Mathematical Geology, Vol. 23, No. 6, 1991

Pseudo-Cross Variograms, Positive-Definiteness, and Cokriging[1]

Donald E. Myers[2]

Cokriging allows the use of data on correlated variables to be used to enhance the estimation of a primary variable or more generally to enhance the estimation of all variables. In the first case, known as the undersampled case, it allows data on an auxiliary variable to be used to make up for an insufficient amount of data. Original formulations required that there be sufficiently many locations where data is available for both variables. The pseudo-cross-variogram, introduced by Clark et al. (1989), allows computing a related empirical spatial function in order to model the function, which can then be used in the cokriging equations in lieu of the cross-variogram. A number of questions left unanswered by Clark et al. are resolved, such as the availability of valid models, an appropriate definition of positive-definiteness, and the relationship of the pseudo-cross-variogram to the usual cross-variogram. The latter is important for modeling this function.

KEY WORDS: cokriging, cross-variograms, positive-definiteness, undersampled.

INTRODUCTION

Cokriging is a generalization of kriging in that it uses not only the spatial correlation of the variable(s) of interest but also the intervariable spatial correlation, and allows estimation of the variable(s) at an unsampled location using not only data for variable but also data from the correlated variables. Many of the early applications used the so-called undersampled formulation, where the data from the correlated variables is used to compensate for a lack of data for the variable of principal interest. However, there is no necessity nor advantage gained from focusing on the estimation of only one of the variables. All may be estimated using all the data with little additional computing.

With the exception of the formulation used by Clark et al. (1989), all of the formulations of cokriging require the use and hence the estimation/modeling of cross-variograms or cross-covariances for each pair of variables used in the analysis. If $Z_1(x), \ldots, Z_m(x)$ are random functions, where x is in 1, 2, or

---

[1]Manuscript received 18 July 1990; accepted 12 February 1991.
[2]Department of Mathematics, University of Arizona, Tucson, AZ 85721.

0882-8121/91/0800-0805$06.50/1

3-space and the objective is to estimate the point value $Z_1(x_0)$, then the weights for the respective data $Z_j(x_i)$ are obtained as the solution to a set of linear equations, and the coefficients are the values of variograms and cross-variograms for pairs of data locations. To estimate the cross-variogram of $Z_j(x)$, $Z_k(x)$ requires a sufficient number of locations where data is available for both variables of interest, a condition which is frequently not satisfied in practice and in particular in an undersampled problem. If the set of locations where data is available for $Z_j(x)$ is a subset of the locations for $Z_k(x)$, then one ad hoc solution is to simply ignore the extra data for $Z_k(x)$ (when estimating the cross-variogram). In general, this will reduce the information available for estimation and hence make the estimation process less reliable; however, more generally, it simply does not use all available information. Note that in an extreme case there may not be any sample locations where data is available for both variables in a pair.

Clark et al. (1989) introduced a pseudo-cross-variogram, instead, whose estimation does not seem to require that data for $Z_j(x)$, $Z_k(x)$ both be available at any of the same locations. They derived the cokriging equations using these pseudo-cross-variograms, although only in an undersampled form, but did not deal with a number of important questions such as what are valid models, what is the relationship to the usual cross-variograms, and what underlying hypotheses or assumptions are needed. It is shown in Myers (1982, 1984, 1985, 1988a) that the usual system for estimation in the undersampled case is a special case of a more general formulation, and that the computer program can incorporate the undersampling. This algorithm will be extended to a formulation using the pseudo-cross-variograms.

THE PSEUDO-CROSS-VARIOGRAMS

Let $Z_j(x)$, $Z_k(x)$ be two random functions of interest defined in 1, 2, or 3-space (it could equally well be in higher dimensional space) and h a vector. Clark et al. (1989) define the pseudo-cross-variogram of Z_j, Z_k as

$$\hat{g}_{jk}(h) = 0.5E[Z_j(x) - Z_k(x + h)]^2 \tag{1}$$

where it is assumed that this function depends only on h. If $j = k$ and the random function satisfies the Intrinsic Hypothesis, then Eq. (1) is the usual variogram. Note that (1) is not the same as half the variance of the difference even if both random functions separately satisfy the Intrinsic Hypothesis since in general they may not have the same means. If the random functions do not have constant means, then the discrepancy may be more serious and hence we begin by generalizing the definition to the following

