the seasonal drift better for 1978 than for the other two years. The reason for this has not been completely identified, but the winter season of 1978 had higher temperatures than the other two years, which may have caused this difference. The cross-semivariograms between the residuals of air temperature and the residuals of surface temperature follow one single spherical model closely. This is a desirable result that needs more research for verification. If after some years of experimentation the same relationship is repeated, the cross-semivariogram can be catalogued for those experimental conditions and can be used with high confidence for sampling design and estimation of unrecorded data.

Soil survey variables had a strong spherical structure with ranges within 300 feet. The cross-semivariogram between depth to 40% clay and clay percent of the plow layer had a linear structure with a range of 600 ft even though the classical correlation coefficient between the two variables was not very high.

Semivariograms for nitrogen content of turkey wheat grains and yield of grains had good structure, with ranges of approximately 20 feet. The cross-semivariogram between these variables had a negative spherical structure with a range of 25 feet. The negative structure does not make much sense, because in general it is not expected that as the nitrogen content of the grains increases, the yield of grain decreases. This was in fact part of the thought of the author (Montgomery, 1913).

Soil nitrate data obtained from samples collected in Davis, California, by Waynick (1918) in a radial sampling scheme (see Fig. 2) produced semivariograms and cross-semivariograms in general not well defined for detecting presence of structure. The causes have not been identified and are not thought to be related to the sampling scheme. On the other hand, carbon and nitrogen contents of the soils of Davis and Oakley have very well defined structure and cross-structure, except for carbon in Davis which showed a pure nugget effect. Semivariograms and cross-semivariograms for carbon and nitrogen in Oakley samples had very similar structures which probably differ only by a constant. Another curious point is that the cross-semivariograms for nitrogen and carbon in Davis had structure, even though the semivariogram for carbon itself showed a pure nugget effect.

Contour maps were drawn for GRCLAY, MOYLD, WABLOD2, WACADA, and WACAOK for measured values alone, kriged-plus-measured values, cokriged-plus-measured values, and estimation variances of cokriging. The contour maps for estimated values are generally different from the ones for measured value alone, which is caused by the addition of unbiased estimated values. However, the contours obtained for kriged-plus-measured values are identical to the ones obtained for cokriged-plus-measured for all variables except WACAOK. For those variables, therefore, there is no gain in using cokriging instead of kriging. For WACAOK,

however, the three maps are different from each other, reflecting the good correlation between WANIOK and WACAOK. The contour maps for the estimation variances generally show what common sense would indicate: high variances for the locations at which samples are far from each other, and low variances for those close to each other.

Jackknifing has been shown to be useful as a guide in the choice of estimation neighborhoods and in validation of assumptions made. This procedure should be used routinely before any estimation is made.

From our view, future research should be moving toward measuring field variables for different environmental conditions, including the soil itself, and computing semivariograms and cross-semivariograms, which we hope will show results similar to the ones shown in Fig. 16 for many years of experiments.

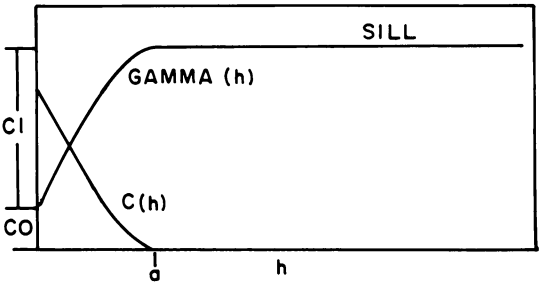

Figure 1
Semivariogram. Range, a, is the distance at which $\gamma(h) + \varepsilon$ = sill.

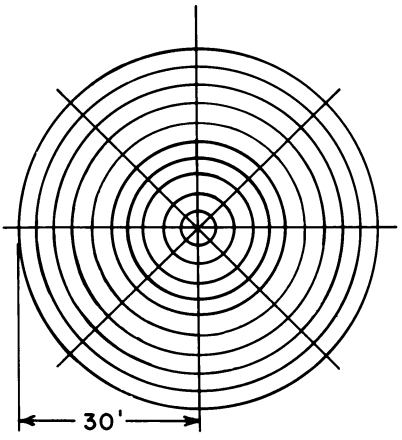

Figure 2
Sampling diagram for Waynick (1918).

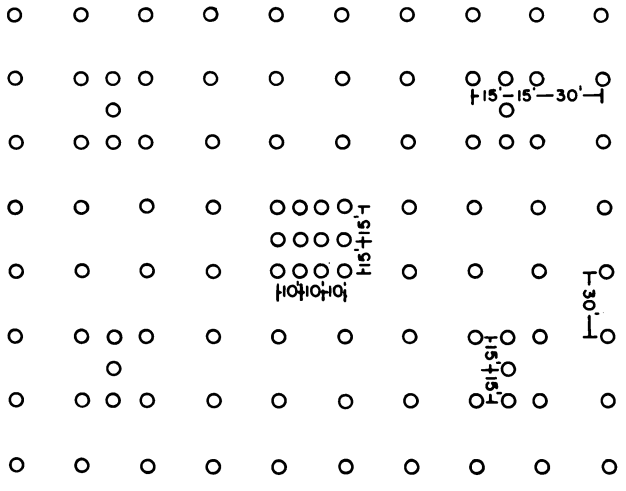

Figure 3
Sampling diagram for Waynick and Sharp (1919)

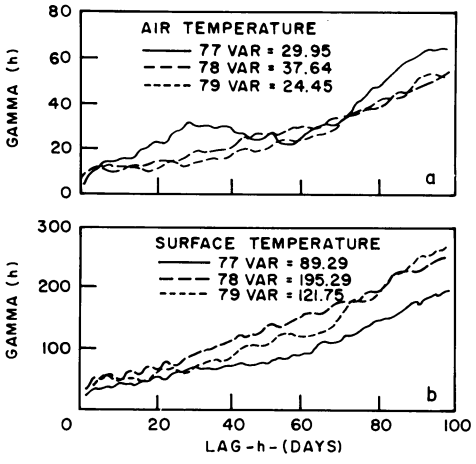

Figure 4
Semivariogram for raw values of a) TA and b) TS for the three years.

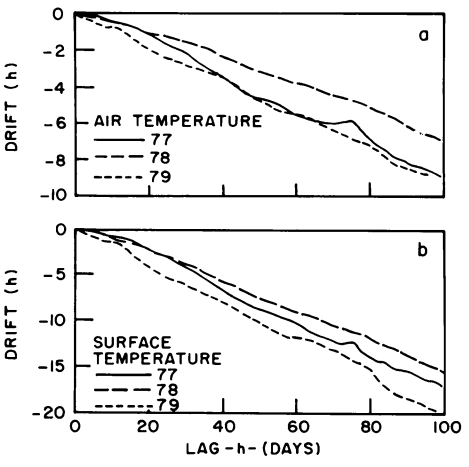

Figure 5
Drifts for the raw values of a) TA and b) TS for the three years.

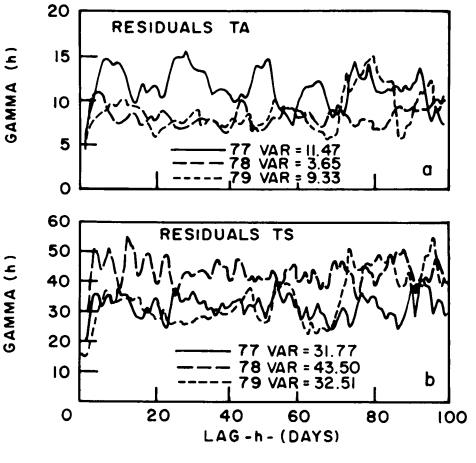

Figure 6
Residuals a) semivariograms for TA and b) semivariograms for TS for the three years.

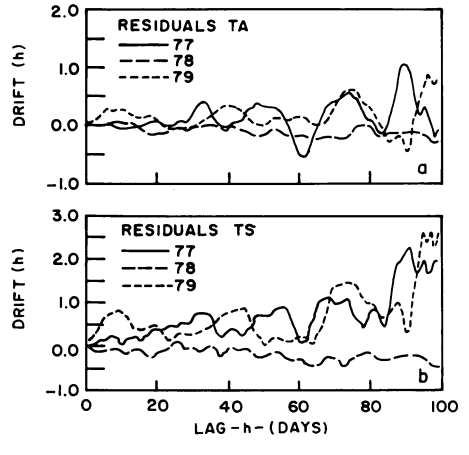

Figure 7
Drift for residuals of a) TA and b) TS for the three years.

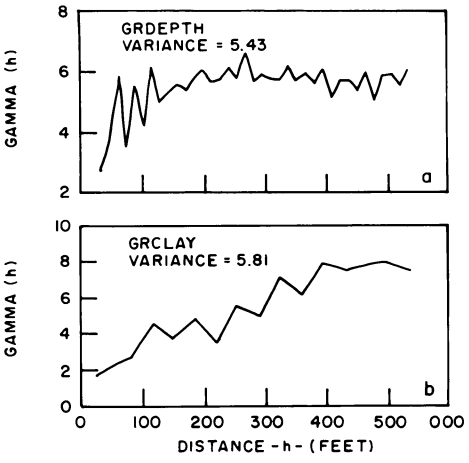

Figure 8

a) Semivariogram for GRDEPTH and b) semivariogram for GRCLAY.

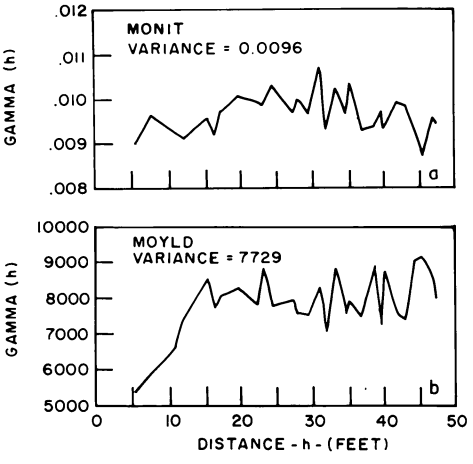

Figure 9

a) Semivariogram for MONIT and b) semivariogram for MOYLD.

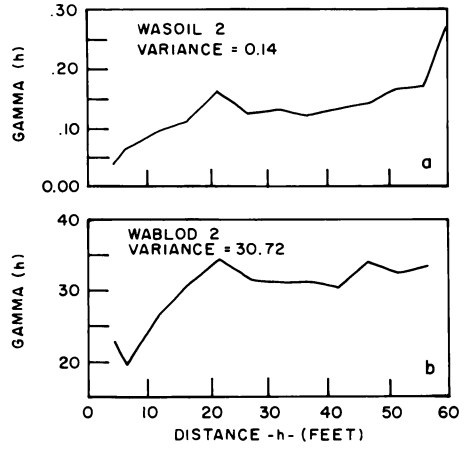

Figure 10

a) Semivariogram for WASOIL2 and b) semivariogram for WABLOD2.

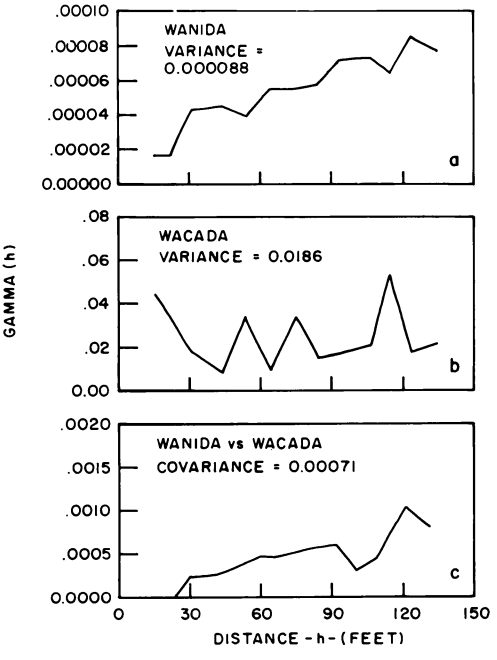

Figure 11

a) Semivariogram for WANIDA and b) semivariogram for WACADA.

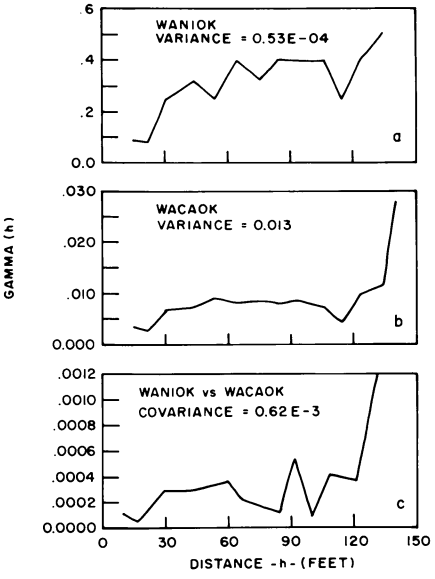

Figure 12

a) Semivariogram for WANIOK and b) semivariogram for WACAOK.

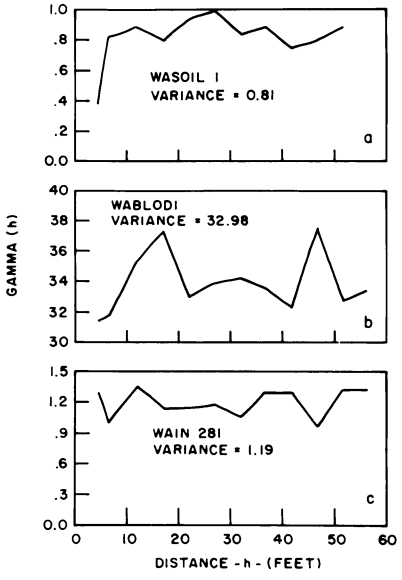

Figure 13

Semivariogram for a) WASOIL1, b) WABLOD1, and c) WAIN281.

