Mathematical [illegible]ogy, Vol. 17, No. 2, 1985

Minimum-Variance Unbiased Quadratic Estimation of Covariances of Regionalized Variables[1]

Peter K. Kitanidis[2]

The parameters of covariance functions (or variograms) of regionalized variables must be determined before linear unbiased estimation can be applied. This work examines the problem of minimum-variance unbiased quadratic estimation of the parameters of ordinary or generalized covariance functions of regionalized variables. Attention is limited to covariance functions that are linear in the parameters and the normality assumption is invoked when fourth moments of the data need to be calculated. The main contributions of this work are (1) it shows when and in what sense minimum-variance unbiased quadratic estimation can be achieved, and (2) it yields a well-founded, practicable, and easy-to-automate methodology for the estimation of parameters of covariance functions. Results of simulation studies are very encouraging.

KEY WORDS: geostatistics, covariance estimation, optimization.

INTRODUCTION

Linear minimum-variance unbiased estimation, also known as kriging, has found many applications in geology, hydrogeology, and mining exploration. It can be used to obtain the best (according to quadratic criteria) unbiased estimates of point values, areal averages, or any other linear function of the regionalized variable. Kriging is also used in cartography for interpolation from nonuniformly distributed data onto the nodes of a regular grid to obtain contour maps. Theoretical and applied aspects of kriging have been extensively discussed in books such as Matheron (1971) and Journel and Huijbregts (1978) and in papers too numerous to mention.

By comparison, less attention has been given to the important problem of optimal estimation of parameters of the covariance function or variogram and (if needed) the drift. Part of the reason is that linear geostatistics is distribution-

[1] Manuscript received 19 December 1983; revised 25 April 1984.

[2] St. Anthony Falls Hydraulic Laboratory, Department of Civil and Mineral Engineering, University of Minnesota, Mississippi River at Third Avenue S.E., Minneapolis, Minnesota 55414, U.S.A.

free, involving only the first two moments of regionalized variables, whereas optimal estimation of covariance functions generally involves greater moments. The problem of estimation of structural parameters has been examined in the works of Matheron (1971), Olea (1975), Delfiner (1976), Davis and David (1978), Journel and Huijbregts (1978), Cressie and Hawkins (1980), Kafritsas and Bras (1981), Starks and Fang (1982), Kitanidis and Vomvoris (1983), Kitanidis (1983), and others.

The objective of this paper is to investigate the problem of estimation of parameters of covariance functions as quadratic functions of the data in such a way that estimates are unbiased and have minimum variance. Biased estimation of variograms or covariance functions has been one of the difficulties in application of universal kriging. Furthermore, interest in estimation of covariance functions is not limited to applications of best linear unbiased estimation. A case in point is the covariance function of hydraulic conductivity which is required in evaluation of solute macrodispersion in aquifers (Gelhar and Axness, 1983). The analysis presented here provides useful insight into problems of parameter estimation of covariance functions or variograms and, at the same time, yields a useful and practicable estimation method. Analysis is limited to the useful class of ordinary or generalized covariance functions which are linear in the parameters. The model is assumed known and the normality assumption is invoked whenever fourth moments of the data need to be calculated. The proposed methodology is an extension and generalization of work reported in Kitanidis (1983) and is related to Rao's (1971) method for estimation of variance components.

THE GENERAL MODEL

The regionalized variable at point i is assumed to be the realization of the general linear model

$$y_i = \mathbf{x}_i \mathbf{b} + \epsilon_i \tag{1}$$

where $\mathbf{x}_i$ is a $1 \times p$ vector of known variables which depend on the spatial coordinates of y_i, $\mathbf{b}$ is a $p \times 1$ vector of coefficients and ϵ_i is a zero-mean random term. The term $\mathbf{x}_i\mathbf{b}$ represents the prior expected value (or drift) and the term ϵ_i represents the random part of y_i.

If $\mathbf{y}$ is the $n \times 1$ vector consisting of all available observations, the general linear model can be written, following standard notation, as

$$\mathbf{y} = X\mathbf{b} + \boldsymbol{\epsilon} \tag{2}$$

where $\mathbf{X}$ is an $n \times p$ known matrix, having $\mathbf{x}_i$ as its ith row, and $\boldsymbol{\epsilon}$ is an $n \times 1$ vector of random terms with covariance matrix $K(\boldsymbol{\theta})$ where $\boldsymbol{\theta}$ is a vector of unknown coefficients.

