An Application of the Geostatistical Approach to the Inverse Problem in Two-Dimensional Groundwater Modeling

ROBERT J. HOEKSEMA and PETER K. KITANIDIS

Institute of Hydraulic Research, The University of Iowa

The geostatistical approach to the estimation of transmissivity from head and transmissivity measurements is developed for two-dimensional steady flow. The field of the logarithm of transmissivity (log-transmissivity) is represented as a zero-order intrinsic random field; its spatial structure is described in this application through a two-term covariance function that is linear in the parameters \( \theta_1 \) and \( \theta_2 \). Linearization of the discretized flow equations allows the construction of the joint covariance matrix of the head and log transmissivity measurements as a linear function of \( \theta_1 \) and \( \theta_2 \). In this particular application, the covariance matrices are calculated numerically in a noniterative fashion. Maximum likelihood estimation is employed to estimate \( \theta_1 \) and \( \theta_2 \) as well as additional parameters from measurements. Linear estimation theory (cokriging) then yields point or block-averaged estimates of transmissivity. The approach is first applied to a test case with favorable results. It is shown that the application of the methodology gives good estimates of transmissivities. It is also shown that when the transmissivities are used in a numerical model they reproduce the head measurements quite well. Results from the application of the methodology to the Jordan aquifer in Iowa are also presented.

1. INTRODUCTION

Computer-based groundwater models employing finite differences or finite elements are often used to simulate the response of the hydraulic head to aquifer excitations, hydrogeologic parameters, and aquifer boundary conditions. Aquifer excitations include pumping, accretion, and evapotranspiration. The pertinent hydrogeologic parameters are permeability or transmissivity and storativity.

One of the difficulties encountered in applying these numerical techniques on a regional (or basinwide) basis is obtaining good estimates of hydrogeologic parameters. To overcome the very real data limitations, current methods of hydrogeologic parameter estimation rely heavily on extracting information from the observed input-output behavior of the system. This is accomplished through the selection of a set of parameters that reproduce, through a numerical model, measured values of hydraulic head, solute concentration, and other measurable system variables. This method has been widely applied as a manual trial and error procedure. Recently, a number of systematic, computerized methods have been proposed. In general these methods treat the parameters as dependent variables in a new problem called the inverse problem of groundwater modeling.

This paper deals with the estimation of transmissivities following the geostatistical solution to the inverse problem. The basic methodology used has been presented in Kitanidis and Vomvoris [1983], where the application of the methodology has been illustrated through one-dimensional simulations. This work applies the methodology to two-dimensional steady-flow fields with emphasis on the use of numerical methods for the calculation of the first two joint moments of hydraulic head and the logarithm of the transmissivity.

The purpose of this paper is to illustrate in detail how the geostatistical approach can be applied to estimate transmissivities in the case of numerical models utilizing the discrete flow equations. While the geostatistical approach is very general, the application described herein deals with steady flow with Dirichlet (given head) boundary conditions and no pumping or accretion terms. The methodology developed herein can readily be applied with most finite-difference or finite-element steady-flow models used in the study of regional groundwater systems and is versatile and computationally efficient. The methodology is tested extensively through simulations and applied to an actual case study.

2. OVERVIEW OF THE METHODOLOGY

The general approach has been described in detail in Kitanidis and Vomvoris [1983], and only a summary will be provided here. (A reviewer brought to our attention that a similar approach has been applied by Carotuoto and Racco [1981].) The natural logarithm of the transmissivity (referred to here as log-transmissivity) and the piezometric head are represented as random fields (or random functions). The first step of the general geostatistical approach involves the selection of a model for the structure of the random field of the log-transmissivity. While a wide variety of models are possible, the one used in this work is that of an intrinsic random field [Matheron, 1971]. The second step involves relating the random structure of the hydraulic head to the assumed random structure of the log-transmissivity. This can be accomplished through analytically derived relationships, Monte Carlo simulation, or linearization of the flow equation. The third approach is applied in this work. The third step involves estimation of the unknown statistical parameters of the assumed log-transmissivity model based on measurements of log-transmissivity and hydraulic head. After testing the validity of the fitted model assumed above, minimum variance unbiased linear estimation (or cokriging) is used to predict point or block-averaged values of the field log-transmissivity anywhere in the model.