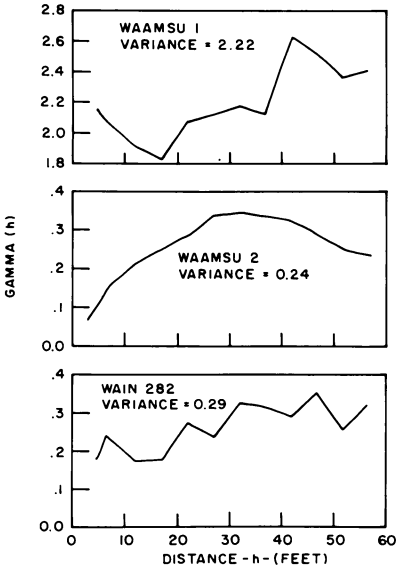

Figure 14

Semivariogram for a) WAAMSU1, b) WAAMSU2, and c) WAIN282.

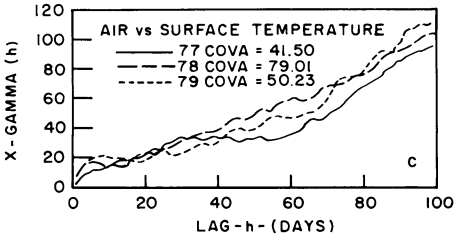

Figure 15

Cross-semivariogram for TA vs TS for the three years.

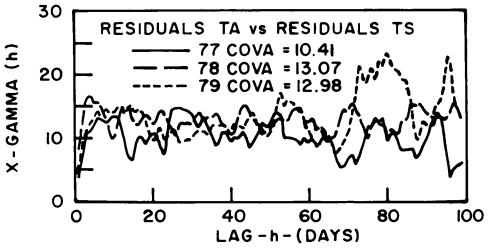

Figure 16

Cross-semivariograms for calculated residuals of TA vs TS for the three years.

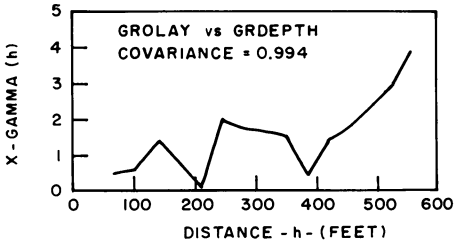

Figure 17

Cross-semivariogram for GRDEPTH vs GRCLAY.

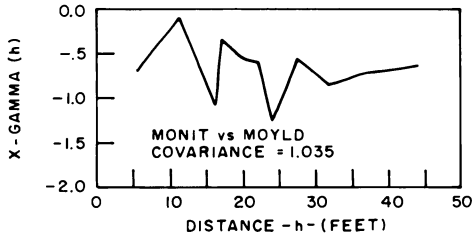

Figure 18

Cross-semivariogram for MONIT vs MOYLD.

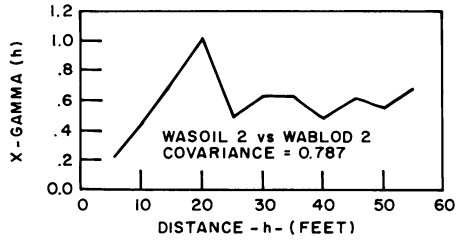

Figure 19

Cross-semivariogram for WASOIL2 vs WABLOD2.

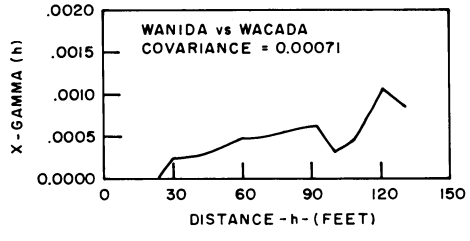

Figure 20

Cross-semivariogram for WANIDA vs WACADA.

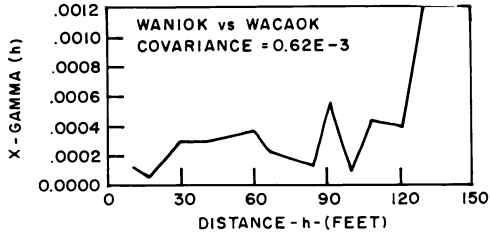

Figure 21

Cross-semivariogram for WANIOK vs WACAOK.

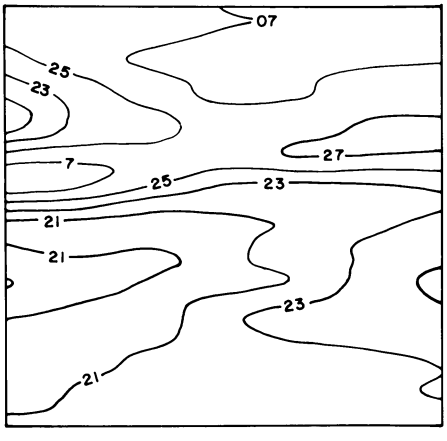

Figure 22a

Contour maps for GRCLAY, measured.

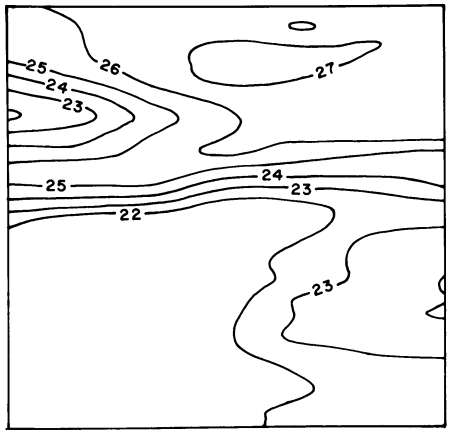

Figure 22b

Contour maps for GRCLAY, measured-plus-kriged.

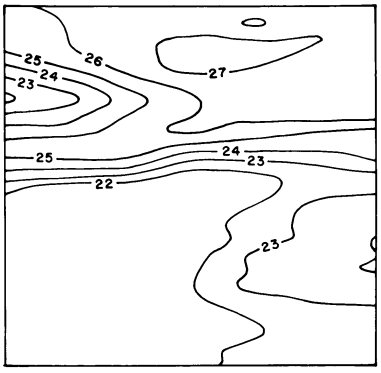

Figure 22c
Contour maps for GRCLAY, measured-plus-cokriged.

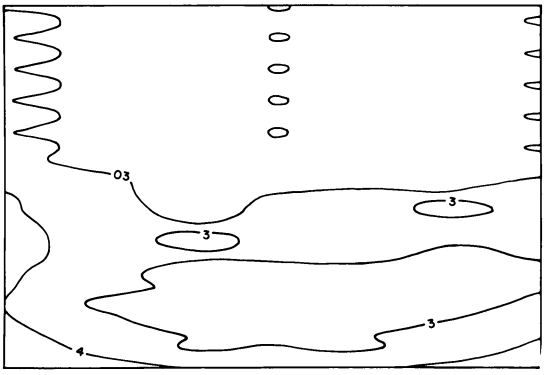

Figure 22d
Contour maps for GRCLAY, estimation variance of cokriging.

Figure 23a
Contour maps for MOYLD, measured.

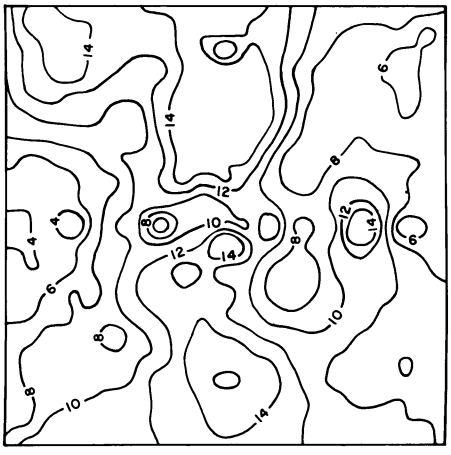

Figure 23b
Contour maps for MOYLD, measured plus kriged.

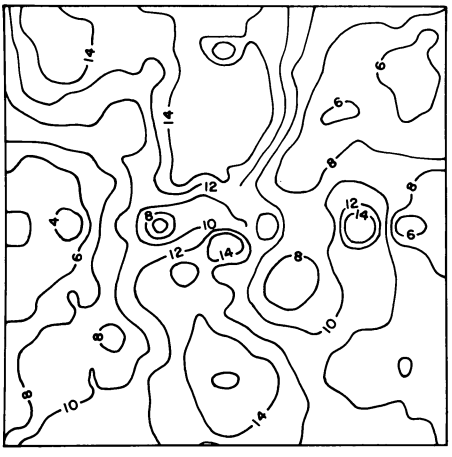

Figure 23c
Contour maps for MOYLD, measured plus cokriged.

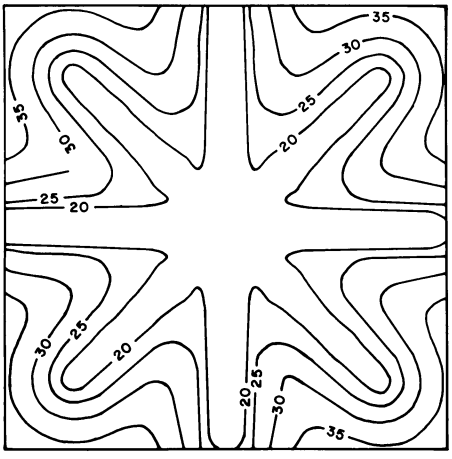

Figure 23d
Contour maps for MOYLD, estimation variance of cokriging.

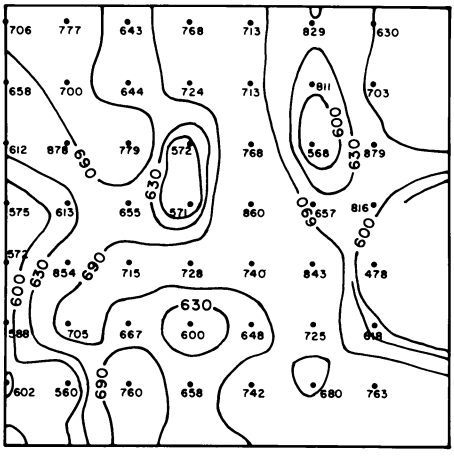

Figure 24a

Contour maps for WABLOD 2, measured.

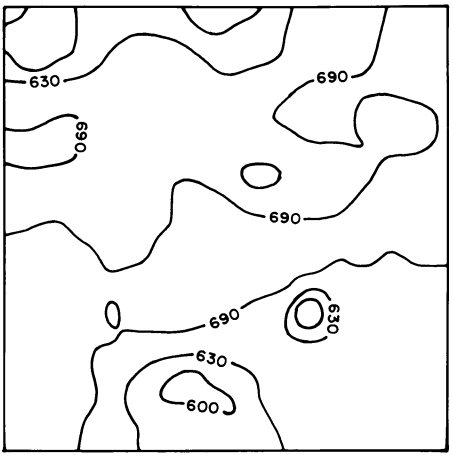

Figure 24b

Contour maps for WABLOD 2, measured plus kriged.

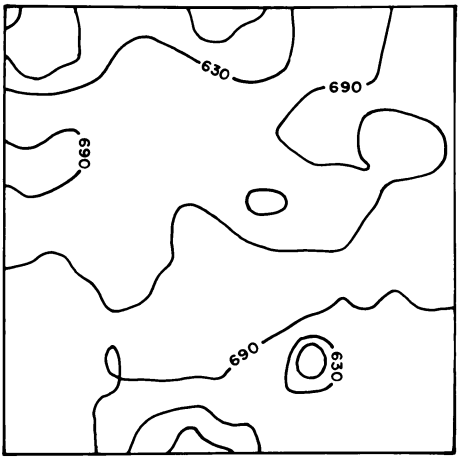

Figure 24c

Contour maps for WABLOD 2, measured plus cokriged.

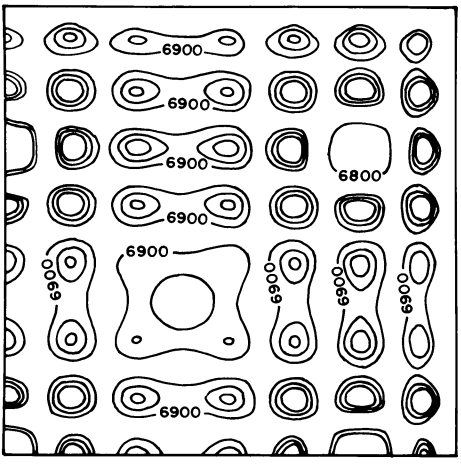

Figure 24d

Contour maps for WABLOD 2, estimation variance of cokriging.

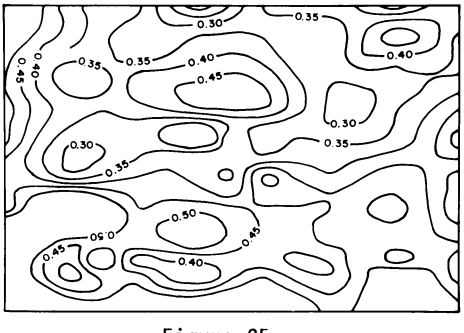

Figure 25a

Contour maps for WACAOK, measured.