The general estimation problem is to determine from available data the unknown drift coefficients **b** and the covariance function parameters $\boldsymbol{\theta}$. Simultaneous estimation of drift and covariance function parameters from spatial data is possible, for example, through maximum likelihood estimation (see discussion in Kitanidis and Vomvoris, 1983). However, this approach is known (Matheron, 1971) to yield biased estimates of variograms or covariance-function parameters. For the linear model given by eq. (2), estimation of parameters of drift may be bypassed so that unbiased estimates of variograms or covariance functions can be obtained. This approach is particularly appropriate when the purpose of parameter identification is application of linear unbiased minimum variance estimation (kriging, universal kriging, or kriging with intrinsic random functions) which do not use drift coefficients ("universality" condition, see Matheron, 1971). This approach is illustrated in the next section.

This paper presents a unified methodology for estimation of parameters of ordinary as well as generalized covariance functions (Matheron, 1973, Delfiner, 1976). The (ordinary or generalized) covariance matrix K of the measurement vector (see Kitanidis, 1983) is assumed in this work to be a linear function of the parameters

$$K = \sum_{i=1}^{m} K_i \theta_i \tag{3}$$

where K_i are known symmetric $n \times n$ matrices and θ_i are unknown parameters to be estimated from available data. Not only is the assumption of linear (in the parameters) covariance functions appropriate as a starting point but also, in the author's opinion, such models are useful in practice. Examples of such covariance functions include the nugget effect; the polynomial generalized covariance function (Matheron, 1973); the interpolating spline (Dubrule, 1983); covariance functions with known correlation structure but unknown variance, such as the spherical with known range; and polygonal and other representations of covariance functions. Extensions to estimation of nonlinear (in the parameters) covariance functions are possible but will not be pursued here.

MINIMUM-VARIANCE UNBIASED QUADRATIC ESTIMATION

Consider that parameters must be estimated as quadratic functions of the data

$$\hat{\theta}_j = \mathbf{y}^T F_j \mathbf{y} \qquad j = 1, \ldots, m \tag{4}$$

where F_j, $j = 1, \ldots, m$ are $n \times n$ matrices to be selected according to the following specifications:

(a) Parameter estimates should be invariant to the addition to the original data of any trend of the assumed form with arbitrary coefficients. That is

$$\hat{\theta}_j = \mathbf{y}^T F_j \mathbf{y} = (\mathbf{y} + X\mathbf{b})^T F_j (\mathbf{y} + X\mathbf{b}) \tag{5}$$

For this condition to hold, it is required that

$$F_j X = 0 \qquad j = 1, \ldots, m \tag{6a}$$

$$F_j^T X = 0 \qquad j = 1, \ldots, m \tag{6b}$$

Such estimators have been referred to as universal quadratic estimators (Matheron, 1971, p. 191). The motivation for condition (6) is to avoid estimation of drift coefficients. In the case of generalized covariance functions, this condition insures that only authorized increments are involved in calculation of $\hat{\theta}_j$. Consequently, $E(\hat{\theta}_j) = Tr[F_j K]$.

(b) Parameter estimates must be unbiased

$$\theta_j = E(\hat{\theta}_j) = E[Tr(F_j \mathbf{y}\mathbf{y}^T)] = \sum_{i=1}^{m} Tr(F_j K_i)\theta_i \qquad j = 1, \ldots, m \tag{7}$$

For this condition to hold, it is required that

$$Tr(F_j K_i) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \tag{8}$$

for every i and j. Note that all that the universality and unbiasedness conditions require are the models of the drift and the covariance function.

(c) Variance of the estimation error for parameter θ_j must be minimum. The variance is

$$E(\hat{\theta}_j - \theta_j)^2 = E(\mathbf{y}^T F_j \mathbf{y} - \theta_j)^2 = E[Tr(F_j \mathbf{y}\mathbf{y}^T)\, Tr(F_j \mathbf{y}\mathbf{y}^T)] - \theta_j^2 \tag{9}$$

where the unbiasedness condition was taken into account. The variance involves the fourth moments of the data. An assumption of normality is often appropriate or, in practice, may be invoked for lack of a better alternative. Gaussian moment factoring and the unbiasedness conditions give

$$E(\hat{\theta}_j - \theta_j)^2 = 2Tr(F_j K F_j K) \tag{10}$$

This variance must be minimized subject to the equality constraints (6) and (8). The Lagrangian of the minimization problem is