$$g_{jk}(h) = g_{kj}(-h) = 0.5 \text{ var } [Z_j(x) - Z_k(x + h)] \tag{1'}$$

and in subsequent sections both Eqs. (1) and (1′) will be used. Sufficient conditions for the existence of (1), (1′) and the required independence with respect to x are easy to formulate. Let $Z_j(x) = Y_j(x) + m_j(x)$, where $Y_j(x)$ is second-order stationary and $m_j(x)$ is the mean of $Z_j(x)$. As usual, $m_j(x)$ will be assumed to be a linear combination of known basis functions but with unknown coefficients. Since the translation of a second-order stationary random function is second-order stationary, and the sum of two second-order stationary random functions is second-order stationary, Eq. (1′) is the variance of a random function and depends only on h. It is not, however, the covariance nor the variogram of a random function unless $j = k$. To see the relationship with the usual cross-variogram, assume that $Z_j(x)$, $Z_k(x)$ have the form above and that the cross-covariance of $Z_j(x)$, $Z_k(x + h)$ is symmetric [this will imply that $g_{jk}(h)$ and $g_{jk}(h)$ are symmetric]. Then (1′) can be rewritten as

$$g_{jk}(h) = 0.5[C_{jj}(0) - 2C_{jk}(0) + C_{kk}(0)] + \gamma_{jk}(h) \tag{2}$$

where $\gamma_{jk}(h)$ is the usual cross-variogram. If moreover $m_j(x)$, $m_k(x)$ are constants m_j, m_k then

$$\hat{g}_{jk}(h) = g_{jk}(h) + 0.5[m_j^2 - 2m_jm_k + m_k^2] \tag{3}$$

and hence, with sufficiently strong assumptions, the pseudo-cross-variogram of Clark et al. differs from the usual cross-variogram by a (positive) constant. The usual cross-variogram $\gamma_{jk}(h)$ is zero for $h = 0$, whereas in general $\hat{g}_{jk}(0) > 0$. This difference is not just a nugget effect. Aside from the advantage that data is not required for both random functions at the same locations, it is also seen from Eqs. (2) and (3) that both (1) and (1′) do quantify the cross-correlation between the random functions Z_j, Z_k. If $m_j = m_k$, then of course Eqs. (1) and (1′) coincide. It is also possible to define a symmetric version of (1) or (1′). Form a new function $W_{jk}(x) = Z_j(x) + Z_k(x)$, then let

$$g_{w_{jk}, w_{jk}}(h) = g_{jk}(h) + g_{kj}(h) - \operatorname{cov}\{Z_j(x) - Z_k(x + h), Z_j(x + h) - Z_k(x)\} \tag{4}$$

The covariance term in Eq. (4) plays the role of a symmetric pseudo-cross-variogram, although it is not clear that this is useful.

Although Eq. (1′) is defined by a variance and (1) by the expected value of a certain square, neither is a true variogram or covariance unless $j = k$. Therefore, the question of appropriate models for either is not easily resolved. A somewhat similar problem arises in the case of cross-variograms, but the latter is directly related to a certain linear combination of variograms as indicated in Myers (1982, 1988a) and the modeling is thus somewhat simpler. The objective in fitting a valid model is to ensure that the estimation variance remains positive and that the system of cokriging equations has a unique solution.

TWO USEFUL RELATIONS

The relations may be stated either in terms of Eq. (1) or (1′)

$$E\{[Z_j(x) - Z_p(u)][Z_k(y) - Z_p(u)]\} = \hat{g}_{jp}(u - x) + \hat{g}_{kp}(u - y) - \hat{g}_{jk}(y - x) \tag{5}$$

$$\mathrm{cov}\,\{[Z_j(x) - Z_p(u)], [Z_k(y) - Z_p(u)]\} = g_{jp}(u - x) + g_{kp}(u - y) - g_{jk}(y - x) \tag{5'}$$

Although not explicitly stated there, Eq. (5) is implicit in Clark et al. Both Eqs. (5) and (5′) are variations of a well-known result for variograms and cross-variograms. They are needed in order to write the estimation variance(s) in terms of variograms and pseudo-cross-variograms.