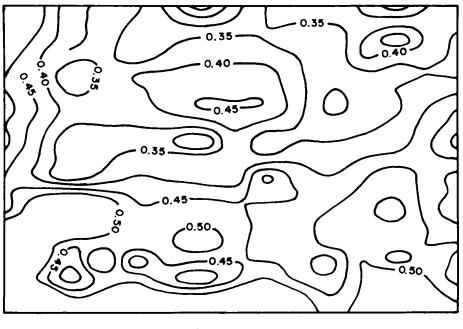

Figure 25b

Contour maps for WACAOK, measured plus kriged.

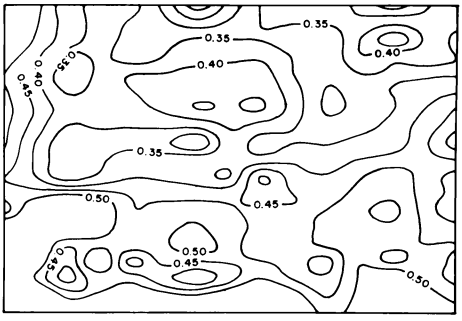

Figure 25c
Contour maps for WACAOK, measured plus cokriged.

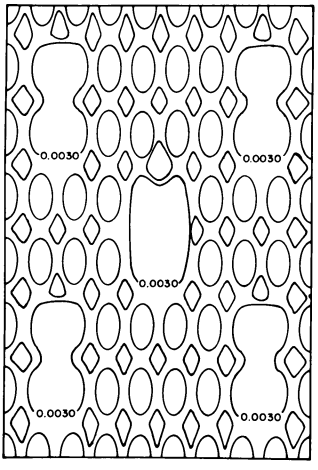

Figure 25d
Contour maps for WACAOK, estimation variance of cokriging.

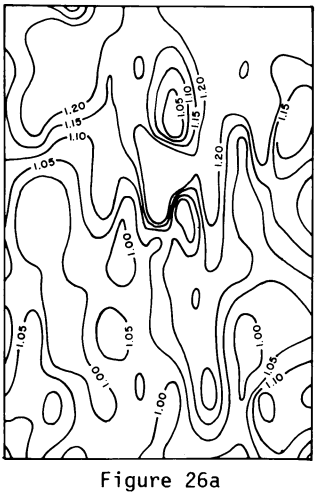

Figure 26a
Contour maps for WACADA, measured.

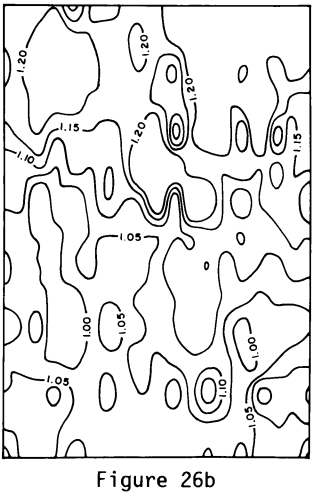

Figure 26b
Contour maps for WACADA, measured plus kriged.

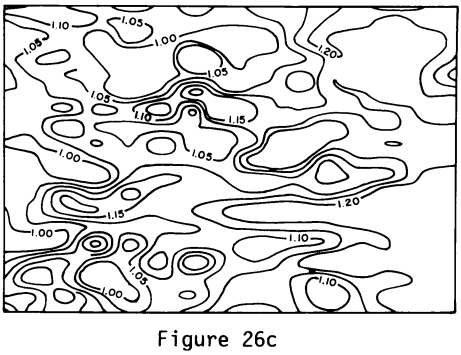

Figure 26c
Contour maps for WACADA, measured plus cokriged.

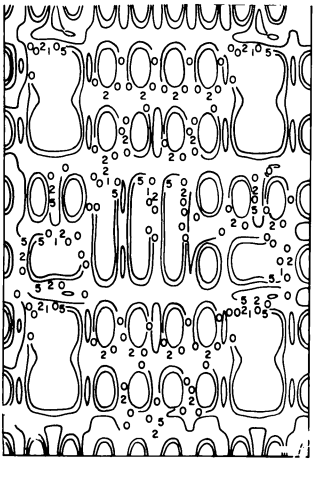

Figure 26d
Contour maps for WACADA, estimation variances of cokriging.