The geostatistical approach to the inverse problem is very general. The particular random field models used in this work for log-transmissivity and hydraulic head are definitely not the only ones possible for application of the geostatistical approach. Their choice was based on the fact that they are flexible enough to represent many real fields and yet produce
good results with a minimum of computational effort. One should always take care to select a model that is appropriate for a given field.

3. APPLICATION

Select Geostatistical Structure and Estimate Parameters

According to field data reviewed by Freeze [1975], the probability density function of permeabilities, and consequently transmissivities, taken from point samples is approximately log-normal. For this reason it will be assumed that the log-transmissivities are normally distributed. The autocovariance model will be selected so that it is simple and computationally tractable but also versatile enough to be able to represent the random field variability at each scale of interest. The log-transmissivity field will be modeled in this work as a zero-order intrinsic random field. A useful model of the generalized covariance function of log-transmissivity is

\[
\text{Cov}(Y_i, Y_j) = \theta_1 \delta_{ij} + \theta_2 d_{ij}
\]

(1)

where \( \delta_{ij} \) is Kronecker's delta (1 if \( i = j \); 0 otherwise); \( d_{ij} \) is the scalar distance between the locations of the point log-transmissivity values \( Y_i \) and \( Y_j \); and \( \theta_1 \geq 0, \theta_2 \leq 0 \) are parameters to be estimated. The first term represents small-scale variability and unobservable measurement error ("nugget effect"), while the second term represents large-scale variability and describes the spatial interdependence of log-transmissivities. Small-scale or white-noise variability is essentially variability at a scale small enough to have no significant effect on the regional variation in hydraulic head.

The two geostatistical parameters, \( \theta_1 \) and \( \theta_2 \), could be estimated by using only measurements of log-transmissivity \( Y \). Following a more general approach, measurements of the log-transmissivity as well as measurements of hydraulic head \( \phi \) will be used to estimate the parameters. The conditions under which the information included in the hydraulic head measurements improves the estimation procedure will be investigated through numerical simulations and a case study.

The complete use of information is accomplished by first constructing the measurement (both head and log-transmissivity) generalized covariance matrix as a function of the parameters \( \theta_1 \) and \( \theta_2 \). Note that the term "measurement covariance matrix" refers to the covariance matrix of the measurement vector. Next, maximum-likelihood parameter estimation is used to estimate \( \theta_1 \) and \( \theta_2 \), based on the actual measurements. That is, the estimates \( \hat{\theta}_1 \) and \( \hat{\theta}_2 \) are the ones that maximize the probability that the given measurements would occur.

The first step mentioned above is the construction of the measurement generalized covariance matrix \( Q \). For computational efficiency the goal is to obtain \( Q \) as a linear function of the parameters of the log-transmissivity generalized covariance given in (1), i.e., \( Q = \theta_1 R + \theta_2 S + P \), where \( (R, S, \text{and} P) \) are matrices of known constants. The \( Q \) matrix can be represented in partitioned form as follows:

\[
Q = \begin{bmatrix}
Q_{\phi\phi} & Q_{\phi T} \\
Q_{T\phi} & Q_{TT}
\end{bmatrix} = \theta_1 \begin{bmatrix}
R_{\phi\phi} & R_{\phi T} \\
R_{T\phi} & R_{TT}
\end{bmatrix} + \theta_2 \begin{bmatrix}
S_{\phi\phi} & S_{\phi T} \\
S_{T\phi} & S_{TT}
\end{bmatrix} + \begin{bmatrix}
P_{\phi\phi} & 0 \\
0 & 0
\end{bmatrix}
\]

(2)

where \( Q_{\phi\phi} \) is head measurement covariance matrix, \( Q_{\phi T} \) is head log-transmissivity measurement generalized covariance matrix, \( Q_{TT} \) is log-transmissivity measurement generalized covariance matrix (given by (1)), and \( Q_{T\phi} = Q_{\phi T}^T \). Allowing the head covariance to be a linear function of \( \theta_2 \) is valid only for small values of \( \theta_2 \), as will be seen in later discussion. While other approaches, such as Monte Carlo simulation, can be used, the linearized approach is computationally very efficient and is the most appropriate, provided that \( \theta_2 \) is small enough. In general, nonlinear estimators may get better results but are much more difficult to use.