THE COKRIGING EQUATIONS

For simplicity we begin by considering the problem of estimating $Z_1(x_0)$ using the data $\bar{Z}(x_i) = [Z_1(x_i), \ldots, Z_m(x_i)]$; $i = 1, \ldots, n$. Although Clark et al. did not assume data were available for all functions at all locations, that assumption has no real impact on their derivation. That special case can be obtained from the full sampled problem in a manner analogous to the algorithm given in Myers (1984, 1988a,b) and implemented in Carr et al. (1985). More recently, this program has been implemented in the Geo-EAS format. For simplicity and without loss of generality we will consider point estimation. The estimator for $Z_1(x_0)$ is given by

$$Z_1^*(x_0) = \Sigma\bar{Z}(x_i)\Gamma_i \tag{6}$$

where Γ_i is a column vector of weights λ_s^i. If the separate components of $Z(x_i)$ satisfy the Intrinsic Hypothesis, then for Eq. (6) to be unbiased it is sufficient for

$$\Sigma\Gamma_i = [1, 0, \ldots, 0]^T \tag{7}$$

In that case, the variance of the error can be written in the form

$$\mathrm{var}\,\{Z_1^*(x_0) - Z_1(x_0)\} = \Sigma\Sigma\Gamma_i^T E\{W_i^T W_j\}\Gamma_j \tag{8}$$

where

$$W_i = [Z_1(x_i) - Z_1(x_0), \ldots, Z_m(x_i) - Z_1(x_0)] \tag{9}$$

The "covariance" matrix in Eq. (8) can be written in a more convenient form using the identity (5)

$$E\{W_i^T W_j\} = \hat{G}_1(x_0 - x_i) + \{\hat{G}_1(x_0 - x_j)\}^T - \hat{G}(x_j - x_i) \tag{10}$$

where

$$\hat{G}_k(x - y) = \begin{bmatrix} \hat{g}_{1k}(x - y) & \dots & \hat{g}_{1k}(x - y) \\ & \dots & \\ & \dots & \\ & \dots & \\ \hat{g}_{mk}(x - y) & \dots & \hat{g}_{mk}(x - y) \end{bmatrix} \tag{11}$$

for $k = 1, \dots, p$.

$$G(x - y) = \begin{bmatrix} \hat{g}_{11}(x - y) & \dots & \hat{g}_{1m}(x - y) \\ & \dots & \\ & \dots & \\ & \dots & \\ \hat{g}_{m1}(x - y) & \dots & \hat{g}_{mm}(x - y) \end{bmatrix} \tag{12}$$

The right-hand side of Eq. (8) can then be written in the form

$$\begin{aligned} \Sigma\Sigma\Gamma_i^T E\{W_i^T W_j\}\Gamma_j = {} & 2\Sigma[\hat{g}_{11}(x_0 - x_j), \dots, \hat{g}_{1m}(x_0 - x_j)]\Gamma_j \\ & - \Sigma\Sigma\Gamma_i^T \hat{G}(x_j - x_j)\Gamma_j \end{aligned} \tag{13}$$

by using (7). To minimize Eq. (8) subject to the constraints given in (7), we must introduce m Lagrange multipliers $\mu_1, \dots, \mu_m$ and form

$$\begin{aligned} & \phi(\lambda_s^i, \mu_s; i = 1, \dots n; s = 1, \dots, m) \\ & \quad = \Sigma\Sigma\Gamma_i' E\{W_i^T W_j\}\Gamma_j - 2\Sigma\Sigma\mu_s(\lambda_s^i - \delta(s - 1)) \end{aligned} \tag{14}$$

where $\delta(u) = 1$ for $u = 0$ and $= 0$ otherwise. As in all the kriging formulations, Eq. (14) is minimized by differentiating with respect to each of the unknowns and setting all derivatives equal to zero. Differentiating with respect to the μ_s will reproduce the constraints given in (7). Differentiating with respect to each the λ_s^i gives the set of equations

$$\begin{aligned} & \Sigma\hat{G}(x_i - x_j)\Gamma_j + [\mu_1, \dots, \mu_m]^T = [\hat{g}_{11}(x_0 - x_i), \dots, \hat{g}_{1m}(x_0 - x_i)]^T \\ & \quad i = 1, \dots, n \end{aligned} \tag{15}$$