$$2Tr(F_j K F_j K) - \sum_{\substack{i=1 \\ i \neq j}}^{m} 4\lambda_i Tr(F_j K_i) - 4\lambda_j [Tr(F_j K_j) - 1]$$

$$- \sum_{i=1}^{n} \sum_{k=1}^{p} 4v_{ik}(F_j X)_{ik} - \sum_{i=1}^{n} \sum_{k=1}^{p} 4u_{ik}(F_j^T X)_{ik} \tag{11}$$

where $\lambda_1, \lambda_2, \ldots, \lambda_m$ are Lagrange multipliers associated with the unbiasedness constraints and v_{ik} and u_{ik}, $i = 1, \ldots, n$, $j = 1, \ldots, p$ are Lagrange multipliers associated with the universality condition. The coefficient 4 is used only for convenience. Note that because K is a function of the unknown parameters the optimization problem is not quadratic with respect to F_j. One may proceed by assuming that a prior estimate of K, denoted by K_o, is available. Then, F_j is calculated by taking derivatives of the Lagrangian with respect to matrix F_j (see Athans and Schweppe, 1965)

$$K_o F_j K_o - \sum_{l=1}^{m} \lambda_l K_l - VX^T - XU^T = 0 \tag{12}$$

where V is $n \times p$ matrix whose ijth element is v_{ij} and U is $n \times p$ matrix whose ijth element is u_{ij}. Taking derivatives with respect to the Lagrange multipliers returns the constraints

$$\begin{aligned} Tr(F_j K_j) &= 1 \\ Tr(F_j K_l) &= 0 \qquad l = 1, \ldots, m, l \neq j \\ F_j X &= 0 \qquad \text{and} \\ F_j^T X &= 0 \end{aligned} \tag{13}$$

Solving eq. (12) with respect to F_j

$$F_j = \sum_{l=1}^{m} \lambda_l K_o^{-1} K_l K_o^{-1} + K_o^{-1} V X^T K_o^{-1} + K_o^{-1} X U^T K_o^{-1} \tag{14}$$

Postmultiplying by X and taking into account condition (6a) and premultiplying by X^T and taking into account condition (6b), the following two equations are obtained

$$\sum_{l=1}^{m} \lambda_l K_o^{-1} K_l K_o^{-1} X + K_o^{-1} V X^T K_o^{-1} X + K_o^{-1} X U^T K_o^{-1} X = 0 \tag{15a}$$

$$\sum_{l=1}^{m} \lambda_l X^T K_o^{-1} K_l K_o^{-1} + X^T K_o^{-1} V X^T K_o^{-1} + X^T K_o^{-1} X U^T K_o^{-1} = 0 \tag{15b}$$

These equations can be solved with respect to V and U^T, under the condition that $(X^T K_o^{-1} X)$ is invertible, yielding

$$V = -\sum_{l=1}^{m} \lambda_l [I - X(X^T K_o^{-1} X)^{-1} X^T K_o^{-1}] K_l K_o^{-1} X (X^T K_o^{-1} X)^{-1} \tag{16}$$

$$U^T = -\sum_{l=1}^{m} \lambda_l (X^T K_o^{-1} X)^{-1} X^T K_o^{-1} K_l \tag{17}$$

Substituting in eq. (14)

$$\begin{aligned} F_j = \sum_{l=1}^{m} \lambda_l [K_o^{-1} - K_o^{-1} X (X^T K_o^{-1} X)^{-1} X^T K_o^{-1}] \\ \cdot K_l [K_o^{-1} - K_o^{-1} X (X^T K_o^{-1} X)^{-1} X^T K_o^{-1}] \end{aligned} \tag{18}$$

Finally using the unbiased condition, eq. (8)

$$\begin{aligned} \sum_{l=1}^{m} Tr\{[K_o^{-1} - K_o^{-1} X (X^T K_o^{-1} X)^{-1} X^T K_o^{-1}] \\ \cdot K_l [K_o^{-1} - K_o^{-1} X (X^T K_o^{-1} X)^{-1} X^T K_o^{-1}] K_i\} \lambda_l = \delta_{ij} \qquad i = 1, \ldots, m \end{aligned} \tag{19}$$

where

$$\delta_{ij} = \begin{cases} 1 & \text{if} \quad i = j \\ 0 & \text{if} \quad i \neq j \end{cases} \tag{20}$$