The submatrices \( R_{\phi\phi} \) and \( S_{\phi\phi} \) are readily evaluated via (1). To calculate matrices that include \( \phi \) in their subscripts, the partial differential equation relating piezometric head to log-transmissivity must be introduced as in the works of Freeze [1975], Smith and Freeze [1979a, b], Baker et al. [1978], Dagan [1979, 1981], Chiling and Dagan [1980], Gutjahr and Gelhar [1981], Mizell et al. [1982], and others. In this work a methodology appropriate for use with numerical models will be developed and applied.

The scope of this paper will be limited to steady flow in a two-dimensional nonhomogeneous isotropic confined aquifer with no inflow and prescribed head as boundary conditions. The submatrices \( R_{\phi\phi}, \) \( S_{\phi\phi}, \) and \( R_{TT} \) are all zero, since small-scale variability in the transmissivity does not affect head variability when the boundary conditions are constant head. The hydraulic head \( \phi \) and the aquifer transmissivity \( T \) must satisfy the following partial differential equation

\[
\frac{\partial T}{\partial x} \frac{\partial \phi}{\partial x} + \frac{\partial T}{\partial x} \frac{\partial \phi}{\partial y} + \frac{\partial T}{\partial y} \frac{\partial \phi}{\partial y} + \frac{\partial T}{\partial y} \frac{\partial \phi}{\partial y} = 0
\]

(3)

In terms of the log-transmissivity \( Y (Y = \ln T) \), (3) is

\[
\frac{\partial Y}{\partial x} \frac{\partial \phi}{\partial x} + \frac{\partial Y}{\partial x} \frac{\partial \phi}{\partial y} + \frac{\partial Y}{\partial y} \frac{\partial \phi}{\partial y} + \frac{\partial Y}{\partial y} \frac{\partial \phi}{\partial y} = 0
\]

(4)

Note that the emphasis of this work is on zero-order intrinsic functions to describe the structure of log-transmissivity. To separate (4) into a deterministic and a stochastic form, the following substitutions are made:

\[
\phi = H + \eta
\]

and

\[
Y = F + \xi
\]

where \( H \) is expected value (or drift) of the head field; \( \eta \) is zero-mean head perturbation; \( F \) is aquifer expected \( Y \) value, a constant for zero-order intrinsic function; and \( \xi \) is zero-mean log-transmissivity perturbation. After substitution, (4) becomes

\[
\frac{\partial \xi}{\partial x} \left( \frac{\partial H}{\partial x} + \frac{\partial \xi}{\partial x} \right) + \frac{\partial \xi}{\partial x} \left( \frac{\partial H}{\partial y} + \frac{\partial \xi}{\partial y} \right) + \frac{\partial \xi}{\partial y} \left( \frac{\partial H}{\partial y} + \frac{\partial \xi}{\partial y} \right) + \frac{\partial \xi}{\partial y} \left( \frac{\partial H}{\partial y} + \frac{\partial \xi}{\partial y} \right) = 0
\]

(5)

If the variations in the transmissivity are small, then second-order terms (products of perturbations) may be neglected [see also Mizell et al., 1982, p. 1054; Dagan, 1982, p. 819]. Note that while (5) implicitly assumes that \( \xi \) is differentiable (in a mean square sense), this assumption is not necessary for the calculation of \( \xi \) (see Appendix A). Linearization analysis holds when \( \theta_2 L \), where \( L \) is the maximum aquifer length, is "small." Several numerical tests have indicated that the linearized analysis gives excellent results for \( \theta_2 L \) up to order 1. After elimination of second-order terms, (5) becomes

\[
\frac{\partial \xi}{\partial x} \left( \frac{\partial H}{\partial x} + \frac{\partial \xi}{\partial x} \right) + \frac{\partial \xi}{\partial x} \left( \frac{\partial H}{\partial y} + \frac{\partial \xi}{\partial y} \right) + \frac{\partial \xi}{\partial y} \left( \frac{\partial H}{\partial y} + \frac{\partial \xi}{\partial y} \right) + \frac{\partial \xi}{\partial y} \left( \frac{\partial H}{\partial y} + \frac{\partial \xi}{\partial y} \right) = 0
\]