It is relatively easy to see how to change the system if we wish to estimate a different component and at a different place. It is also easy to see that we can simultaneously estimate several components and even estimate at different locations but using the same data for all the estimations. The weight vectors Γ_j will have to become matrices (i.e., a column for each component estimated), likewise the column of Lagrange multipliers will have to become a matrix and the column on the right-hand side of (15) likewise will become a matrix. This simultaneous estimation corresponds to minimizing the sum of estimation variances in exactly the same way as in the formulation given in Myers (1982, 1988a,b). If the entire system is written in matrix form, these properties are more evident and likewise the appropriate conditions on the matrix function $G(h)$. Let $\hat{G}_{k0}(x_0 - x_i) = [\hat{g}_{k1}(x_0 - x_i), \ldots, \hat{g}_{km}(x_0 - x_i)]^T$ for $k = 1, \ldots, p$, then the system can be written in the form

$$\begin{bmatrix} \hat{G}(x_1 - x_1) & \ldots & \hat{G}(x_1 - x_n) & I \\ & \ldots & & I \\ & \ldots & & I \\ \hat{G}(x_n - x_1) & \ldots & \hat{G}(x_n - x_n) & I \\ I & \ldots & I & 0 \end{bmatrix} \begin{bmatrix} \Gamma_1 \\ \cdot \\ \cdot \\ \Gamma_n \\ \mu \end{bmatrix} = \begin{bmatrix} \hat{G}_{10}(x_0 - x_1) \\ \cdot \\ \cdot \\ \hat{G}_{10}(x_0 - x_n) \\ I_0 \end{bmatrix}$$

where I is an $m \times m$ identity matrix, I_0 is the column on the right-hand side of Eq. (7) and μ is the vector of Lagrange multipliers. Unlike the matrix variogram function used in Myers (1982), $\hat{G}(h)$ need not be a symmetric matrix. To extend the system so that all variables can be estimated, it is only necessary to adjoin additional columns to the weight and Lagrange multiplier vectors, and similarly to adjoin columns on the right-hand side of the system of equations corresponding to each of the variables being estimated. Sufficient conditions for the invertibility of the coefficient matrix will be given subsequently.

The similarity with the system given in Myers (1982) is reenforced if we consider the special case resulting in Eqs. (2) and (3) [i.e., $g_{jk}(h) = \gamma_{jk}(h) + a_{jk}$ with $a_{jj} = 0$]. In fact, the constants a_{jk} have no effect on the solution of the system given in Eq. (15). If this simplification is used in Eq. (15), then the usual system of cokriging equations in terms of variograms and cross-variograms is obtained.

POSITIVE-DEFINITENESS

Let $G(h)$ be an $m \times m$ matrix valued function with $G(-h) = G(h)^T$, defined on 1, 2, or 3-space (or higher dimension). G is said to be conditionally (strictly) positive-definite if

$$\Sigma\Sigma\Gamma_i^T G(x_i - x_j)\Gamma_j > 0 \tag{16}$$

for all points $x_1, \ldots, x_n$ and all weight vectors $\Gamma_1, \ldots, \Gamma_n$ not all identically zero but whose sum is the zero vector. As shown in Myers (1988c), this is exactly the right condition to ensure that the coefficient matrix for the system of co-kriging equations is invertible. While this definition of positive-definiteness provides a sufficient condition, it does not directly indicate how to model the individual entries in the matrix function except that the diagonal entries are variograms and can be estimated/modeled in (one of) the usual way(s). Clearly, the pseudo-cross-variograms cannot be modeled independently of the associated pair of variograms. To see that conditional positive-definiteness as given above is sufficient for the invertibility of the coefficient matrix to abbreviate the coefficient matrix as follows

$$\begin{bmatrix} K & E \\ E^T & 0 \end{bmatrix}$$

where E is a column of identity matrices. If this matrix is not invertible, then there is a vector $[U^T \quad V^T]^T$ not identically zero such that

$$\begin{bmatrix} K & E \\ E^T & 0 \end{bmatrix} \begin{bmatrix} U \\ V \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

and hence

$$\text{(i) } KU + EV = 0 \qquad \text{and} \qquad \text{(ii) } E^T U = 0$$

Condition (ii) is exactly that imposed on the weight vectors in the definition of conditional positive-definiteness—that is, that they sum to zero. In addition (ii) implies that $U^T E = 0$ and hence $U^T EV = 0$. Now if both sides of (i) are premultiplied by U^T, then we have $U^T KU = 0$, which would contradict the assumption of conditional positive-definiteness unless the vector U is identically zero. In that case, KU would be zero and hence $EV = 0$, which implies that $V = 0$. This argument will extend easily to the case of universal cokriging.