APPENDIX

A. Computer Programs

1. GAMA1D--Semivariogram in one-dimension.
2. AVARIO--Average semivariogram.
3. XGAMA--Average cross-semivariogram.
4. KRIGE--Kriging program.
5. COKRI--Cokriging program.
6. CORSP--Spectral analysis program.
7. SUBROUTINES--Called by KRIGE, COKRI, and SPECT.
   a. NBHOOD--Selects the neighborhood of estimation.
   b. KRIX--Builds the kriging system.
   c. COKMS--Builds the cokriging system.
   d. SIMQ--Solves a linear system.
   e. REG--Calculates the jackknifing parameters.
   f. CIV--Calculates gamma and x-gamma values.
   g. SORT--Sorts neighbors by distance.
   h. AUTO--Calculates autocorrelation function.
   i. SPECT--Calculates spectral density.

```
      PROGRAM GAMA1D
C ************************************************************
C
C                  VARIOGRAM:ONE DIMENSIONAL
C
C                  BY SIDNEY VIEIRA
C
C  INPUT VARIABLES:
C
C  ZIN( ):EXPERIMENTAL VALUES
C  XIN( ):COORDINATE POSITION
C
C   INPUT CONSTANTS
C
C   NR:NUMBER OF ROWS ( Y DIRECTION )
C   NLAG:NUMBER OF LAGS CALCULATED
C   NSTP:INCREMENT IN CALCULATION OF GAMMAS
C
C   OUTPUT VARIABLES:
C
C   NCPL( ):NUMBER OF COUPLES(PAIRS) IN EACH CALCULATION
C   G( ): VARIOGRAM VALUES.
C   D( ): AVERAGE DISTANCE BETWEEN COUPLES.
C
C   OUTPUT CONSTANTS
C
C   AMEAN:ARITMETIC MEAN OF EXPERIMENTAL VALUES
C   VAR:VARIANCE
C   STDEV:STANDARD DEVIATION
C   VARCOF:COEFFICIENT OF VARIATION
C
C   OBSERVATIONS:
C
C   1. INPUT VARIABLE ZIN IS <NR> LONG.
C   2. OUTPUT VARIABLES ARE 'NLAG' LONG.
C   3. MAXIMUN DISTANCE GAMMA IS CALCULATED IS
C      <MAX=(NLAG-1)*NSTP>
C
C ************************************************************
C
C  INITIALIZATION
C
      DIMENSION ZIN(500),XIN(500),NCPL(201),
     1G(201),D(201)
      INTEGER IN,OUT
      BYTE FILNAM(28)
      DATA IN/1/,OUT/6/
      WRITE(5,999)
  999 FORMAT(' ENTER FILENAME ',$)
      READ(5,1001) FILNAM
 1001 FORMAT(28A1)
```



```
      FILNAM(28)=0
      WRITE(5,1000)
 1000 FORMAT(' ENTER # OF LAGS,AND STEP ',$)
      READ(5,*) NLAG,NSTP
      SUMZ=0.0
      SUMZ2=0.0
      DO 1 I=1,NLAG
      NCPL(I)=0
      D(I)=0.0
      G(I)=0.0
    1 CONTINUE
C
C OPEN INPUT AND OUTPUT FILES
C
      OPEN(UNIT=IN,NAME=FILNAM,TYPE='OLD',FORM='FORMATTED')
      OPEN(UNIT=OUT,NAME='SY:GAMA.LST',TYPE='NEW')
C
C CALCULATE VARIOGRAM(GAMMA) VALUES
C
      MAX=(NLAG-1)*NSTP
      J=1
      DO 2 I=1,500
      READ(IN,99,END=2000,ERR=2001) XIN(J),ZIN(J)
      J=J+1
    2 CONTINUE
 2000 NR=J-1
      WRITE(5,600) NR
      WRITE(5,3000)(XIN(I),ZIN(I),I=1,NR)
 3000 FORMAT(2(3X,F6.2))
C
C ACCUMULATE SUMS OF EXPERIMENTAL VALUES
C
      DO 42 K42=1,NR
      SUMZ2=SUMZ2+ZIN(K42)*ZIN(K42)
      SUMZ=SUMZ+ZIN(K42)
   42 CONTINUE
      N2=NR-1
      DO 51 K51=1,N2
      L=K51+1
      ZI2=ZIN(K51)
      XI2=XIN(K51)
         DO 61 K61=L,NR
         DX=XIN(K61)-XI2
         H=IFIX(DX)
         IF(H.GT.MAX) GO TO 51
         L4=H/NSTP+0.5
         NCPL(L4)=NCPL(L4)+1
         D(L4)=D(L4)+DX
         DZ=ZI2-ZIN(K61)
         G(L4)=G(L4)+0.5*DZ*DZ
   61    CONTINUE
```

```
   51 CONTINUE
      DO 45 K45=1,NLAG
      CP=FLOAT(NCPL(K45))
      IF(CP.EQ.0.) GO TO 45
      D(K45)=D(K45)/CP
      G(K45)=G(K45)/CP
   45 CONTINUE
C
C CALCULATE STATISTICS
C
      AMEAN=SUMZ/FLOAT(NR)
      VAR=SUMZ2/FLOAT(NR)-AMEAN*AMEAN
      STDEV=SQRT(VAR)
      VARCOF=STDEV/AMEAN
C
C  OUTPUT:WRITING
C
      WRITE(6,100) FILNAM
      WRITE(6,200)NR,AMEAN,VAR,STDEV,VARCOF
      WRITE(6,300)
      WRITE(6,400)
      DO 30 I=1,NLAG
      IF(G(I).EQ.0.0)GO TO 30
      WRITE(6,500) I,NCPL(I),D(I),G(I)
   30 CONTINUE
      GO TO 9
 2001 WRITE(5,700)
C
C  FORMATS
C
   99 FORMAT(2(3X,F6.2))
  100 FORMAT(10X,28A1)
  200 FORMAT(//3X,'# OF SAMPLES=',I6/3X,'MEAN=',F11.4/3X,'VA
     1F11.4/3X,'STANDARD DEVIATION=',F11.4/3X,'COEFFICIENT O
     2IATION=',F11.4)
  300 FORMAT(//15X,'VARIOGRAM : ONE DIMENSION'//)
  400 FORMAT(10X,1H/'STEP    # OF COUPLES    AVERAGE DISTANCE
  500 FORMAT(10X,I3,3X,I7,5X,F14.2,3X,F17.4)
  600 FORMAT(3X,' END OF FILE REACHED AFTER ',I4,' NUMBERS R
  700 FORMAT(3X,'ERROR DURING READING')
    9 CLOSE(UNIT=OUT)
      CLOSE(UNIT=IN)
      STOP
      END
```



```
      PROGRAM AVARIO
C ********************************************************
C
C                 VARIOGRAM:ALL DIRECTIONS (AVERAGE)
C
C                 BY SIDNEY VIEIRA
C
C  INPUT VARIABLES:
C
C  X(  ):X-COORDINATE OF EXPERIMENTAL VALUES
C  Y(  ):Y-COORDINATE OF EXPERIMENTAL VALUES
C  Z(  ):EXPERIMENTAL VALUES
C
C   INPUT CONSTANTS
C
C   N :NUMBER OF EXP. VALUES ( GENERATED IN THE READ LOOP)
C   NLAG:NUMBER OF LAGS CALCULATED
C   NSTP:INCREMENT IN CALCULATION OF GAMMAS
C
C   OUTPUT VARIABLES:
C
C   PAIR( ):NUMBER OF COUPLES(PAIRS) IN EACH CALCULATION
C   X1(I,1):AVERAGE DISTANCE BETWEEN PAIRS
C   X1(I,2): VARIOGRAM VALUE
C
C   OUTPUT CONSTANTS
C
C   AMEAN:ARITMETIC MEAN OF EXPERIMENTAL VALUES
C   VAR:VARIANCE
C   STDEV:STANDARD DEVIATION
C   VARCOF:COEFFICIENT OF VARIATION
C
C   OBSERVATIONS:
C
C   1. INPUT VARIABLES ARE '(N)' LONG (1-D ARRAY).
C   2. OUTPUT VARIABLES ARE 'NLAG' LONG.
C   3. MAXIMUN DISTANCE GAMMA IS CALCULATED IS
C      <MAX=(NLAG-1)*NSTP>
C   4. CAPACITY
C             SAMPLES: 200
C           LAGS   : 100
C   5. OPTIONS
C            ILOG=1 VARIOGRAM CALCULATED ON ORIGINAL DATA
C          ILOG=2 VARIOGRAM CALCULATED ON LOGARITHMS OF DATA
C          IPLOT=1 NO PLOT
C          IPLOT=2 GAMMAS VERSUS DISTANCES IS PLOTTED ON LINEPR
C
C ********************************************************
C
C  INITIALIZATION
C
```

```
      DIMENSION X(200),Y(200),Z(200),PAIR(100),
     1X1(100,2)
      BYTE FILNAM(28)
      INTEGER IN1,IN2,OUT
      DATA IN1/1/,IN2/2/,OUT/6/
C
C   OPEN FILE FOR INPUT PARAMETERS
C
      OPEN(UNIT=IN2,NAME='SY:PARAVARIO.DAT',TYPE='OLD',FORM=
C
C   READ IN INPUT PARAMETERS
C
      READ(2,1000) FILNAM
      READ(2,*) NLAG,NSTP
      READ(2,*) IPLOT,ILOG
      FILNAM(28)=0
C
C   CLOSE FILE FOR INPUT PARAMETERS
C
      CLOSE(UNIT=IN2)
      SUMZ=0.0
      SUMZ2=0.0
      DO 12 J=1,NLAG
      PAIR(J)=0.0
      DO 1 I=1,2
      X1(J,I)=0.0
    1 CONTINUE
   12 CONTINUE
C
C   OPEN FILES FOR INPUT VARIABLES AND FOR OUTPUT VARIABLES
C
      OPEN(UNIT=IN1,NAME=FILNAM,TYPE='OLD',FORM='FORMATTED')
      OPEN(UNIT=OUT,NAME='SY:GAMA.LST',TYPE='NEW')
C
C READ IN INPUT VARIABLES
C
      WRITE(6,1001) FILNAM
      J=0
      DO 15 I=1,200
      READ(1,500,END=2000,ERR=2001) X(I),Y(I),Z(I)
      J=J+1
   15 CONTINUE
 2000 N=J
      WRITE(5,3000) N
      GOTO (201,200) ILOG
  200 WRITE(6,401)
      DO 11 I=1,N
      Z(I)=ALOG(Z(I))
      WRITE(6,501) X(I),Y(I),Z(I)
   11 CONTINUE
      GO TO 300
```



```
  201 WRITE(6,400)
      WRITE(6,501)(X(I),Y(I),Z(I),I=1,N)
C
C ACCUMULATE SUMS OF EXPRIMENTAL VALUES
C
  300 DO 20 J=1,N
      SUMZ=SUMZ+Z(J)
      SUMZ2=SUMZ2+Z(J)*Z(J)
   20 CONTINUE
C
C CALCULATE STATISTICS
C
      AMEAN=SUMZ/FLOAT(N)
      VAR=SUMZ2/FLOAT(N)-AMEAN*AMEAN
      STDEV=SQRT(VAR)
      VARCOF=STDEV/AMEAN
C
C CALCULATE VARIOGRAM(GAMMA) VALUES
C
      N1=N-1
      MAX=(NLAG-1)*NSTP
      DO 3 K3=1,N1
      XT=X(K3)
      YT=Y(K3)
      ZT=Z(K3)
      L=K3+1
         DO 4 K4=L,N
         DX=XT-X(K4)
         DY=YT-Y(K4)
         DZ=ZT-Z(K4)
         DHS=DX*DX+DY*DY
         IF(DHS.LT.1.0E-6) GO TO 4
         DH=SQRT(DHS)
         H=FLOAT(DH)
         IF(H.GT.MAX)GO TO 4
         L4=H/NSTP+1
         PAIR(L4)=PAIR(L4)+1.0
         X1(L4,1)=X1(L4,1)+DH
         X1(L4,2)=X1(L4,2)+0.5*DZ*DZ
    4    CONTINUE
    3 CONTINUE
      DO 5 K5=1,NLAG
      IF(PAIR(K5).EQ.0)GO TO 5
      X1(K5,1)=X1(K5,1)/PAIR(K5)
      X1(K5,2)=X1(K5,2)/PAIR(K5)
    5 CONTINUE
C  OUTPUT:WRITING
C
      WRITE(6,502)AMEAN,VAR,STDEV,VARCOF
      WRITE(6,503) FILNAM
      WRITE(6,504)
```

```
      DO 30 I=1,NLAG
      IF(X1(I,2).EQ.0.0)GO TO 30
      WRITE(6,505) I,PAIR(I),X1(I,1),X1(I,2)
   30 CONTINUE
      GOTO(203,202) IPLOT
C
C   2) PLOTTING
C
  202 CALL PLOT1(X1,NLAG,100)
  203 GO TO 301
 2001 WRITE(5,3001)
C
C  FORMATS
C
  400 FORMAT(/8X,'X',10X,'Y',10X,'Z',/)
  401 FORMAT(/8X,'X',10X,'Y',6X,'LOG(Z)',/)
  500 FORMAT(3(2X,F6.2))
  501 FORMAT(3(4X,F6.2))
  502 FORMAT(3X,'MEAN=',F11.4,/3X,'VARIANCE=',F11.4/3X,
     1'STANDARD DEVIATION=',F11.4/3X,'VARIATION COEFFICIENT=
     2,F11.4)
  503 FORMAT(//15X,'VARIOGRAM : ',28A1//)
  504 FORMAT(9X,'STEP   # OF COUPLES   AVERAGE DISTANCE    G
  505 FORMAT(10X,I3,3X,F7.2,5X,F14.2,3X,F17.4)
 1000 FORMAT(28A1)
 1001 FORMAT(10X,' EXPERIMENTAL VALUES FOR : ',28A1/)
 3000 FORMAT(' END OF FILE REACHED AFTER ',I3,' NUMBERS READ
 3001 FORMAT(' ERROR DURING READING ')
  301 CLOSE(UNIT=IN1)
      CLOSE(UNIT=OUT)
      STOP
      END
```



```
      PROGRAM XGAMA
C ********************************************************
C
C                CROSS-VARIOGRAM:ONE DIMENSIONAL
C
C                BY SIDNEY VIEIRA
C
C  INPUT VARIABLES:
C
C  Z1 AND Z2:EXPERIMENTAL VALUES FOR THE TWO TRANSECTS
C  X1 :COORDINATE POSITION OF TWO TRANSECTS
C
C   INPUT CONSTANTS
C
C   NR1 AND NR2:NUMBER OF ROWS ( Y DIRECTION ) FOR TWO TRANS
C   NLAG:NUMBER OF LAGS CALCULATED
C   NSTP:INCREMENT IN CALCULATION OF GAMMAS
C
C   OUTPUT VARIABLES:
C
C   NCPL( ):NUMBER OF COUPLES(PAIRS) IN EACH CALCULATION
C   G( ): CROSS-VARIOGRAM VALUES.
C   D( ): AVERAGE DISTANCE BETWEEN COUPLES.
C
C   OUTPUT CONSTANTS
C
C   AVE1 AND AVE2:ARITMETIC MEANS OF EXPERIMENTAL VALUES
C   VAR1 AND VAR2:VARIANCES
C   DEV1 AND DEV2:STANDARD DEVIATIONS
C   CV1 AND CV2:COEFFICIENTS OF VARIATION
C
C   OBSERVATIONS:
C
C   1. INPUT VARIABLES Z1 AND Z2 ARE <NR1> AND <NR2> LONG.
C   2. OUTPUT VARIABLES ARE 'NLAG' LONG.
C   3. MAXIMUN DISTANCE X-GAMMA IS CALCULATED IS
C      <MAX=(NLAG-1)*NSTP>
C   4. X-COORDINATE POSITIONS MUST BE THE SAME FOR
C      TWO TRANSECTS.
C   5. SHORTER TRANSECT MUST BE READ FIRST.
C
C ********************************************************
C
C  INITIALIZATION
C
      DIMENSION Z1(500),Z2(500),NCPL(100),X1(500),X2(500),
     1G(100),D(100)
      BYTE FN1(28),FN2(28)
      INTEGER IN1,IN2,OUT
      DATA IN1/1/,IN2/2/,OUT/6/
      WRITE(5,999)
```

```
  999 FORMAT(' ENTER FILENAME FOR FISRT SERIES ',$)
      READ(5,1001) FN1
 1001 FORMAT(28A1)
      FN1(28)=0
      WRITE(5,1010)
 1010 FORMAT(' ENTER FILENAME FOR 2ND SERIES ',$)
      READ(5,1001) FN2
      FN2(28)=0
      WRITE(5,1000)
 1000 FORMAT(' ENTER # OF LAGS,AND STEP ',$)
      READ(5,*) NLAG,NSTP
      SUM1=0.0
      SUM2=0.0
      SUMZ1=0.0
      SUMZ2=0.0
      DO 1 I=1,NLAG
      NCPL(I)=0
      D(I)=0.0
      G(I)=0.0
    1 CONTINUE
C
C OPEN INPUT AND OUTPUT FILES
C
      OPEN(UNIT=IN1,NAME=FN1,TYPE='OLD',FORM='FORMATTED')
      OPEN(UNIT=IN2,NAME=FN2,TYPE='OLD',FORM='FORMATTED')
      OPEN(UNIT=OUT,NAME='SY:GAMA.LST',TYPE='NEW')
C
C CALCULATE VARIOGRAM(GAMMA) VALUES
C
      MAX=(NLAG-1)*NSTP
      J=0
      DO 2 I=1,500
      READ(IN1,99,END=3000,ERR=2001) X1(I),Z1(I)
      J=J+1
    2 CONTINUE
 3000 NR1=J
      WRITE(5,600) NR1
      K=0
      DO 11 I=1,500
      READ(IN2,99,END=2000,ERR=2001) X2(I),Z2(I)
      K=K+1
   11 CONTINUE
 2000 NR2=K
      WRITE(5,600) NR2
      IF(NR1.EQ.NR2) GO TO 17
      DO 16 I=1,NR1
      IF(X1(I).EQ.X2(I)) GO TO 16
      DO 18 J=I,NR2-1
      Z2(J)=Z2(J+1)
      X2(J)=X2(J+1)
   18 CONTINUE
```



```
   16 CONTINUE
   17 NR=NR1
C
C ACCUMULATE SUMS OF EXPERIMENTAL VALUES
C
      DO 42 K42=1,NR
      SUM1=SUM1+Z1(K42)
      SUM2=SUM2+Z2(K42)
      SUMZ1=SUMZ1+Z1(K42)*Z1(K42)
      SUMZ2=SUMZ2+Z2(K42)*Z2(K42)
   42 CONTINUE
      N2=NR-1
      DO 51 K51=1,N2
      L=K51+1
      XI1=X1(K51)
      ZI1=Z1(K51)
      ZI2=Z2(K51)
         DO 61 K61=L,NR
      DX=X1(K61)-XI1
         H=IFIX(DX)
         IF(H.GT.MAX) GO TO 51
         L4=H/NSTP+0.5
         NCPL(L4)=NCPL(L4)+1
         D(L4)=D(L4)+DX
         DZ1=ZI1-Z1(K61)
      DZ2=ZI2-Z2(K61)
         G(L4)=G(L4)+0.5*DZ1*DZ2
   61    CONTINUE
   51 CONTINUE
      DO 45 K45=1,NLAG
      CP=FLOAT(NCPL(K45))
      IF(CP.EQ.0.) GO TO 45
      D(K45)=D(K45)/CP
      G(K45)=G(K45)/CP
   45 CONTINUE
C
C CALCULATE STATISTICS
C
      AVE1=SUM1/FLOAT(NR)
      AVE2=SUM2/FLOAT(NR)
      VAR1=SUMZ1/FLOAT(NR)-AVE1*AVE1
      VAR2=SUMZ2/FLOAT(NR)-AVE2*AVE2
      DEV1=SQRT(VAR1)
      DEV2=SQRT(VAR2)
      CV1=DEV1/AVE1
      CV2=DEV2/AVE2
C
C  OUTPUT:WRITING
C
      WRITE(6,100) FN1
      WRITE(6,200)NR,AVE1,VAR1,DEV1,CV1
```

```
      WRITE(6,100) FN2
      WRITE(6,200)NR,AVE2,VAR2,DEV2,CV2
      WRITE(6,300)
      WRITE(6,400)
      DO 30 I=1,NLAG
      IF(G(I).EQ.0.0)GO TO 30
      WRITE(6,500) I,NCPL(I),D(I),G(I)
   30 CONTINUE
      GO TO 15
 2001 WRITE(5,700)
C
C  FORMATS
C
   99 FORMAT(2(3X,F6.2))
  100 FORMAT(10X,28A1)
  200 FORMAT(//3X,'# OF SAMPLES=',I6/3X,'MEAN=',F11.4/3X,'VA
     1F11.4/3X,'STANDARD DEVIATION=',F11.4/3X,'COEFFICIENT O
     2IATION=',F11.4)
  300 FORMAT(//15X,'CROSS-VARIOGRAM : ONE DIMENSION'//)
  400 FORMAT(10X,1H 'STEP   # OF COUPLES   AVERAGE DISTANCE
  500 FORMAT(10X,I3,3X,I7,5X,F14.2,3X,F17.4)
  600 FORMAT(3X,' END OF FILE REACHED AFTER ',I4,' NUMBERS R
  700 FORMAT(3X,'ERROR DURING READING')
   15 CLOSE(UNIT=IN1)
      CLOSE(UNIT=IN2)
      CLOSE(UNIT=OUT)
      STOP
      END
```



```
      PROGRAM KRIGE
C  SIDNEY VIEIRA
C  SIMPLE KRIGING
C
C  THIS PROGRAM COMPUTES KRIGED VALUES AND ESTIMATION VARIAN
C  REQUIRES A MODEL OF VARIOGRAM AS INPUT (SEE THE PROGRAM C
C  THE VARIOGRAM).  'SLIDING NEIGHBORHOOD' MEANS THAT ONLY N
C  POINTS ARE TAKEN INTO ACCOUNT DURING THE CALCULATION OF T
C  THIS TECHIQUE IS APPROPRIATE FOR LARGE # OF DATA (OVER 10
C  PROGRAM CONSIDERS ALL NEIGHBORS WITHIN 10 METERS .  THE S
C  MINV  COMES FROM THE DIGITAL SYSTEM LIBRARY.
C  IT IS USED FOR INVERTING THE KRIGING MATRIX. ALTHOUGH ANY
C  INVERTING ROUTINE WORKS, THE ONES FOR SYMETRIC POSITIVE D
C  MATRICES ARE MORE EFFICIENT .
C
C  Z1( ) = EXPERIMENTAL VALUES
C  X1( ),Y1( ) = COORDINATES OF EXPERIMENTAL VALUES
C  Z2( ) = KRIGED VALUES
C  X2( ),Y2( ) = COORDINATES OF KRIGED POINTS
C  F( ) = KRIGING MATRIX(STORED COLUMNWISE).
C  G( ) = MATRIX OF LAMBDAS
C  B( ) = MATRIX OF RIGHT HAND SIDE OF KRIGING SYSTEM
C  SK( ) = KRIGING VARIANCES (OR ESTIMATION VARIANCES)
C  ISA( ) = STORAGE OF THE INDICES OF THE NEIGHBORING POINTS
C  NEV = # OF EXPERIMENTAL VALUES
C  NKP = # OF POINTS TO BE KRIGED
C  GAMMA(X) = FUNCTION, MODEL OF VARIOGRAM
C  A0,B0 = PARAMETERS OF THE VARIOGRAM MODEL
      GAMMA(X)=X/(A0+B0*X)
      DATA  HLM1,A0,B0,NEV,NKP/4,1.2113,0.1058,1280,1280/
      DIMENSION X1(1280),Y1(1280),Z1(1280),
     1F(900),G(30),B(30),ISA(30),A(8)
     INTEGER HLM1
C
C  INPUT
C
     OPEN (UNIT=1,NAME='ALLDATA.DAT',TYPE='OLD',ACCESS='SEQUENTI
     1  FORM='FORMATTED',CARRIAGECONTROL='NONE')
     OPEN (UNIT=4,NAME='SY:KRIGED.DAT',TYPE='NEW',FORM='FORMATTE
     OPEN (UNIT=6,NAME='SY:LISTOUT.LST',TYPE='NEW')
C
C  AUTOGENERATION OF X,Y COORDINATES
C
      DATA A/0.0,5.0,20.0,25.0,30.0,35.0,54.0,55.0/
      M=0
      DO 1 I=1,8
      K=0
      AX=A(I)