(6)
By taking the expected value of both sides of (6), the following deterministic partial differential equation is obtained:

\[
\frac{\partial^2 H}{\partial x^2} + \frac{\partial^2 H}{\partial y^2} = 0
\]  

(7)

Subtracting (7) from (6) yields the stochastic equation

\[
\frac{\partial^2 h}{\partial x^2} + \frac{\partial^2 h}{\partial y^2} = - \frac{\partial f}{\partial x} \frac{\partial H}{\partial x} - \frac{\partial f}{\partial y} \frac{\partial H}{\partial y}
\]  

(8)

Equation (8) is next written in finite difference form

\[
h_{j-1} + h_{i-1} - 4h_i + h_{j+1} + h_{i+1} = -\frac{\Delta x}{2} (f_{j+1} - f_{j-1}) \frac{\partial H}{\partial x} - \frac{\Delta x}{2} (f_{i+1} - f_{i-1}) \frac{\partial H}{\partial y}
\]  

(9)

where the subscripts indicate the relative location of the finite-difference blocks associated with each variable. For simplicity, square finite-difference blocks with dimensions \(\Delta x\) are used in this analysis. Figure 1 shows the relative location of the head perturbation variable \(h\).

The (prior) expected value of the piezometric head \(H\) is calculated by solving (7), using a numerical model with constant transmissivity and the prescribed boundary conditions. The expected-head slopes \(\partial H/\partial x\) and \(\partial H/\partial y\) in (9) can then be approximated by using the discrete values of \(H\) resulting from the solution of (7). Note that the expected heads can be evaluated at a different level of discretization than that for \(h\) and \(f\). This allows use of previously developed models to determine the average head field.

Equation (9) can then be written for each block. The head perturbations for blocks on the constant head boundary can be placed on the right side of (9). In matrix form this set of equations is

\[
A b = B f + C b_h
\]

where \(A\), \(B\), and \(C\) are constant matrices, \(b\) is the vector of nonboundary block head perturbations, \(f\) is the vector of block-averaged log-transmissivity perturbations, and \(b_h\) is the vector of boundary block head perturbations. When perfectly known constant head boundaries are considered, \(b_h\) is a zero vector. In this analysis, \(b_h\) is included to account for possible uncertainty in the estimates of the head along the boundary.

Solving for \(h\) yields

\[
h = (A^{-1} B) f + (A^{-1} C) b_h
\]  

(10)

Note that the form of (10) would not be affected if a finite-element instead of a finite-difference scheme were chosen as a point of departure (equation (9)).

The covariance matrix for measurement-point heads is

\[
Q_{pp} = E[ h_p h_p^T ]
\]  

(11)

where \(h_p\) is the vector of measurement-point head perturbations, i.e., measured head minus expected head, and \(E[ \ ]\) denotes the expectation operator. Similarly, the head log-transmissivity measurement cross-covariance is

\[
Q_{pf} = E[ h_p f_p^T ]
\]  

(12)

where \(f_p\) is the vector of measurement point log-transmissivity perturbations.

Each head perturbation at a measurement point can be expressed approximately as a linear combination of the head perturbations at the centers of the four finite-difference blocks nearest to it. If \(h_p\) is a measurement point head perturbation located as shown in Figure 2, and \(h_1, h_2, h_3,\) and \(h_4\) are the finite-difference head perturbations at the centers of the four blocks nearest to the particular head measurement point, then \(h_p\) is expressed in terms of nodal values through the following simple interpolation scheme:

\[
h_p = \left( \frac{\Delta x - u}{\Delta x} \right) \left( \frac{\Delta y - u}{\Delta y} \right) h_1 + \left( \frac{\Delta x - v}{\Delta x} \right) \left( \frac{\Delta y - v}{\Delta y} \right) h_2
\]

\[
+ \left( \frac{\Delta x - v}{\Delta x} \right) \left( \frac{\Delta y - u}{\Delta y} \right) h_3 + \left( \frac{\Delta x - u}{\Delta x} \right) \left( \frac{\Delta y - v}{\Delta y} \right) h_4
\]  

(13)

Given that the piezometric head is a smooth and continuous function, this interpolation scheme is adequate as long as the block size \(\Delta x\) is reasonably small. The vector of measurement point head perturbations \(b_p\) can now be expressed as

\[
b_p = W_1 h + W_2 b_h
\]  

(14)

where \(W_1\) and \(W_2\) are transformation matrices of constants.