It was noted above, that in the case of second-order stationarity and a symmetric covariance, that the pseudo-cross-variogram differs from the usual cross-variogram by a constant. Hence, the matrix function whose diagonal entries are variograms and whose off-diagonal entries are pseudo-cross-variograms can be written in the form

$$G(h) = \overline{\gamma}(h) + A$$

where A is a symmetric matrix with zeros on the diagonal. G is conditionally positive-definite in the sense given above if and only if $\overline{\gamma}$ is conditionally positive-definite—this means that in the case of symmetry pseudo-cross-variograms can be modeled by *valid* cross-variograms plus a constant.

ESTIMATION AND MODELING

Since Eqs. (1), (1′) are expected values, it would seem plausible to estimate the pseudo-cross-variogram in the same way as an ordinary variogram. For a given pair of random functions $Z_j(x)$, $Z_k(x)$, estimate $\hat{g}_{jk}(h)$ by

$$\hat{g}^*_{jk}(h) = \{1/2N(h)\}\Sigma[Z_j(x_i) - Z_k(x_i + h)]^2 \tag{17}$$

where as usual $N(h)$ is the number of pairs of points x_i, $x_i + h$ with data for Z_j at x_i *and* data for Z_k at $x_i + h$. It is likely that as in estimating variograms, distance classes and angle windows would have to be used. Because neither Eq. (1) nor (1′) is symmetric, it is necessary to examine whether $\hat{g}^*_{jk}(h)$ is approximately the same as $\hat{g}^*_{jk}(-h)$. Having obtained estimated values for the pseudo-cross-variogram for different lags, it is still necessary to select a valid model. For as noted above, neither Eq. (1) nor (1′) is the variogram of a random function, nor is either simply related to a linear combination of variograms.

The relationship given by Eqs. (2) and (3) does suggest a method for modeling the pseudo-cross-variograms. If the plots of Eq. (16) appear to be symmetric, then the pseudo-cross-variogram differs from a cross-variogram by a positive constant (i.e., the value at $h = 0$). Note that this constant is not the same as a nugget effect. First, estimate the constant from the plots, and then model the cross-variogram as described in Myers (1982, 1988a,b). This method is then simply a way to improve the plotting of cross-variograms and is not a direct way to model pseudo-cross-variograms. This can be seen in another way. If $G(0)$ is the matrix of constants $g_{st}(0)$, then $G(h)$ satisfies (16) if and only if $G(h)$-$G(0)$ satisfies (16).

Particularly in early applications, the cross-variograms were frequently modeled by using a linear representation. Let $\overline{U}(x) = [U_1(x), \ldots, U_p(x)]$ be a vector of uncorrelated random functions with variogram matrix function $\overline{\gamma}_U(h)$ and means $m_1, \ldots, m_p$. If the vector $\overline{Z}(x) = [Z_1(x), \ldots, Z_m(x)]$ has the representation

$$\overline{Z}(x) = \overline{U}(x)A = \overline{U}(x)[A_1, \ldots, A_m] \tag{18}$$

where the A_i's are $1 \times p$ vectors of constants, then the variogram matrix function for $\overline{Z}$ can be written as

$$\overline{\gamma}_z(h) = A^T\overline{\gamma}_U(h)A \tag{19}$$

where the cross variogram for Z_i, Z_j—that is, the ij entry in (19) is given by

$$\gamma_{ij}(h) = A_i^T\overline{\gamma}_U(h)A_j \tag{20}$$

and hence the cross-variogram is a certain linear combination of the variograms for the U's. To obtain the analog of the variogram matrix function, it is necessary to assume that the components $Y_i(x)$ are second-order stationary. Then the analogous function would be given by Eq. (12), and an entry in (12) would be given by

$$\hat{g}_{ij}(h) = A_i^TDA_i + A_j^TDA_j - A_i^TDA_j - A_j^TDA_i - A_i^T\overline{\gamma}_Y(h)A_j - A_j^T\overline{\gamma}_Y(h)A_i \tag{21}$$

where $D = C(0) + M^2 = E\{\overline{U}^T(x)\overline{U}(x + h)\}$. The representation for the pseudo-cross-variograms is then more complicated than for cross-variograms and the conditions for verifying positive-definiteness are also more complicated. Unfortunately, it is clear that there is no real substitute for data at an adequate number of locations for all the variables of interest.