      IBEG=160*I-159
      IBEN=160*I
         DO 2 J=IBEG,IBEN
```

```
            K=K+1
          M=M+1
          X1(M)=AX
          Y1(M)=K
        2         CONTINUE
        1 CONTINUE
          WRITE(6,*) HLM1,AO,BO,NEV,NKP
          J=0
          DO 10 I=1,2000
         READ(1,599,END=2000,ERR=2001) Z1(I)
         J=J+1
       10  CONTINUE
     2000  NEV=J
          WRITE(5,3000) NEV
     3000  FORMAT(' END OF FILE REACHED AFTER ',I4,' NUMBERS REA
          SUM=0.
          SUM2=0.
C
C  COMPUTATION OF KRIGED POINTS
C
          WRITE(4,602)
          DO 80 K80=1,NKP
          X2T=X1(K80)
          Y2T=Y1(K80)
          Z1T=Z1(K80)
C
C  LOOK FOR ALL NEIGHBORS WITHIN HLM1 METERS.
C
          L1=0
          DO 82 K82=1,NEV
          IF(K82.EQ.K80) GO TO 82
          DX1=X2T-X1(K82)
          DY1=Y2T-Y1(K82)
          DHS=DX1*DX1+DY1*DY1
          HT=SQRT(DHS)
          IF(HT.GT.HLM1)GO TO 82
          L1=L1+1
          ISA(L1)=K82
       82 CONTINUE
           NEW=L1+1
           NONO=L1
          NTOT=NEW*NEW
          IF (Y2T.GT.5.0.AND.Y2T.LT.155.0) GO TO 12
C
C  START BUILDING THE KRIGING SYSTEM
C  ALL DIAGONAL TERMS ARE 0
C
          K=1
          DO 11 K1=1,NONO
          F(K)=0.
          K=K+NEW+1
```



```
       11 CONTINUE
          F(NTOT)=0.
C
C  LAST ROW AND LAST COLUMN TERMS ARE 1
C
          K=0
          DO 22 K2=1,NONO
          K=K+NEW
          F(K)=1.
       22 CONTINUE
          IN=NEW*NONO+1
          IEND=NTOT-1
          DO 23 K=IN,IEND
          F(K)=1.
       23 CONTINUE
C
C  COMPUTE MATRIX F
C
          DO 4 K4=2,NONO
          M=(K4-2)*NEW
          K=1
          DO 3 K3=K4,NONO
          IP=ISA(K3)
          IP1=ISA(K4-1)
          DX=X1(IP)-X1(IP1)
          DY=Y1(IP)-Y1(IP1)
          DHS=DX*DX+DY*DY
          X=SQRT(DHS)
          J=M+K3
          IN=J+K*NONO
          F(J)=GAMMA(X)
          F(IN)=F(J)
          K=K+1
        3 CONTINUE
        4 CONTINUE
C
C  INVERT F
C
          CALL MINV(F,NEW,D,L,M)
          IF (D) 48,46,48
       46 WRITE(6,50) D
       50 FORMAT(' MATRIX IS NOT INVERTIBLE '/5X,'DET[ F ]=',G11
          GO TO 15
       48 WRITE(6,51) D
       51 FORMAT(5X,' THE RESULTING DETERMINANT IS= ',G11.4)
C
C  COMPUTE THE KRIGED POINTS AND ESTIMATION VARIANCES
C  (SOLVE THE SYSTEM F*G=B)
C
C  COMPUTE B
C
```

```
      B(NEW)=1.
      DO 6 K6=1,NONO
      ISV=ISA(K6)
      DX=X1(ISV)-X2T
      DY=Y1(ISV)-Y2T
      DHS=DX*DX+DY*DY
      X=SQRT(DHS)
      B(K6)=GAMMA(X)
    6 CONTINUE
C
C  COMPUTE G (=INVERSE(F)*B)
C  (VECTOR OF UNKNOWNS-NONO LAMBDAS AND ONE MU)
C
      DO 7 K7=1,NEW
      GX=0.
      DO 8 K8=1,NEW
      I=K8+(K7-1)*NEW
    8 GX=GX+F(I)*B(K8)
    7 G(K7)=GX
C
C  COMPUTE KRIGED VALUES
C
   12 ZX=0.
      DO 9 K9=1,NONO
      ILS=ISA(K9)
    9 ZX=G(K9)*Z1(ILS)+ZX
C
C  COMPUTE ESTIMATION VARIANCES
C
      SKX=0.
      DO 111 K10=1,NEW
  111 SKX=SKX+G(K10)*B(K10)
      WRITE(5,604) K80,ZX,Z1T,SKX
      WRITE(4,604) K80,ZX,Z1T,SKX
      ANORM=(Z1(K80)-ZX)/SQRT(SKX)
      SUM=SUM+ANORM
      SUM2=SUM2+ANORM*ANORM
   80 CONTINUE
      AMEAN=SUM/FLOAT(NKP)
      VAR=SUM2/FLOAT(NKP)-AMEAN*AMEAN
      DEV=SQRT(VAR)
      WRITE(6,600) NEV,AMEAN,VAR,DEV
      GO TO 15
 2001 WRITE(5,1001)
 1001 FORMAT(' ERROR DURING READING ')
  599 FORMAT(F5.2)
  600 FORMAT(//5X,'NB. OF SAMPLES=',2X,I4/5X,'MEAN
     1/5X,'VARIANCE      =',G11.4/5X,'STD. DEVIATION=',G11.4
  602 FORMAT(' ',T10,'INDEX',T20,'KRIGED',T35,'MEASURED',T50
     1ION VARIANCE')
  604 FORMAT(' ',T14,I4,T20,F7.2,T35,F7.2,T50,G11.4)
```



```
   15 CLOSE(UNIT=1)
      CLOSE(UNIT=4)
      CLOSE(UNIT=6)
      STOP
      END
```

```
      PROGRAM COKRI
C
C
C  COKRIGING PROGRAM.
C
       DIMENSION XO(500),YO(500),ZO(500),X1(200),Y1(200),Z1(2
     1X2(50),Y2(50),Z2(50),C(700),R(30),RR(30),ISA1(8,16),IS
     2DIST1(8,16),DIST2(8,8),SK(500),NN1(20),NN2(10)
       BYTE F1(28),F2(28)
       INTEGER IN1,IN2,IN3,IN4,OUT
       DATA IN1/1/,IN2/2/,IN3/3/,IN4/4/,OUT/6/,PI/3.141592/
       COMMON PI,N1,N2,IDIM,DMAX1,DMAX2,DINC1,DINC2,LIM1,LIM2
     1CO,C1,A,IGAMA,IPLAN,IDIV,IDIR,DM1,DM2
C
C  OPEN FILE FOR INPUT PARAMETERS AND READ THEM IN
C
       OPEN(UNIT=IN3,NAME='SY:PARAM.DAT',TYPE='OLD',FORM='FOR
       READ(3,1001) F1
       READ(3,1001) F2
       READ(3,*) IDIM,IDIV,IPLAN,IGAMA,IDIR
       READ(3,*) MINY,MAXY,INCY,MINX,MAXX,INCX
       READ(3,*) DMAX1,DINC1,DMAX2,DINC2,LIM1,LIM2
       READ(3,*) CO,C1,A
 1001  FORMAT(28A1)
       F1(28)=0.
       F2(28)=0.
      WRITE(5,1001) F1
      WRITE(5,1001) F2
      WRITE(5,*) IDIM,IDIV,IPLAN,IGAMA,IDIR
      WRITE(5,*) MINY,MAXY,INCY,MINX,MAXX,INCX
      WRITE(5,*) DMAX1,DINC1,LIM1,DMAX2,DINC2,LIM2
      WRITE(5,*) CO,C1,A
       GOTO (800,820,821) IPLAN
  800  J=0
       DO 400 K=MINY,MAXY,INCY
       J=J+1
       YO(J)=FLOAT(K)
  400  CONTINUE
       NKP=J
       WRITE(5,*) NKP
       DO 500 I=1,NKP
  500  WRITE(5,5000) I,YO(I)
 5000  FORMAT(3X,I3,3X,F7.2)
       GO TO 302
  820  J=0
       DO 299 K=MINX,MAXX,INCX
       DO 301 L=MINY,MAXY,INCY
       J=J+1
       XO(J)=FLOAT(K)
       YO(J)=FLOAT(L)
  301  CONTINUE
```



```
  299  CONTINUE
       NKP=J
       WRITE(5,*) NKP
       DO 501 I=1,NKP
  501  WRITE(5,3005) I,XO(I),YO(I)
       GO TO 302
C
C  OPEN FILE FOR X,Y COORDINATES OF POINTS TO COKRIGE
C
  821  OPEN(UNIT=IN4,NAME='SY:GRID.DAT',TYPE='OLD',FORM='FORM
       J=0
       DO 322 I=1,500
       READ(IN4,3001,END=2005,ERR=2001) XO(I),YO(I)
       J=J+1
  322  CONTINUE
 2005  NKP=J
       CLOSE(UNIT=IN4)
C
C  OPEN INPUT FILES FOR VARIABLES 1 AND 2
C
  302  OPEN(UNIT=IN1,NAME=F1,TYPE='OLD',FORM='FORMATTED')
       OPEN(UNIT=IN2,NAME=F2,TYPE='OLD',FORM='FORMATTED')
       GOTO(803,802) IDIM
  802  J=0
       DO 303 I=1,500
       READ(1,3000,END=2000,ERR=2001) X1(I),Y1(I),Z1(I)
       J=J+1
  303  CONTINUE
 2000  N1=J
       WRITE(5,1005) N1
 1005  FORMAT(' END OF FILE REACHED AFTER ',I3,' NUMBERS READ
       J=0
       DO 304 I=1,500
       READ(2,3000,END=2500,ERR=2001) X2(I),Y2(I),Z2(I)
       J=J+1
  304  CONTINUE
 2500  N2=J
       WRITE(5,1005) N2
      WRITE(5,3000)(X1(I),Y1(I),Z1(I),I=1,N1)
      WRITE(5,3000)(X2(I),Y2(I),Z2(I),I=1,N2)
       GO TO 804
  803  J=0
       DO 305 I=1,500
       READ(1,3001,END=2002,ERR=2001) Y1(I),Z1(I)
       J=J+1
  305  CONTINUE
 2002  N1=J
       WRITE(5,1005) N1
       J=0
       DO 306 I=1,500
       READ(2,3001,END=2004,ERR=2001) Y2(I),Z2(I)
```

```
      J=J+1
  306 CONTINUE
 2004 N2=J
      WRITE(5,1005) N2
      DO 602 I=1,N1
  602 WRITE(5,3005) I,Y1(I),Z1(I)
      DO 603 J=1,N2
  603 WRITE(5,3005) J,Y2(J),Z2(J)
 3005 FORMAT(3X,I3,2(2X,F7.2))
C
C  CLOSE INPUT FILES
C
  804 CLOSE(UNIT=IN1)
      CLOSE(UNIT=IN2)
      CLOSE(UNIT=IN3)
C
C  BEGIN COKRIGING PROCESS
C
      DO 1 L1=1,NKP
      XOT=XO(L1)
      YOT=YO(L1)
      WRITE(5,6000) XOT,YOT
 6000 FORMAT(' NOW COKRIGING POINT ',F6.2,' , ',F6.2)
C
C  SELECT DESIRED NEEIGHBORHOOD OF ESTIMATION
C
      CALL NBHOOD(XOT,YOT,X1,Y1,X2,Y2,DIST1,DIST2,NN1,NN2,IS
      NB1=0
      NB2=0
      DO 604 I=1,IDIR
      NB1=NB1+NN1(I)
      NB2=NB2+NN2(I)
  604 CONTINUE
      WRITE(5,6001) NB1,DM1,NB2,DM2
 6001 FORMAT(' NBHOOD :'/4X,I3,' Z1 WITHIN ',F4.1/4X,I3,' Z2
     1F4.1)
      DO 605 I=1,IDIR
      N1B=NN1(I)
      WRITE(5,6003)
      DO 606 J=1,N1B
      IJ=ISA1(I,J)
      YW1=Y1(IJ)
      XW1=X1(IJ)
      WRITE(5,6002) IJ,XW1,YW1,DIST1(I,J)
  606 CONTINUE
      N2B=NN2(I)
      WRITE(5,6004)
      DO 607 K=1,N2B
      IK=ISA2(I,K)
      YW2=Y2(IK)
      XW2=X2(IK)
```



```
      WRITE(5,6002) IK,XW2,YW2,DIST2(I,K)
  607 CONTINUE
  605 CONTINUE
 6002 FORMAT(3X,I3,3(3X,F7.2))
 6003 FORMAT(' THE INDECES AND NEIGHBORS FOR Z1 ARE : ')
 6004 FORMAT(' THE INDICES AND NEIGHBORS FOR Z2 ARE : ')
C
C  BUILD COKRIGING SYSTEM
C
      CALL COKMS(X1,X2,Y1,Y2,NN1,NN2,ISA1,ISA2,DIST1,DIST2,R
      NN=NB1+NB2+2
      NT=NN*NN
      WRITE(5,6009)
 6009 FORMAT(' THE COEFFICIENT MATRIX IS :'/)
      DO 610 I=1,NN
      K=NT-NN+I
      WRITE(5,6010)(C(J),J=I,K,NN)
  610 CONTINUE
 6010 FORMAT(10(1X,F7.4))
      WRITE(5,6011)
 6011 FORMAT(' THE RIGHT HAND SIDE IS : ')
      DO 120 I=1,NN
      WRITE(5,6012) R(I)
  120 CONTINUE
 6012 FORMAT(2X,F6.2)
C
C  SOLVE COKRIGING SYSTEM [C]*[LAMBDAS]=[R]. ON SOLUTION
C  [LAMBDAS]=[INVERSE C]*[R], STORED ON THE VECTOR [R].
C
      CALL MINV(C,NN,D,L,M)
      IF(D) 200,46,200
   46 WRITE(5,80)
   80 FORMAT(' NO SOLUTION:SINGULAR MATRIX ')
      GO TO 399
C
C  COMPUTE COKRIGED VALUES
C
  200 WRITE(5,6013)
 6013 FORMAT(' THE INVERSE IS: ')
      DO 613 I=1,NN
      K=NT-NN+I
      WRITE(5,6010) (C(J),J=I,K,NN)
  613 CONTINUE
      DO 7 K7=1,NN
      GX=0.
      DO 8 K8=1,NN
      I=K8+(K7-1)*NN
    8 GX=GX+C(I)*R(I)
    7 RR(K7)=GX
      DO 19 I=1,IDIR
      N1B=NN1(I)
```

```
          DO 20 J=1,N1B
          KJ=ISA1(I,J)
          ZX=ZX+RR(J)*Z1(KJ)
       20 CONTINUE
          KK=NB1+1
          N2B=NN2(I)
          DO 21 L=1,N2B
          KL=ISA2(I,L)
          ZX=ZX+RR(KK)*Z2(KL)
          KK=KK+1
       21 CONTINUE
       19 CONTINUE
          Z0(L1)=ZX
C
C  COMPUTE THE ESTIMATION VARIANCES
C
          SKX=0.
          DO 22 I=1,NN
          SKX=SKX+R(I)*RR(I)
       22 CONTINUE
          SK(L1)=SKX
        1 CONTINUE
C
C  WRITE RESULTS TO DISC
C
C  OPEN OUTPUT FILE
C
          OPEN(UNIT=OUT,NAME='SY:COKRIGED.LST',TYPE='NEW')
          GOTO (807,808) IDIM
      807 DO 23 I=1,NKP
          WRITE(6,3002) I,Y0(I),Z0(I),SK(I)
       23 CONTINUE
          GO TO 300
      808 DO 24 I=1,NKP
          WRITE(6,3003) I,X0(I),Y0(I),Z0(I),SK(I)
       24 CONTINUE
          GO TO 300
     2001 WRITE(5,1014)
     1014 FORMAT(' ERROR DURING READING ')
C
C  FORMATS
C
     3000 FORMAT(3(2X,F7.2))
     3001 FORMAT(2(3X,F6.2))
     3002 FORMAT(3X,I3,2(3X,F7.2),3X,G13.4)
     3003 FORMAT(3X,I3,3(3X,F7.2),3X,G13.4)
C
C  CLOSE OUTPUT FILE
C
      300 CLOSE(UNIT=OUT)
      399 STOP
```



```
          END
```

```
      program corsp
c
c
c     program written by Sidney Vieira
c
c
c     function:
c          Calculates the autocorrelation function for an e
c          spaced series and then calculates the spectral d
c          and frequency of possible periods on the series.
c          of the calculations are found in the subroutines
c          by this program.
c
c
c     Variables:
c          NR     Number of measurements
c           Z     Variable
c           M     Length of autocorrelation
c           R     Autocorrelation
c          SP     Spectral density
c          FR     Frequency
c          SM     Smooth spectral density
c          MW     Number of terms in the series
c
c
c     Subroutines required:
c         Auto    Autocorrelation function (IBM SSP functio
c         Spect   Spectral density (Written by M. Vauclin)
c
c
c     Observations:
c         1) The series Z must be equally spaced by "deflix
c         2) The length of the autocorrelation "M" must be
c            than "NR", the number of observations
c
c
c.....................................................
c
c
c
c     Initialization
c
      dimension z(200),r(100),sp(100),fr(100),sm(100)
      byte fn(28)
      integer in,out
      data in/1/,out/6/
c
c
c     Output
c
      write(5,1000)
```



```
 1000 format(' enter filename ',$)
      read(5,999) fn
 999  format(28a1)
      fn(28)=0
      write(5,1002)
 1002 format(' enter number of lags "l" and spacing "deltx"
      read(5,*) l,deltx
      open(unit=im,name=fn,type='old')
      j=0
      do 1 i=1,200
      read(1,*,end=2000,err=2001) xu,z(i)
      j=j+1
 1    continue
 2000 nr=j
      write(5,1001) nr
 1001 format(' eop after ',i3,' records read ')
      close(unit=in)
c
c     calculate autocorrelation
c
      call auto(z,nr,l,r)
c
c     calculate number of terms on the series
c
      l=l-1
      do 5 i=1,l
      if (r(i).gt.0.0) go to 5
      mw=i-1
      go to 6
 5    continue
      mw=i-1
c
c     calculate spectral density
c
 6    call spect(r,l,deltx,mw,sp,fr,sm)
c
c     output
c
      open(unit=out,name='sy:cor.lst',type='new')
      write(6,1006) fn,l,deltx,mw
 1006 format(' autocorrelation for ',28a1/1x,'number of lag
     1 i3/1x,'spacing (deltx)= ',f4.2/1x,'number of terms= '
      do 2 i=1,l
      write(6,*) i,r(i),fr(i),sp(i),sm(i)
 2    continue
      close(unit=out)
      go to 15
 2001 write(5,1005)
 1005 format(' error during reading ')
 15   stop
      end
```

```
      SUBROUTINE NBHOOD(XOT,YOT,X1,Y1,X2,Y2,DIST1,DIST2,NN1,
     1NN2,ISA1,ISA2)
C
C  SELECTS NEIGHBORHOOD AROUND THE LOCATION OF THE VALUE TO
C  COKRIGED
C
      DIMENSION X1(200),Y1(200),X2(50),Y2(50),DIST1(8,16),
     1DIST2(8,8),ISA1(8,16),ISA2(8,8),NN1(20),NN2(10)
      COMMON PI,N1,N2,IDIM,DMAX1,DMAX2,DINC1,DINC2,LIM1,LIM2
     1CO,C1,A,IGAMA,IPLAN,IDIV,IDIR,DM1,DM2
      DM1=DMAX1
      DM2=DMAX2
      GOTO (800,801,802,803) IDIV
  801 T1=COS(PI/2.)
      AMULT=4.
      GO TO 9
  802 T1=COS(PI/4.)
      AMULT=2.
      GO TO 9
  803 T1=COS(PI/8.)
      AMULT=1.
    9 DO 2 L2=1,IDIR
      PHI=(2.*FLOAT(L2)-1.0)*AMULT*PI/8.
      CA=COS(PHI)
      SA=SIN(PHI)
      NN2(L2)=0
   10 NN1(L2)=0
      DO 3 L3=1,N1
      DY1=Y1(L3)-YOT
      DX1=X1(L3)-XOT
      D1=SQRT(DY1*DY1+DX1*DX1)
      IF (D1.GT.DM1.OR.D1.LT.1.0E-6) GO TO 3
      CC1=DY1*SA/D1+DX1*CA/D1
      IF (CC1.LT.T1) GO TO 3
      IF (L2.EQ.1) GO TO 50
      IF (L3.EQ.ISA1(L2-1,NT1)) GO TO 3
   50 NN1(L2)=NN1(L2)+1
      NT1=NN1(L2)
      ISA1(L2,NT1)=L3
      DIST(L2,NT1)=D1
    3 CONTINUE
      IF (NN2(L2).GE.LIM2) GO TO 99
   11 NN2(L2)=0
      DO 4 L4=1,N2
      DY2=Y2(L4)-YOT
      DX2=X2(L4)-XOT
      D2=SQRT(DY2*DY2+DX2*DX2)
      IF (D2.GT.DM2.OR.D2.LT.1.0E-6) GO TO 4
      CC2=DY2*SA/D2+DX2*CA/D2
      IF (CC2.LT.T1) GO TO 4
      IF (L2.EQ.1) GO TO 51
```



```
      IF (L4.EQ.ISA2(L2-1,NT2)) GO TO 4
   51 NN2(L2)=NN2(L2)+1
      NT2=NN2(L2)
      ISA2(L2,NT2)=L4
      DIST2(L2,NT2)=D2
    4 CONTINUE
   99 IF (NN1(L2)-LIM1) 100,101,102
  100 DM1=DM1+DINC1
      GO TO 10
  101 IF (NN2(L2)-LIM2) 103,2,104
  103 DM2=DM2+DINC2
      GO TO 11
  102 LDIR=L2
      CALL SORT(DIST1,NN1,ISA1,LDIR,LIM1)
      GO TO 101
  104 LDIR=L2
      CALL SORT(DIST2,NN2,ISA2,LDIR,LIM2)
    2 CONTINUE
      RETURN
  800 NN2(1)=0
   12 NN1(1)=0
      DO 6 L6=1,N1
      DY1=ABS(Y1(L6)-YOT)
      IF (DY1.GT.DM1.OR.DY1.LT.1.0E-6) GO TO 6
      NN1(1)=NN1(1)+1
      NT1=NN1(1)
      ISA1(1,NT1)=L6
      DIST1(1,NT1)=DY1
    6 CONTINUE
      IF (NN2(1).GE.LIM2) GO TO 199
   13 NN2(1)=0
      DO 7 L7=1,N2
      DY2=ABS(Y2(L7)-YOT)
      IF (DY2.GT.DM2.OR.DY2.LT.1.0E-6) GO TO 7
      NN2(1)=NN2(1)+1
      NT2=NN2(1)
      ISA2(1,NT2)=L7
      DIST2(1,NT2)=DY2
    7 CONTINUE
  199 IF (NN1(1)-LIM1) 200,201,202
  200 DM1=DM1+DINC1
      GO TO 12
  201 IF (NN2(1)-LIM2) 203,5,204
  203 DM2=DM2+DINC2
      GO TO 13
  202 LDIR=1
      CALL SORT(DIST1,NN1,ISA1,LDIR,LIM1)
      GO TO 201
  204 LDIR=1
      CALL SORT(DIST2,NN2,ISA2,LDIR,LIM2)
    5 CONTINUE
```