Using (14) and (10), (11) and (12) yield

\[
Q_{pp} = (W_1 A^{-1} B) E[f f^T] (W_1 A^{-1} B)^T
\]

\[
+ (W_1 A^{-1} C + W_2) E[b_h b_h^T] (W_1 A^{-1} C + W_2)^T
\]  

(15)
and

\[ Q_{yy} = (W_1 A^{-1} B) \delta (W_1 A^{-1} B)^T \]  

(16)

Note that it has been assumed that \( E[h_y^T] = 0 \) and \( E[h_y h_y^T] = 0 \).

In (16), \( E[h_y^T] \) is the generalized cross-covariance matrix for finite-difference block-averaged log-transmissivities and measurement point log-transmissivities, while (1) gives the generalized log-transmissivity covariance model for any two points in the field. To obtain the generalized covariance between a point log-transmissivity and a block log-transmissivity, (1) must be integrated over the block. If \( E[f_{lml}] \) represents generalized log-transmissivity covariance between block \( k \) and measurement point \( l \), then

\[ E[f_{lml}] = \theta_i d_{ml} \]

where \( d_{ml} \) is the average distance between block \( k \) and point \( l \). If \( \delta \) is the matrix of block to measurement point average distances, \( (\delta(k,l) = d_{ml}) \), then

\[ E[h_y^T] = \theta_i \delta \]

(17)

In (15), \( E[h_y^T] \) is the generalized covariance matrix for block log-transmissivities. If \( E[f_{lml}] \) is the generalized log-transmissivity covariance for blocks \( k \) and \( l \), then, as before,

\[ E[f_{lml}] = \theta_i d_{ml} \]

where \( d_{ml} \) is now the average distance between block \( k \) and block \( l \). If \( \delta \) is the matrix of block to block average distances, then

\[ E[h_y^T] = \theta_i \delta \]

(18)

The methods used to evaluate \( \delta \) and \( \delta \) are given in Appendix B. Note that the expressions for \( E[h_y^T] \) and \( E[h_y^T] \) do not involve \( \theta_i \), the parameter that represents small-scale log-transmissivity variability.

In (15), \( E[h_y h_y^T] \) is the covariance matrix for the block boundary heads. The existence of this covariance matrix is due to the modeler's uncertainty about the assumed head along the boundary. One possible way to evaluate this matrix is to assume some reasonable value of the boundary head variance \( \text{Var} (\phi_h) \) and then assume that the correlation decays exponentially with the distance. That is, the covariance between boundary head \( i \) and boundary head \( j \) is given by

\[ E[h_i h_j] = \text{Var} (\phi_h) \exp \left( - \frac{d_{ij}}{l} \right) \]

where \( d_{ij} \) is the scalar distance between the two points and \( l \) is approximately the distance at which the correlation disappears. This approach has been followed in this work. An alternative approach would be to consider \( \text{Var} (\phi_h) \) and \( l \) as parameters to be estimated from the data.

Equations (15) and (16) can now be combined with (17) and (18) to yield

\[ Q_{xy} = \theta_i (W_1 A^{-1} B) \delta (W_1 A^{-1} B)^T \]

\[ + (W_1 A^{-1} C + W_2) E[h_y h_y^T] (W_1 A^{-1} C + W_2)^T \]

and

\[ Q_{yy} = \theta_i (W_1 A^{-1} B) \delta \]

Therefore, the remaining four submatrices in (2) can be expressed as

\[ S_{xy} = (W_1 A^{-1} B) \delta (W_1 A^{-1} B)^T \]

\[ S_{yy} = (W_1 A^{-1} B) \delta \]

\[ P_{yy} = (W_1 A^{-1} C + W_2) E[h_y h_y^T] (W_1 A^{-1} C + W_2)^T \]

\[ S_{xy} = S_{yy}^T \]