COMPUTER PROGRAMS

Since there is essentially no difference in the form of the cokriging equations using pseudo-cross-variograms, any existing program for solving cokriging equations can be used nearly unchanged. The differences and changes are easy to identify. First of all, the pseudo-cross-variogram need not be symmetric, as noted above, and hence the filling in of the entries in the matrix function and in the subsequent coefficient matrix (15) is slightly different. The class of valid models is also slightly different for two reasons; one being the possibility of nonsymmetry, and secondly the positive constants a_{jk} in $g_{jk}(h) = \gamma_{jk} + a_{jk}$ discussed above. In any case, the changes could be viewed as a generalization and a modified program would still function for cokriging using the usual cross variograms.

The estimation of pseudo-cross variograms is, of course, quite different. Although the proposed estimator given by (17) is the obvious analogue of the usual sample or empirical variogram, the search process in the algorithm for finding the pairs is considerably more complicated. Each data location must in some sense be coded to indicate which variables have data at that location. In particular, for a given pair of variables Z_j, Z_k the sample location pair $(x_i, x_i + h)$ is not the same as the pair $(x_i + h, x_i)$. In fact, it would be possible for the set of data locations for Z_j to be completely disjoint from those for Z_k. Algorithms and programs for computing the sample variogram are essentially of two kinds; the most common sorts the pairs into distance classes and angle windows as the sample variogram is computed, and thus no list is constructed of the pairs utilized. Several recently released programs (e.g., PREVAR in the GEOEAS package) first sort the data locations and only afterward use this file to compute the sample variogram. Converting the first kind of progam to compute the sample pseudo-cross-variogram is fairly easy, and this has been implemented. In the case where the sample locations are sorted and the file of sorted pairs retained, it would seem that such a file would have to be partitioned into two files (i.e., corresponding to h and to $-h$). While the second approach has clear advantages, it may require more memory than is available on the particular computer, and it also will produce very large files. Subsequent plotting of the sample pseudo-cross-variogram will also be slightly more complicated in that it will be necessary to plot for both h and $-h$, but it is necessary in order to check for nonsymmetries.

BLOCK ESTIMATION

If instead of wanting to estimate $Z_1(x_0)$ the objective is to estimate $Z_{1,A}$, where

$$Z_{1,A} = (1/A) \int Z_1(u)\, du \tag{22}$$

then certain terms in Eqs. (8) and (9) must be changed. In particular, the terms of the form

$$E\{Z_s(x_i) - Z_1(x_0)\}\{Z_t(x_j) - Z_1(x_0)\} \tag{23}$$

must be replaced by

$$E\{Z_s(x_i) - Z_{1,A}\}\{Z_t(x_j) - Z_{1,A}\} \tag{24}$$

Using the identity Eq. (5) and the same technique as is used in converting point kriging equations to block kriging, Eq. (24) becomes

$$\begin{aligned}(1/A) \int \hat{g}_{1s}(u - x_i)\, du + (1/A) \int \hat{g}_{1t}(v - x_j)\, dv \\ - \hat{g}_{st}(x_i - x_j) - (1/A^2) \int\int \hat{g}_{11}(u - v)\, du\, dv\end{aligned} \tag{25}$$

which, in keeping with common notation, we will write as

$$\hat{g}_{1s}(A, x_i) + \hat{g}_{1t}(A, x_j) - \hat{g}_{st}(x_i - x_j) - \hat{g}_{11}(A, A) \tag{26}$$

This will not result in a change in the basic form of the equations but it will add a term to the kriging variance, namely, $-\hat{g}_{11}(A, A)$. The terms on the right-hand side of (15) will then have to be replaced by average values.