Builds the Kriging System Page 1

```
      subroutine krix(x,y,n,is,dis,b,f)
c
c
c     function:
c
c     builds kriging system:coefficient matrix [f], and
c     right hand side [r]. matrices can be filled with
c     semivariances (iv=1) or covariances (iv=2).
c
c
c     variables:
c
c     x,y    coordinates of neighbors used on one estim-
c            tion.
c     n      number of neighbors.
c     is     array index of neighbors.
c     dis    distances from neighbors to kriged value .
c     nt     number of equations and unknowns.
c     f      coefficient matrix.
c     r      right hand side.
c
c
c     commons:
c
c     c0,c1,range,sill,igama,iv - parameters for structu-
c     ral model (semivariogram or covariogram).
c
c
c     subroutine required : cov
c
      dimension x(1),y(1),is(1),dis(1),b(1),f(1)
      common co,c1,range,sill,igama,iv
c
c     fill in matrix b- the right hand side
c
      nt=n+1
      ntot=nt*nt
      do 1 i=1,n
      xd=dis(i)
      b(i)=cov(xd,c0,c1,range,sill,igama,iv)
    1 continue
c
c     weights sum to 1.0 .
c
      b(nt)=1.0
c
c     initialize matrix f with 1's
c
      do 2 i=1,ntot-1
      f(i)=1.0
    2 continue
```