Coefficients $\lambda_1, \ldots, \lambda_m$ are determined by solving this system of m equations. Then F_j is calculated from eq. (18). The solution is given explicitly as follows. Let M be the $m \times m$ symmetric matrix whose klth element is

$$\begin{aligned} M_{kl} = Tr\{[K_o^{-1} - K_o^{-1} X (X^T K_o^{-1} X)^{-1} X^T K_o^{-1}] \\ \cdot K_k [K_o^{-1} - K_o^{-1} X (X^T K_o^{-1} X)^{-1} X^T K_o^{-1}] K_l\} \end{aligned} \tag{21}$$

If the inverse of M exists

$$C = M^{-1} \tag{22}$$

then the Lagrange coefficients $\lambda_1, \ldots, \lambda_m$ can be uniquely determined and F_j is given explicitly below

$$\begin{aligned} F_j = \sum_{l=1}^{m} C_{lj} [K_o^{-1} - K_o^{-1} X (X^T K_o^{-1} X)^{-1} X^T K_o^{-1}] \\ \cdot K_l [K_o^{-1} - K_o^{-1} X (X^T K_o^{-1} X)^{-1} X^T K_o^{-1}] \end{aligned} \tag{23}$$

One may readily verify that the solution satisfies the universality and unbiasedness conditions, eqs. (6) and (8), respectively. Matrix F_j is obviously symmetric. In the case of generalized covariance functions, it is useful to consider the transformation

$$\mathbf{z} = [I - X(X^T K_o^{-1} X)^{-1} X^T K_o^{-1}] \mathbf{y} \tag{24}$$

Vector $\mathbf{z}$ consists of authorized increments because

$$[I - X(X^T K_o^{-1} X)^{-1} X^T K_o^{-1}] X = 0$$

(see Kitanidis, 1983). Then

$$\hat{\theta}_j = \sum_{i=1}^{m} C_{ij} \mathbf{z}^T K_o^{-1} K_i K_o^{-1} \mathbf{z} \tag{25}$$

Consequently, the calculation of $\hat{\theta}_j$ involves only authorized increments.

The ijth element of the covariance matrix of the parameter estimates after accounting for the unbiasedness condition, is

$$E[(\hat{\theta}_i - \theta_i)(\hat{\theta}_j - \theta_j)] = E[Tr(F_i \mathbf{y}\mathbf{y}^T)\, Tr(F_j \mathbf{y}\mathbf{y}^T)] - \theta_i\theta_j \tag{26}$$

Through Gaussian moment factoring for the calculation of fourth moments and accounting for the unbiasedness condition

$$E[(\hat{\theta}_i - \theta_i)(\hat{\theta}_j - \theta_j)] = 2Tr(F_i K F_j K) \tag{27}$$

In the applications, these variances and covariances may be calculated by substituting $K = \Sigma_{i=1}^{m} K_i \hat{\theta}_i$. Thus, the developed methodology yields measures of precision of parameter estimates.

The solution becomes simpler in the special case when only one parameter θ_j is to be estimated, all other parameters being fixed to zero. In this case, no assumptions about K_o are made and

$$\hat{\theta}_j = \mathbf{y}^T K_j^{-1} [I - X(X^T K_j^{-1} X)^{-1} X^T K_j^{-1}]\, \mathbf{y} / Tr[I - X(X^T K_j^{-1} X)^{-1} X^T K_j^{-1}] \tag{28}$$

with variance of estimation

$$\mathrm{Var}\,(\hat{\theta}_j) = 2\theta_j^2 / Tr[I - X(X^T K_j^{-1} X)^{-1} X^T K_j^{-1}] \tag{29}$$

Note that the trace appearing in the denominator is equal to $n - p$, where n is the number of data and p is the number of drift coefficients. Thus

$$\mathrm{Var}\,(\hat{\theta}_j) = 2\theta_j^2 / (n - p) \tag{30}$$

A particular case is the familiar problem of estimating the variance of the residuals in ordinary linear regression. Then $K_j = I$, for uncorrelated homoscedastic residuals, and the estimate given by eq. (28) is simply the sum of the squares of the fitted residuals divided by the degrees of freedom (number of data minus the number of fitted regression coefficients). This estimator is widely used (Draper and Smith, 1966) in least-squares theory regardless of the normality of the data (see also discussion in Rao, 1973, p. 228).