Variability caused by measurement error can also be accounted for. Since measurement errors can reasonably be assumed to be independent of each other, log-transmissivity measurement error is already included in \( \theta_i \). The head measurement error can be included in the measurement covariance matrix by simply adding a diagonal matrix to \( P_{yy} \):

\[ P_{yy} = (W_1 A^{-1} C + W_2) E[h_y h_y^T] (W_1 A^{-1} C + W_2)^T + \sigma^2 I \]

(19)

In (19), \( I \) is the identity matrix and \( \sigma^2 \) is the head variance caused by measurement error. This head variance is determined either by considering the accuracy of methods used to measure the water surface depth and the ground elevation at the well, as well as the numerical error of the numerical scheme, or can be considered as a parameter to be estimated from the data.

At this point, all of the submatrices in (2) have been accounted for. The next step, then, involves estimation of \( \theta_i \) and \( \theta_j \), given the measurements. Note that \( \theta_i \) can still be estimated, since \( R_{yy} \) in (2) is nonzero.

First, the measurement vector and the generalized measurement covariance matrix must be transformed so that increments instead of point values of log-transmissivity appear in the measurement vector. Thus the measurement vector expected values will be known (the mean of a log-transmissivity increment is zero), and the covariance will become positive definite. This is done by subtracting one of the log-transmissivity measurements from each of the others. If \( z \) is the vector of \( n \) head measurements and \( m \) log-transmissivity measurements

\[ [\phi_1, \phi_2, \ldots, \phi_m, y_1, y_2, \ldots, y_m]^T \]

then the modified measurement vector \( z_{mod} \) is

\[ z_{mod} = [\phi_1, \phi_2, \ldots, \phi_m, y_1 - y_m, y_2 - y_m, \ldots, y_{m-1} - y_m]^T \]

This transformation can be represented as

\[ z_{mod} = Uz \]

where \( U \) is the transformation matrix. The covariance matrix for the modified measurements \( z_{mod} \) can then be expressed as

\[ Q_{mod} = UQU^T \]

Note that \( Q_{mod} \) is positive definite and, consequently, an appropriate covariance function.

The parameter estimation procedure can be summarized as follows. The \( n \) measurement point expected heads \( H_1, H_2, \ldots, H_n \) are determined via a numerical simulation model of the aquifer with the assumed boundary conditions and a constant transmissivity. This produces expected head values for discrete points throughout the aquifer. The measurement point values are then obtained by using an interpolation procedure similar to that of (13). Next, the measurement covariance matrix is constructed as a function of \( \theta_1 \) and \( \theta_2 \), and the maximum-likelihood parameter estimation procedure is used to determine the most probable values of \( \theta_1 \) and \( \theta_2 \), given the
measurements [see Kitanidis and Vomvoris, 1983; Kitanidis and Lane, 1984].

The basic computational advantage of the proposed approach, in which the average log-transmissivity is not estimated and the covariance function is linear in the parameters, is that the numerical models for the calculation of the expected head vector and the covariance matrix (which is the part of the procedure requiring most of the computer time and storage) need only be run once.

Minimum Variance Unbiased Linear Estimation of Hydrogeologic Parameters (Cokriging)

Once the statistical parameters of the log-transmissivity covariance have been estimated, the next step is to use this information to predict values of log-transmissivity at any point in the field. Linear, minimum variance, unbiased estimation theory, or cokriging, will be employed. Using this procedure, the estimate of the log-transmissivity at any point or over any block in the field $\hat{Y}_0$ is considered to be a linear combination of the $n$ head measurements $\phi_1, \phi_2, \ldots, \phi_n$ and the $m$ log-transmissivity measurements $Y_1, Y_2, \ldots, Y_m$ plus a constant $\xi$:

$$\hat{Y}_0 = \sum_{i=1}^{n} \mu_i \phi_i + \sum_{j=1}^{m} \lambda_j Y_j + \xi$$

If $Y_0$ is the true value of log-transmissivity at the point or block of estimation, then the variance of the estimate is $E[(Y_0 - \hat{Y}_0)^2]$ and can be expressed as

$$E[(Y_0 - \hat{Y}_0)^2] = E\left[\left(Y_0 - \sum_{i=1}^{n} \mu_i \phi_i - \sum_{j=1}^{m} \lambda_j Y_j - \xi\right)^2\right]$$