Selects Estimation Neighborhood Page 3

```
      RETURN
      END
```

```
      subroutine cokms(x1,x2,y1,y2,nb1,nb2,isa1,isa2,dist1,
     1dist2,r,c)
c
c  builds the cokriging system: [c]* [lambdas]=[r]
c
      dimension x1(1),y1(1),x2(1),y2(1),dist1(1),dist2(1),
     1isa1(1),isa2(1),c(1),r(1)
      common c011,c111,a11,sil11,iga11,c012,c112,a12,sil12,i
     1c022,c122,a22,sil22,iga22,iv
c
c  initialize matrix c
c
      nn=nb1+nb2+2
      nt=nn*nn
      do 5 i=1,nt
    5 c(i)=0.0
c
c  fill in the ones in matrix c
c
      nint=nn*(nn-2)+1
      nend=nint+nb1-1
      do 6 i=nint,nend
    6 c(i)=1.0
      nint=nt-nb2-1
      nend=nt-2
      do 7 i=nint,nend
    7 c(i)=1.0
      k=nn-1
      do 8 i=1,nb1
      c(k)=1.0
      k=k+nn
    8 continue
      k=nn*(nb1+1)
      do 9 i=1,nb2
      c(k)=1.0
      k=k+nn
    9 continue
c
c  fill in the covas in matrix c
c
c
c       upper left nb1 by nb1( covariance function 1-1 )
c
        do 13 j=1,nb1
        kk=j
        kj=isa1(j)
        xt=x1(kj)
        yt=y1(kj)
        do 14 l=1,nb1
        kl=isa1(l)
        dx=x1(kl)-xt
```



```
c
c      very last entry is zero
c
      f(ntot)=0.0
c
c      fill the remaining entries with semivariances
c      (iv=1) or covariances (iv=2).
c
c      upper triangular part.
c
      do 3 i=1,n
      ip1=is(i)
      xt=x(ip1)
      yt=y(ip1)
      do 4 J=i,n
      ip=is(j)
      dx=x(ip)-xt
      dy=y(ip)-yt
      xd=sqrt(dx*dx+dy*dy)
      k=i+(j-1)*nt
      f(k)=cov(xd,co,c1,range,sill,igama,iv)
    4 continue
    3 continue
c
c      matrix is symetrical: lower triangular part is
c      equal to upper triangular part.
c
      do 5 i=2,n
      k=(i-1)*(nt+1)
      do 6 j=i,n
      m=j+(i-2)*nt
      f(m)=f(k)
      k=k+nt
    6 continue
    5 continue
      return
      end
```

```
            dy=y1(kl)-yt
            x=sqrt(dx**dx+dy**dy)
            c(kk)=cov(x,c011,c111,a11,sil11,iga11,iv)
            kk=kk+nn
     14     continue
     13     continue
      c
      c     upper right nb1 by nb2( covariance function 2-1 )
      c
            do 15 j=1,nb2
            kk=nn*(nb1+j-1)
            kj=isa2(j)
            xt=x2(kj)
            yt=y2(kj)
            do 16 l=1,nb1
            kl=isa1(l)
            dx=xt-x1(kl)
            dy=yt-y1(kl)
            x=sqrt(dx**dx+dy**dy)
            kk=kk+1
            c(kk)=cov(x,c012,c112,a12,sil12,iga12,iv)
     16     continue
     15     continue
      c
      c     lower left nb2 by nb1( covariance function 1-2 )
      c
            do 17 j=1,nb2
            kj=isa2(j)
            xt=x2(kj)
            yt=y2(kj)
            kk=nb1+j
            do 18 l=1,nb1
            kl=isa1(l)
            dx=x1(kl)-xt
            dy=y1(kl)-yt
            x=sqrt(dx**dx+dy**dy)
            c(kk)=cov(x,c012,c112,a12,sil12,iga12,iv)
            kk=kk+nn
     18     continue
     17     continue
      c
      c     lower right nb2 by nb2( covariance function 2-2 )
      c
            do 19 j=1,nb2
            kk=nb1*(nn+1)+J
            kj=isa2(j)
            xt=x2(kj)
            yt=y2(kj)
            do 20 l=1,nb2
            kl=isa2(l)
            dx=xt-x2(kl)
```



```
            dy=yt-y2(kl)
            x=sqrt(dx**dx+dy**dy)
            c(kk)=cov(x,c022,c122,a22,sil22,iga22,iv)
            kk=kk+nn
     20     continue
     19     continue
      c
      c  fill in matrix r
      c
      c       upper nb1( covariance function 2-1 )
      c
          do 21 j=1,nb1
          x=dist1(j)
          r(j)=cov(x,c012,c112,a12,sil12,iga12,iv)
     21   continue
      c
      c       lower nb2 ( covariance function 2-2 )
      c
          do 22 k=1,nb2
          x=dist2(k)
          kk=nb1+k
          r(kk)=cov(x,c022,c122,a22,sil22,iga22,iv)
     22   continue
      c
      c       weights on z2 sum to 1.0
      c
          r(nn)=1.0
      c
      c       weights on z1 sum to 0.0
      c
          r(nn-1)=0.0
          return
          end
```

```
C
C     ..................................................
C
C        SUBROUTINE SIMQ
C
C        PURPOSE
C           OBTAIN SOLUTION OF A SET OF SIMULTANEOUS LINEAR
C           AX=B
C
C        USAGE
C           CALL SIMQ(A,B,N,KS)
C
C        DESCRIPTION OF PARAMETERS
C           A - MATRIX OF COEFFICIENTS STORED COLUMNWISE.  T
C               DESTROYED IN THE COMPUTATION.  THE SIZE OF M
C               N BY N.
C           B - VECTOR OF ORIGINAL CONSTANTS (LENGTH N). THE
C               REPLACED BY FINAL SOLUTION VALUES, VECTOR X.
C           N - NUMBER OF EQUATIONS AND VARIABLES. N MUST BE
C           KS - OUTPUT DIGIT
C                0 FOR A NORMAL SOLUTION
C                1 FOR A SINGULAR SET OF EQUATIONS
C
C        REMARKS
C           MATRIX A MUST BE GENERAL.
C           IF MATRIX IS SINGULAR , SOLUTION VALUES ARE MEAN
C           AN ALTERNATIVE SOLUTION MAY BE OBTAINED BY USING
C           INVERSION (MINV) AND MATRIX PRODUCT (GMPRD).
C
C        SUBROUTINES AND FUNCTION SUBPROGRAMS REQUIRED
C           NONE
C
C        METHOD
C           METHOD OF SOLUTION IS BY ELIMINATION USING LARGE
C           DIVISOR. EACH STAGE OF ELIMINATION CONSISTS OF I
C           ROWS WHEN NECESSARY TO AVOID DIVISION BY ZERO OR
C           ELEMENTS.
C           THE FORWARD SOLUTION TO OBTAIN VARIABLE N IS DON
C           N STAGES. THE BACK SOLUTION FOR THE OTHER VARIAB
C           CALCULATED BY SUCCESSIVE SUBSTITUTIONS. FINAL SO
C           VALUES ARE DEVELOPED IN VECTOR B, WITH VARIABLE
C           VARIABLE 2 IN B(2),........., VARIABLE N IN B(N).
C           IF NO PIVOT CAN BE FOUND EXCEEDING A TOLERANCE O
C           THE MATRIX IS CONSIDERED SINGULAR AND KS IS SET
C           TOLERANCE CAN BE MODIFIED BY REPLACING THE FIRST
C
C     ..................................................
C
      SUBROUTINE SIMQ(A,B,N,KS)
      DIMENSION A(1),B(1)
C
```



```
C        FORWARD SOLUTION
C
      TOL=0.0
      KS=0
      JJ=-N
      DO 65 J=1,N
      JY=J+1
      JJ=JJ+N+1
      BIGA=0
      IT=JJ-J
      DO 30 I=J,N
C
C        SEARCH FOR MAXIMUM COEFFICIENT IN COLUMN
C
      IJ=IT+I
      IF(ABS(BIGA)-ABS(A(IJ))) 20,30,30
   20 BIGA=A(IJ)
      IMAX=I
   30 CONTINUE
C
C        TEST FOR PIVOT LESS THAN TOLERANCE (SINGULAR MATRIX
C
      IF(ABS(BIGA)-TOL) 35,35,40
   35 KS=1
      RETURN
C
C        INTERCHANGE ROWS IF NECESSARY
C
   40 I1=J+N*(J-2)
      IT=IMAX-J
      DO 50 K=J,N
      I1=I1+N
      I2=I1+IT
      SAVE=A(I1)
      A(I1)=A(I2)
      A(I2)=SAVE
C
C        DIVIDE EQUATION BY LEADING COEFFICIENT
C
   50 A(I1)=A(I1)/BIGA
      SAVE=B(IMAX)
      B(IMAX)=B(J)
      B(J)=SAVE/BIGA
C
C        ELIMINATE NEXT VARIABLE
C
      IF(J-N) 55,70,55
   55 IQS=N*(J-1)
      DO 65 IX=JY,N
      IXJ=IQS+IX
      IT=J-IX
```

Calculates the Jackknifing Parameters Page 1

```
      subroutine reg(ifile,n,nb,c0,inb,dm,z1,z2,sk)
c
c.....................................................
c
c     function:
c
c     calculates the jack-knifing parameters and writes th
c     to a disk file called 'syo:jack.lst'.
c
c
c     input variables:
c
c     z1      measured values
c     z2      estimated values
c     sk      estimation variance
c
c
c     input constants:
c
c     ifile   logical file unit number
c     n       number of values z1,z2,sk
c     nb      number of neighbors (inb=2) used on estimati
c     c0      nugget effect
c     inb     neighborhood option
c     dm      distance within which neighbors are (inb=1)
c
c
c     output :
c
c     ave     mean values
c     var     variances
c     cov     covariance (z1,z2)
c     a,b,r12 regression parameters
c     amer    mean for absolute errors
c     vaer    variance of absolute errors
c     arer    mean for reduced errors
c     vrer    variance for reduced errors
c
c
c
c.....................................................
c
c
      dimension z1(1),z2(1),sk(1)
      ser1=0.0
      ser2=0.0
      ser3=0.0
      ser4=0.0
      do 57 i=1,n
      aber=z2(i)-z1(i)
      ser1=ser1+aber
```