AN ALTERNATE INTERPRETATION OF MVUQ ESTIMATION

Consider a linear transformation of the data

$$\mathbf{z} = W\mathbf{y} \tag{31}$$

where W is a given $n \times n$ transformation matrix. One can always find a W matrix

so that the drift is cancelled out or $\mathbf{z}$ is a vector of authorized increments. According to the model the covariance matrix of $\mathbf{z}$, the average of $W\mathbf{y}\mathbf{y}^T W^T$, is equal to WKW^T. A reasonable approach is to fit the coefficients so that the difference between $W\mathbf{y}\mathbf{y}^T W^T$ and its expected value WKW^T is minimum in some sense. The criterion which minimizes the sum of the squares of all the elements of the matrix

$$W\mathbf{y}\mathbf{y}^T W^T - WKW^T \tag{32}$$

is

$$\min L = Tr[W(K - \mathbf{y}\mathbf{y}^T)\, W^T W(K - \mathbf{y}\mathbf{y}^T)\, W^T]$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} [W(K - \mathbf{y}\mathbf{y})\, W^T]_{ij}^2 \tag{33}$$

The criterion may be written as

$$\min_{\theta} L = Tr[W^T W(K - \mathbf{y}\mathbf{y}^T)\, W^T W(K - \mathbf{y}\mathbf{y}^T)] \tag{34}$$

Given eq. (3), the necessary conditions for the optimum are a linear system of m equations with m unknowns

$$\sum_{j=1}^{m} Tr[W^T W K_i W^T W K_j]\, \theta_j = \mathbf{y}^T W^T W K_i W^T W \mathbf{y} \qquad i = 1, \ldots, m \tag{35}$$

If W is selected so that

$$W^T W = K_o^{-1} - K_o^{-1} X (X^T K_o^{-1} X)^{-1} X^T K_o^{-1} \tag{36}$$

where K_o is an estimate of K (multiplied by an arbitrary constant), the solution is identical to the $MVUQ$ estimation eqs. (21) through (25). However, even if a good estimate of K is not available, the solution would still be a good estimate because it is based on the minimization of a reasonable fitting criterion. For example, in the absence of better information, one may always set $K_o = I$, the identity matrix. This solution, although generally suboptimal, is computationally appealing because it avoids inversion of $n \times n$ matrices.

Also, even if the data are not Gaussian, the solution is still optimal in the sense of minimizing the quadratic fitting criterion of eq. (33). The estimates may not have minimum variance but are unaffected by fitted drift coefficients and are generally unbiased.

DISCUSSION

The analysis provides valuable insight into the problem of quadratic estimation of covariance parameters. In particular, it illustrates that for finite samples minimum-variance and unbiased quadratic estimation can readily be achieved in the single-parameter case. It can also be achieved in the multiple parameter case if the $m - 1$ ratios of the parameters over one of them are known. Then K_o (multiplied by an arbitrary constant) can be determined and F_j can be selected independently of the observations from the solution of a quadratic programming problem. In practice, however, K_o can be selected only as an approximation to the actual covariance matrix K. The obtained parameter estimates are unbiased but only approximately minimum variance (unless, of course, K_o is equal to the actual covariance matrix). These results can be called locally minimum-variance unbiased quadratic estimates. Simulation indicates that they are not sensitive to values of the assumed ratios (in the sense that differences in results are well within the sampling error).

In the case of universal kriging, the most common method for estimation of the variogram (or the corresponding covariance function) apparently is through analysis of residuals obtained from the data by subtracting the fitted drift. However, as Matheron (1971) and others have pointed out, such estimators may be seriously biased. This has been considered to be a major drawback of universal kriging. The presented methodology shows how unbiased estimates of the variogram can be obtained even in the presence of unknown drift parameters. In the author's opinion, this approach removes an objection to the application of universal kriging. After unbiased estimates of the variogram parameters have been obtained, drift parameters may be estimated, if appropriate, using weighted least-squares or kriging.

A practical solution to the problem of unknown K_o is to apply the procedure iteratively by updating at the end of each iteration the estimate of K_o based on the new values of the parameters. This procedure makes the weighting matrices F_j functions of the data so that, strictly speaking, the unbiasedness condition (eq. 7) may not necessarily hold. The parameter estimates may thus be slightly biased. However, this bias should not be confused with the bias caused by the use of fitted drift parameters to estimate the covariance function, as common practice in applications of universal kriging. Nevertheless, it can be shown that the unbiasedness solution should hold for large samples (asymptotically) even without extending the area over which measurements are taken.