SUMS, DIFFERENCES, AND INHOMOGENEITIES

Unlike the ordinary cross-variogram wherein the differences are formed for only the same variable, in the pseudo-cross-variogram differences of pairs of variables are required. At least in terms of an interpretation there can be a problem when the variables are given in entirely different units or represent very different phenomena. Although such a disparity does not lead to mathematical problems, the relevance of such a statistic may be open to question. There may be practical problems when the variability of the data for one variable is vastly different from that of the other or where the scale of magnitudes is quite different (e.g., if the sample pseudo-cross-variogram is computed from data for Z_1, Z_2, and the latter has values several orders of magnitude larger than the former, or vice versa). In such a case, the values of the sample function are essentially those of a sample variance of the larger-valued variable. The data can easily be rescaled in the case of the ordinary sample cross-variogram or the sample variogram of the sum or difference. For example, if a constant m_1 is subtracted from all the data values for Z_1, and a constant m_2 is subtracted from all the data values for Z_2, the sample cross-variogram is unchanged. The same is true for the sample variogram of the sum of the variables. If the data for Z_1 is multiplied by a constant b_1 and the data for Z_2 is multiplied by a constant b_2, then the sample cross-variogram is multiplied by $b_1 b_2$. The theoretical cross-variogram has analogous properties. Neither of these properties is true for the sample pseudo-cross-variogram, although one could "normalize" the data and then (after cokriging) remove the normalization from the cokriged values.

A disparity in the units is not crucial in the cokriging estimator since the weights (i.e., the coefficients in the estimator) can be interpreted as reflecting the change in units. If data is transformed, by subtraction of a constant or multiplication by a constant, the retransformation can be applied either before or after cokriging, when variograms and cross-variograms are used.

SUMMARY AND CONCLUSIONS

While the pseudo-cross-variogram defined by Clark et al. (1989) is not a true cross-variogram nor a variogram, it can be used to derive the cokriging equations, and the resulting equations are nearly identical to the ones using cross-covariances or cross-variograms. This cross-variogram has the advantage that it is not necessary to have data on both variables at any common locations and hence the empirical pseudo-cross-variogram can be computed in many instances where the usual sample cross-variogram cannot be. Because it is intended for problems wherein one variable has little data, or there are few locations where data is available for both variables, it would not be easy in that case to use the sum and difference variograms in order to determine valid models for the cross-variogram. Modification of existing software to compute the empirical pseudo-cross-variogram, or cokriging software to utilize the pseudo-cross-variogram, is relatively easy. This pseudo-cross-variogram will have significant importance in connection with modeling both spatial and temporal dependence and correlation.

REFERENCES

Carr, J., Myers, D. E., and Glass, C., 1985, Co-Kriging: A Computer Program: Computers and Geosciences, v. 11, p. 111-127.

Clark, I., Basinger, K., and Harper, W., 1989, MUCK—A Novel Approach to Co-Kriging. *in* B. E. Buxton (Ed.), Proceedings of the Conference on Geostatistical, Sensitivity, and Uncertainty: Methods for Ground-Water Flow and Radionuclide Transport Modeling: Batelle Press, Columbus, p. 473-494.

Myers, D. E., 1982, Matrix Formulation of Cokriging: Math. Geol., v. 14, p. 249-257.

Myers, D. E., 1983, Estimation of Linear Combinations and Cokriging: Math. Geol., v. 15, p. 633-637.

Myers, D. E., 1984, Cokriging: New Developments, *in* G. Verly et al., (Eds.), Geostatistics for Natural Resource Characterization: D. Reidel, Dordrecht.

Myers, D. E., 1985, Co-Kriging: Methods and Alternatives, *in* P. Glaeser, (Ed.), The Role of Data in Scientific Progress: Elsevier, New York.

Myers, D. E., 1988a, Some Aspects of Multivariate Analysis, *in* C. F. Chung et al. (Eds.), Quantitative Analysis of Mineral and Energy Resources: D. Reidel, Dordrecht, p. 669-687.

Myers, D. E., 1988b, Multivariate Geostatistics for Environmental Monitoring: Sciences de la Terre, v. 27, p. 411-428.

Myers, D. E., 1988c, Interpolation with Positive Definite Functions: Sciences de la Terre, v. 28, p. 251-265.

Myers, D. E., 1989, Vector Conditional Simulation, *in* M. Armstrong (Ed.), Geostatistics: D. Reidel, Dordrecht, p. 283-293.