Solves a Linear System of Equations Page 3

```
      DO 60 JX=JY,N
      IXJX=N*(JX-1)+IX
      JJX=IXJX+IT
   60 A(IXJX)=A(IXJX)-(A(IXJ)*A(JJX))
   65 B(IX)=B(IX)-(B(J)*A(IXJ))
C
C        BACK SOLUTION
C
   70 NY=N-1
      IT=N*N
      DO 80 J=1,NY
      IA=IT-J
      IB=N-J
      IC=N
      DO 80 K=1,J
      B(IB)=B(IB)-A(IA)*B(IC)
      IA=IA-N
   80 IC=IC-1
      RETURN
      END
```

```
          ser2=ser2+aber*aber
          if (sk(i).le.0.0) go to 57
          rder=aber/(sqrt(sk(i)))
          ser3=ser3+rder
          ser4=ser4+rder*rder
   57     continue
          an=float(n)
          amer=ser1/an
          vaer=ser2/an-amer*amer
          arer=ser3/an
          vrer=ser4/an-arer*arer
          s1=0.0
          s2=0.0
          s3=0.0
          s4=0.0
          s5=0.0
          do 5 i=1,n
          s1=s1+z2(i)
          s2=s2+z1(i)
          s3=s3+z2(i)*z2(i)
          s4=s4+z1(i)*z1(i)
          s5=s5+z1(i)*z2(i)
    5     continue
          ave1=s1/an
          ave2=s2/an
          var1=s3/an-ave1*ave1
          var2=s4/an-ave2*ave2
          cov12=s5/an-ave1*ave2
          b=(an*s5-s2-s1)/an*s3-s1*s1)
          a=(s2-b*s1)/an
          r12=cov12/(sqrt(var1)*sqrt(var2))
          open(unit=ifile,name='sy0:jack.lst',type='new')
          write(ifile,103) ave1,var1
          write(ifile,104) ave2,var2
          write(ifile,105) cov12
          write(ifile,106) a,b,r12
          write(ifile,107) c0
          write(ifile,108) amer,vaer
          write(ifile,109) arer,vrer
          goto (7,6), inb
    6     write(ifile,110) nb
          close(unit=ifile)
          return
    7     write(ifile,111) dm
          close(unit=ifile)
          return
  103     format(1x,'kriged',/4x,'mean     =',g11.4,/4x,
         1  'variance=',g11.4)
  104     format(//1x,'measured',/4x,'mean     =',g11.4,/4x,
         1  'variance=',g11.4)
  105     format(//1x,'the covariance is: ',g11.4)
```



```
  106     format(//1x,'the regression coefficients for "y=a+b*
         1  'a= ',g11.4,/4x,'b= ',g11.4,/4x,'r= ',f7.4)
  107     format(//1x,'the nugget effect value is :',g11.4)
  108     format(//1x,' absolute error  :'/4x,'mean    =',g11.
         1  'variances',g11.4)
  109     format(//1x,' reduced error   :'/4x,'mean    =',g11.
         1  'variances',g11.4)
  110     format(/1x,' the number of neighbors is :',i2)
  111     format(/1x,' all neighbors within ',f4.1)
          end
```

```
      function cov(x,c0,c1,range,sill,igama,iv)
c
c
c*************************************************************
c
c     Function:
c
c     computes the semivariogram (iv=1) or covariogram (iv=
c     for distances x.  Models available are: 1) linear; 2)
c     3) spherical; 4) exponential; 5) gaussian; and 6) hol
c     selection is transferred on igama.
c
c     Variables:
c
c     x       distance
c     c0      nugget effect
c     c1      type of model
c     range   range of variogram
c     sill    sill (plateau) of variogram
c     igama   choice of the model for the data
c     iv      choice of the structural tool
c
c
c*************************************************************
c
      if(x.gt.0.0) go to 40
      go to (41,42), iv
   41 cov=0.0
      return
   42 cov=sill
      return
   40 if(igama.gt.3) go to 20
      if(x.lt.range) go to 20
      go to (30,31), iv
   30 cov=sill
      return
   31 cov=0.0
      return
   20 go to (1,2,3,4,5,6), igama
c
c
c     linear model
c
    1 cov=c0+c1*x
      go to (50,51), iv
c
c     power model
c
    2 cov=c0*x**c1
      go to (50,51), iv
c
```



```
c     spherical model
c
    3 cov=c0+c1*(1.5*x/range-0.5*(x/range)**3)
      go to (50,51), iv
c
c     exponential model
c
    4 cov=c0+c1*(1-exp(-x/range))
      go to (50,51), iv
c
c     gaussian model
c
    5 x=(x/range)**2
      cov=c0+c1*(1-exp(-x))
      go to (50,51), iv
c
c     hole effect model
c
    6 x=x/range
      cov=c0+c1*(1-(sin(x)/x))
      if(iv.eq.1) return
   51 cov=sill-cov
   50 return
      end
```

```
      SUBROUTINE SORT(DIST,NUM,ISA,L,LIM)
C
C  SORTS OUT DISTANCES FROM EACH POINT WITHIN THE NEIGHBOR
C  OF ESTIMATION TO THE POINT TO BE COKRIGED IN INCREASING O
C  DER.INDICES ASSOCIATED WITH LOCATION OF EACH NEIGHBOR   A
C  ORDERED ALONG WITH DISTANCES.
C
      DIMENSION DIST(1,1),NUM(1),ISA(1,1)
      WRITE(5,1000) NUM(L)
      DO 1 I=1,NUM(L)
    1 WRITE(5,*) I,ISA(L,I),DIST(L,I)
 1000       FORMAT(I2,' NEIGHBORS NOT SORTED ')
      NN=NUM(L)-1
      DO 100 J=1,NN
C
C  FIND LOCATION LL OF SMALLEST DIST.
C
      LL=J
      JJ=J+1
      DO 200 I=JJ,NUM(L)
      IF (DIST(L,LL).LT.DIST(L,I)) GO TO 200
      LL=I
  200 CONTINUE
C
C  INTERCHANGE DIST(L,LL),ISA(L,LL) WITH DIST(L,J),ISA(L,J),
C  PECTIVELLY.
C
      T=DIST(L,LL)
      U=ISA(L,LL)
      DIST(L,LL)=DIST(L,J)
      ISA(L,LL)=ISA(L,J)
      DIST(L,J)=T
      ISA(L,J)=U
  100 CONTINUE
      WRITE(5,1001) NUM(L)
 1001       FORMAT(I2,' SORTED ')
      DO 2 I=1,NUM(L)
    2 WRITE(5,*) I,ISA(L,I),DIST(L,I)
      NUM(L)=LIM
      RETURN
      END
```



```
C
C     ..................................................
C
C        SUBROUTINE AUTO
C
C        PURPOSE
C           TO FIND AUTOCOVARIANCES OF SERIES A FOR LAGS 0 T
C
C        USAGE
C           CALL AUTO (A,N,L,R)
C
C        DESCRIPTION OF PARAMETERS
C           A    - INPUT VECTOR OF LENGTH N CONTAINING THE T
C                  WHOSE AUTOCOVARIANCE IS DESIRED.
C           N    - LENGTH OF THE VECTOR A.
C           L    - AUTOCOVARIANCE IS CALCULATED FOR LAGS OF
C                  L-1.
C           R    - OUTPUT VECTOR OF LENGTH L CONTAINING AUTO
C                  OF SERIES A.
C
C        REMARKS
C           THE LENGTH OF R IS DIFFERENT FROM THE LENGTH OF
C           BE GREATER THAN L.  IF NOT, R(1) IS SET TO ZERO
C           IS MADE TO THE CALLING PROGRAM.
C
C        SUBROUTINES AND FUNCTION SUBPROGRAMS REQUIRED
C           NONE
C
C        METHOD
C           DESCRIBED IN R.B. BLACKMAN AND J.W. TUKEY, 'THE
C        OF POWER SPECTRA', DOVER PUBLICATIONS INC., NEW YOR
C
C     ..................................................
C
      SUBROUTINE AUTO (A,N,L,R)
      DIMENSION A(1),R(1)
C
C     CALCULATE AVERAGE OF TIME SERIES A
C
      AVER=0.0
      IF(N-L) 50,50,100
   50 R(1)=0.0
      RETURN
  100 DO 110 I=1,N
  110 AVER=AVER+A(I)
      FN=N
      AVER=AVER/FN
C
C     CALCULATE AUTOCOVARIANCES
C
      DO 130 J=1,L
```

```
      subroutine spect(r,m,deltx,mw,s,f,smth)
c
c
c*************************************************************
c
c       written by M. Vauclin
c
c      function:
c
c      calculates spectral density function- fourier transfo
c      of autocorrelation.  The spectral density is calculat
c      using "Bartletts" window and the frequency is calcula
c      using the "Nyquist" frequency.  A smooth density func
c      is also calculated using the "Hanning" procedure.
c
c     Variables:
c
c      r        autocorrelation
c      m        number of lags on the autocorrelation
c      deltx    spacing between measurements
c      mw       number of terms in the series
c      s        spectral density
c      f        frequency
c      smth     smooth spectral density from the "Hanning" pro
c                                                            '
c
c*************************************************************
c
c
      dimension r(1),s(1),f(1),smth(1)
      om=2.*3.14159*deltx
      fn=0.5/deltx
c
c     calculate the spectral density function
c
      sdum=0.0
      do 1 k=2,mw
    1 sdum=sdum+r(k)*(1.-(k-1)/float(mw))
      s(1)=deltx*(r(1)+2.*sdum)
      do 3 j=2,m+1
      f(j)=(j-1)*fn/m
      sdum=0.0
      do 2 k=2,mw
    2 sdum=sdum+r(k)*cos(om*k*f(j))*(1.-(k-1)/float(mw))
    3 s(j)=deltx*(r(1)+2.*sdum)
c
c     calculate smooth spectral density function
c
      smth(1)=0.5*(s(1)+s(2))
      smth(m+1)=0.5*(s(m)+s(m+1))
      do 4 j=2,m
```



```
      NJ=N-J+1
      SUM=0.0
      DO 120 I=1,NJ
      IJ=I+J-1
  120 SUM=SUM+(A(I)-AVER)*(A(IJ)-AVER)
      FNJ=NJ
  130 R(J)=SUM/FNJ
      RETURN
      END
```

```
      4 smth(j)=0.25*s(j-1)+0.5*s(j)+0.25*s(j+1)
        smt=0.0
        do 5 j=2,m
      5 smt=smt+smth(j)
        smt=smt+0.5*(smth(1)+smth(m+1))
        return
        end
```

REFERENCES

Ball, D. F., and W. M. Williams. 1968. Variability of soil chemical properties in two uncultivated brown earths. J. Soil Sci. 19:379-391.

Bartlett, M. S. 1966. Stochastic Processes. Cambridge Univ. Press, 2nd Ed.

Beckett, P. H. T., and R. Webster. 1971. Soil variability: a review. Soils Fert. 34:1-15.

Blais, R. A., and P. A. Carlier. 1968. Applications of geostatistics in ore evaluation. Canadian Institute of Mining and Metalurgy, Ore Research Estimation and Grade Control, Special Volume 9, Montreal, Quebec. p. 41-48.

Burgess, T. M., and R. Webster. 1980a. Optimal interpolation and isarithmic mapping of soil properties. I. The semivariogram and punctual kriging. J. Soil Sci. 31:315-331.

Burgess, T. M., and R. Webster. 1980b. Optimal interpolation and isoarithmic mapping of soil properties. II. Block kriging. J. Soil Sci. 31:333-341.

Campbell, J. B. 1978. Spatial variation of sand content and pH within single contiguous delineation of two soil mapping units. Soil Sci. Soc. Am. J. 42:460-464.

David, M. 1970. The geostatistical estimation of porphyry-type deposits and scale factor problems. Proc. Pribram Science and Technique M5: 91-109.

David, M. 1977. Geostatistical ore reserve estimation. Elsevier Scientific Publ. Co. 364 pp.

Delhomme, J. P. 1976. Kriging in hydrosciences. Centre D'Informatique Geologique, Fontainebleau, France.

Grossman, R. B. 1978. Partial Soil Inventory of Agronomy Research Center. Boone County, Missouri. 30 p.

Hajrasuliha, S., N. Baniabbassi, J. Metthey and D. R. Nielsen. 1980. Spatial variability of soil sampling for salinity studies in southwest Iran. Irrigation Sci. 1:197-208.

Harradine, F. F. 1949. The variability of soil properties in relation to stage of profile development. Soil Sci. Soc. Am. Proc. 14:302-311.

Harris, J. A. 1920. Practical universality of field heterogeneity as a factor influencing plot yields. J. Agr. Res. XIX(7).

Journel, A. G. 1974. Geostatistics for conditional simulation of ore bodies. Eco. Geology 69:673-687.

Journel, A. G., and Ch. J. Huijbregts. 1978. Mining Geostatistics. Academic Press, London.

Krige, D. G. 1951. A statistical approach to some basic mine evaluation problems on the witwatersrand. J. Chem. Metall. Min. Soc. S. Afi. 52: 119-139.

Matheron, G. 1963. Principles of geostatistics. Econ. Geology 58:1246-1266.

Matheron, G. 1971. The Theory of Regionalized Variables and Its Application. Les Cahiers du Centre de Morphologie Mathematique, Fas. 5, C. G. Fontainebleau.

Montgomery, E. G. 1913. Experiments in wheat breeding: Experimental error in the nursery and variation in nitrogen and yield. U.S. Dept. Agri. Bur. Plant Indust. Bul. 269. 61 p.

Olea, R. A. 1975. Optimum mapping techniques using regionalized variable theory. Kansas Geol. Survey, Series on Spatial Analysis No. 2, Univ. Kansas, Lawrence, Kansas. 137 p.

Olea, R. A. 1977. Measuring spatial dependence with semivariograms. Kansas Geol. Survey, Series on Spatial Analysis No. 2, Univ. Kansas, Lawrence, Kansas.

Pendleton, R. L. 1919. Are soils mapped under a given type name by the Bureau of Soils method closely similar to one another? Univ. Calif. Publ. Agric. Sci. 3:369-498.

Robinson, G. W., and W. E. Lloyd. 1915. On the probable error of sampling in soil surveys. J. Agric. Sci. 7:144-153.

Smith, L. H. 1910. Plot arrangement for variety experiment with corn. Proc. Amer. Soc. Agron. V. I. 1907/09. p. 84-89.

Snedecor, G. W., and W. G. Cochran. 1967. Statistical Methods. 6th Ed., Iowa State Univ. Press, Ames. 593 p.

Ugarte, A. 1972. Ejemploes de modelor de estimacion a corto e largo plans. Bol. Geostadist. No. 4:3-22.

Vieira, S. R., and J. L. Hatfield. 1982. Temporal variability of air temperature and remotely sensed surface temperature for bare soil. Inter. J. Remote Sensing (Submitted).

Vieira, S. R., D. R. Nielsen and J. W. Biggar. 1981. Spatial variability of field-measured infiltration rate. Soil Sci. Soc. Am. J. 45:1040-1048.

Waynick, D. D. 1918. Variability in soils and its significance to past and future soil investigations. I. Statistical study of nitrification in soils. Univ. Cal. Publ. Agr. Sci. 3(9):243-270.

Waynick, D. D., and L. T. Sharp. 1919. Variability in soils and its significance to past and future soil investigations. II. Variation in nitrogen and carbon in field soils and their relation to the accuracy of field trials. Univ. Cal. Publ. Agr. Sci. 4(5):121-139.

Webster, R. 1973. Automatic soil boundary location for transect data. Mathematical Geology 5(1):27-37.

Webster, R., and H. E. de la C. Cuanalo. 1975. Soil transects correlograms of north Oxfordshire and their interpretation. J. Soil Sci. 26(2): 176-194.

Webster, R., and T. M. Burgess. 1980. Optimal interpolation and isarithmic mapping of soil properties. III. Changing drift and universal kriging. J. Soil Sci. 31:505-524.