Another important issue is the requirement that parameter estimates satisfy certain constraints so that the covariance function be positive definite or conditionally positive definite. The most common constraint type is linear. From the viewpoint of optimization, no difficulty is encountered in accounting for these constraints. However, the constraints make F_j dependent on the data and may interfere with the unbiasedness condition. Nevertheless, as the sample size increases this bias diminishes in magnitude as these constraints become binding less frequently.

SIMULATION RESULTS

The proposed methodology has been tested extensively through computer simulations and applied to the estimation of covariance parameters of real data with encouraging results. Here only some indicative simulation results are presented.

Consider a one-dimensional intrinsic random function of zero order with generalized covariance function.

$$K_{ij} = \theta_1 \delta(|h_{ij}|) - \theta_2 |h_{ij}|$$

where h_{ij} is the distance between points i and j and δ is Kronecker's delta

$$\delta(|h_{ij}|) = \begin{cases} 1 & \text{if} \quad |h_{ij}| = 0 \\ 0 & \text{if} \quad |h_{ij}| > 0 \end{cases}$$

Fifty different realizations of the aforementioned intrinsic random field with parameters

$$\theta_1 = 1$$

$$\theta_2 = 5$$

were generated. Increments of each realization were generated using Gaussian variates. Each realization of the random function is sampled at 30 points, the coordinates of which were selected randomly from a uniform distribution [0, 1]. For each realization, the 30 measurements were employed to determine the value of θ_1 and θ_2. For purposes of comparison four different methods were employed

(1) Minimum-variance unbiased quadratic estimation with assumed $\theta_2/\theta_1 = 5$.
(2) Iterative minimum-variance unbiased quadratic where K_o was adjusted based on the most recent estimates of the parameters and the procedure was repeated.
(3) As in (1), but setting $K_o = I$, the identity matrix. This approach corresponds to the extreme case of assuming $\theta_2/\theta_1 = 0$.
(4) Iterative regression, based on Delfiner's (1976) approach as applied by Kafritsas and Bras (1981).

For all four methods the constraints $\hat{\theta}_1 \geqslant 0$ and $\hat{\theta}_2 \geqslant 0$ were enforced using quadratic programming to minimize the expression of eq. (33) or (34). In the author's experience, this method is preferable to simply setting $\theta_i = 0$ if the algorithm gives $\hat{\theta}_i < 0$. The first constraint became binding in five realizations for method (3). The second constraint became binding in 3, 1, 4, and 30 realizations for methods (1), (2), (3), and (4), respectively. Enforcing the constraints reduced the spread of the estimates, in addition, of course, to yielding acceptable results.

A summary of the 50 estimates of each parameter with each method can be given through five numbers. The smallest and the largest value, the quartiles, and

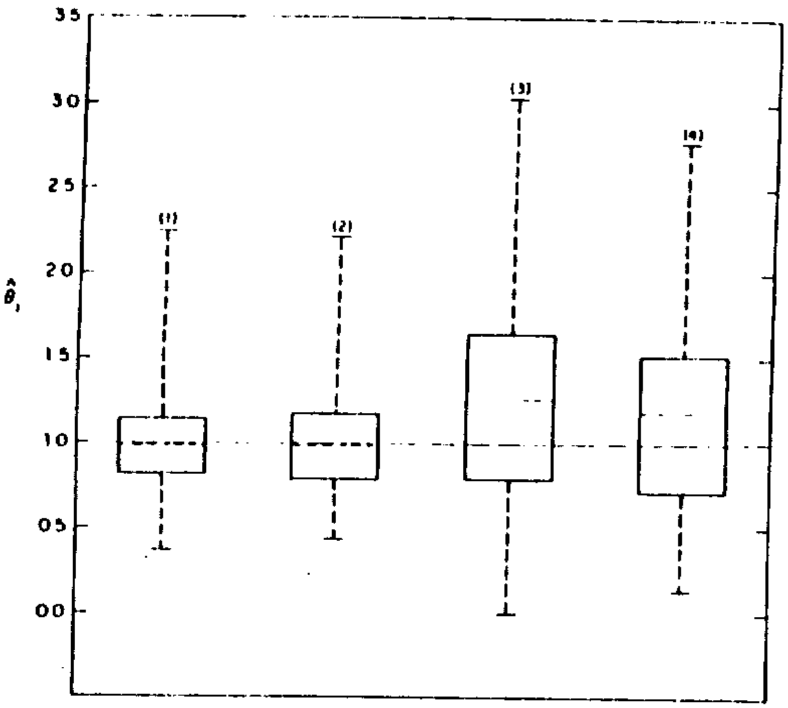

Fig. 1. Median, quartiles, and extreme values of estimates $\hat{\theta}_1$ obtained from 50 realizations using each of the four estimation methods.

the median. Schematic plots of the five-number summaries of the results are given in Figs. 1 and 2. The vertical scale is that of parameter estimates. The solid bars joined to form a box are at the quartiles. The positions of the extreme values are indicated through solid crossbars tied to the box with dashed whiskers. The dashed crossbar inside the box represents the median value. The mean, variance and mean squared error are given in Table 1.

The *MVUQ* estimation and the iterative *MVUQ* estimation (methods 1 and 2) presented in this paper performed markedly better than the other two methods. Although the order with which each method performed varied from realization to realization, the mean square error of estimation was about 3 and 5 times smaller for $\hat{\theta}_1$ and $\hat{\theta}_2$, respectively, for methods (1) and (2) as compared to methods (3) and (4). In fact, (3) and (4) may have given slightly biased estimates of θ_1 as a result of the nonnegativity constraints (see Table 1). When constraints were not enforced, estimates were unbiased but had a larger variance.

The performance of the iterative *MVUQ* estimation is particularly encouraging. This method does not rely on prior estimates of θ_1/θ_2. Even though the method does not guarantee small-sample unbiasedness, no bias is apparent from

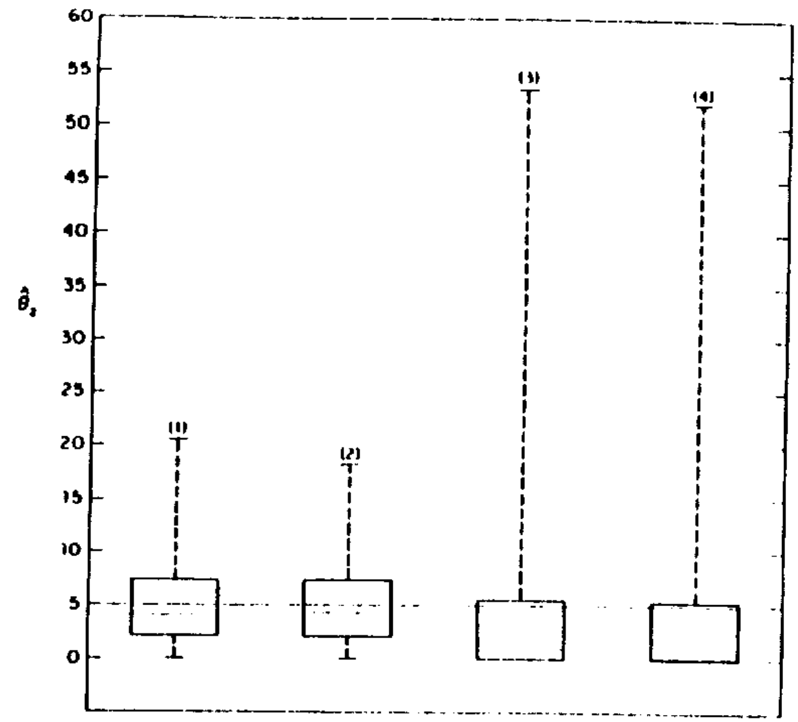

Fig. 2. Median, quartiles, and extreme values of estimates $\hat{\theta}_2$ obtained from 50 realizations using each of the four estimation methods.

the simulations, and the method performs as well as the *MVUQ* estimation with the actual value of θ_1/θ_2.

The results did not change qualitatively when the simulations were repeated first with $\theta_2 = 2$ and then with realizations generated with non-Gaussian variates, although the differences in performance became slightly less pronounced. Preliminary results indicate that the proposed *MVUQ* estimation method which em-

Table 1

	MVUQ		Iterative *MVUQ*		*MN*		Iterative regression	
	θ_1	θ_2	θ_1	θ_2	θ_1	θ_2	θ_1	θ_2
Actual value	1	5	1	5	1	5	1	5
Mean estimate	1.03	5.19	1.04	4.97	1.23	5.12	1.24	4.64
Variance	0.141	17.8	0.146	15.9	0.469	96.1	0.381	98.3
Mean squared error	0.142	17.8	0.148	15.9	0.522	96.1	0.438	98.4

ploys the Gaussian assumption gives reasonable results with non-Gaussian fields. Simulation studies are underway to evaluate fully the proposed methodology.

CONCLUDING REMARKS

Despite some similarities in the formulation, minimum-variance unbiased quadratic estimation of the parameters of ordinary or generalized covariance functions is considerably more difficult than best linear unbiased estimation. Unlike *BLU* estimation, *MVUQ* estimation requires that certain assumptions be made about the third and fourth moment of the data and, except in special cases, depends on some prior estimates of the parameters. If the sample is small and more than one parameter must be estimated, only locally minimum-variance unbiased quadratic estimation is possible. However, unbiased estimation of covariance function or variogram parameters is always possible.

Nevertheless, simulation results indicate that for practical purposes the proposed quadratic estimation procedure can achieve both minimum variance and unbiasedness for moderately large samples. Preliminary results also indicate that results are insensitive to mild violations of the normality assumption. After all, neither the universality nor the unbiasedness conditions depend on the validity of the normality assumption, and results are always optimal in the sense of minimizing a reasonable fitting criterion.

ACKNOWLEDGMENTS

The material is based upon work supported by the National Science Foundation under grant CME-8106577. Additional support was provided by the U.S. Department of the Interior through the Iowa Water Resources Research Institute (Project No. CT 381702). Useful comments made by anonymous reviewers and the editor on an earlier version of the paper are appreciated.

REFERENCES

Athans, M. and Schweppe, F. G., 1965, Gradient matrices and matrix calculations: Lincoln Laboratory, MIT Technical Note 1965-53, 34 p.

Cressie, N. and Hawkins, D. M., 1980, Robust estimation of the variogram. I, Jour. Math. Geol., v. 12, n. 2, p. 115-125.

Davis, M. W. D. and David, M., 1978, Automatic kriging and contouring in the presence of trends: Jour. Canad. Pet. Tech., January-March, p. 90-99.

Delfiner, P., 1976, Linear estimation of nonstationary spatial phenomena, *in* M. Guarascio, M. David, and C. Huijbregts (Eds.), Advanced geostatistics in the mining industry: Dordrecht, Holland, K. Reidel Publishing Co., p. 49-68.

Dubrule, O., 1983, Two methods with different objectives: Splines and kriging: Jour. Math. Geol., v. 15, n. 2, p. 245-257.

Draper, N. R. and Smith, H., 1966, Applied regression analysis: New York, John Wiley & Sons, Inc., 407 p.

Gelhar, L. W. and Axness, C. L., 1983, Three-dimensional stochastic analysis of macrodispersion in aquifers: Water Resour. Res., v. 19, n. 1, p. 161-180.

Journel, A. G. and Huijbregts, C. J., 1978, Mining geostatistics: New York, Academic Press, 600 p.

Kafritsas, J. and Bras, R. L., 1981, The practice of kriging: MIT, R. M. Parsons Laboratory, Technical Report Number 263, 107 p.

Kitanidis, P. K. and Vomvoris, E. G., 1983, A geostatistical approach to the inverse problem in groundwater modeling (steady state) and one-dimensional simulations: Water Resour. Res., v. 19, n. 3, p. 677-690.

Kitanidis, P. K., 1983, Statistical estimation of polynomial generalized covariance functions and hydrologic applications: Water Resour. Res., v. 19, n. 4, p. 901-921.

Matheron, G., 1971, The theory of regionalized variables and its applications: Fontainbleau, France, Ecole de Mines, 212 p.

Matheron, G., 1973, The intrinsic random functions and their applications: Adv. Appl. Prob., v. 5, p. 439-468.

Olea, R. A., 1975, Measuring spatial dependence with semivariograms: Series on spatial analysis no. 3, Kansas Geological Survey, 24 p.

Rao, C. R., 1971, Minimum variance quadratic unbiased estimation of variance components: Jour. Multivar. Anal., v. 1, p. 445-456.

Rao, C. R., 1973, Linear statistical inference and its applications: John Wiley & Sons, New York, 625 p.

Starks, T. H. and Fang, J. H., 1982, On the estimation of the generalized covariance function: Jour. Math. Geol., v. 14, n. 1, p. 57-64.