This relation is expanded to yield

$$E[(Y_0 - \hat{Y}_0)^2] = \sigma_y^2 + \xi^2 + F^2\left[1 - \sum_{j=1}^{m} \lambda_j \right]^2 - 2F\left[1 - \sum_{j=1}^{m} \lambda_j \right]\left[\xi + \sum_{i=1}^{n} \mu_i H_i\right]$$

$$+ \left[2\xi + \sum_{i=1}^{n} \mu_i H_i\right]\left[\sum_{j=1}^{m} \mu_j H_j\right] - 2\sum_{i=1}^{n} \mu_i \text{Cov}(\phi_i, Y_0)$$

$$- 2\sum_{j=1}^{m} \lambda_j \text{Cov}(Y_j, Y_0) + \sum_{j=1}^{m} \sum_{k=1}^{m} \mu_i \mu_k \text{Cov}(\phi_i, \phi_k)$$

$$+ 2\sum_{j=1}^{m} \sum_{j'=1}^{m} \mu_j \lambda_{j'} \text{Cov}(\phi_i, Y_{j'}) + \sum_{j=1}^{m} \sum_{j'=1}^{m} \lambda_j \lambda_{j'} \text{Cov}(Y_j, Y_{j'})$$

$$= \left(1 - \sum_{j=1}^{m} \lambda_j\right) + 2\xi + \sum_{i=1}^{n} \mu_i H_i$$

(20)

where $\sigma_y^2$ is the generalized variance of the log-transmissivity and $F$ is the mean log-transmissivity. To eliminate the dependence of $\hat{Y}_0$ on $F$, i.e., to make $\hat{Y} - \hat{Y}_0$ a generalized increment, the term $\left[1 - \sum_{j=1}^{m} \lambda_j\right]$ must be set to zero. This yields the condition

$$\sum_{j=1}^{m} \lambda_j = 1$$

(21)

Next, the estimate variance must be minimized. Define $Z$ as the following:

$$Z = E[(Y_0 - \hat{Y}_0)^2] + 2\xi\left(1 - \sum_{j=1}^{m} \lambda_j\right) + \sum_{i=1}^{n} \lambda_i \nu_i \phi_i$$

where $\nu$ is the Lagrange multiplier. The estimate variance will be minimum, subject to the constraint of (21), when the following is true:

$$\frac{\partial Z}{\partial \lambda_i} = 0 \quad i = 1, \ldots, m$$

$$\frac{\partial Z}{\partial \nu_i} = 0$$

Applying the above relation results in the following set of equations:

$$\sum_{i=1}^{n} \mu_i Q_{\phi}(i, k) + \sum_{j=1}^{m} \lambda_j Q_{\phi}(k, j) = \text{Cov}(\phi_i, Y_0)$$

(22)

$$\sum_{i=1}^{n} \mu_i Q_{\phi}(i, l) + \sum_{j=1}^{m} \lambda_j Q_{\phi}(l, j) + \nu = \text{Cov}(Y_l, Y_0)$$

(23)

$$\sum_{j=1}^{m} \lambda_j = 1$$

(24)

(the subscripts in the parenthesis associated with the $Q$ matrices refer to the row and column of an element in the particular matrix). The right sides of (22) and (23) are determined by using techniques similar to those presented earlier in this work. If $w_1$ and $w_2$ are constant transformation vectors (similar to $W_1$ and $W_2$) such that

$$h_k = w_1^T h + w_2^T h$$

then the following relationships are valid for the generalized covariance terms on the right sides of (22) and (23).

$$\text{Cov}(\phi_0, Y_0) = \begin{cases} \theta_1 w_1^T A^{-1} \bar{d}_{\phi0} & \text{(point estimation)} \\ \theta_2 w_1^T A^{-1} \bar{d}_{\phi0} & \text{(block estimation)} \end{cases}$$

(25)

$$\text{Cov}(Y_l, Y_0) = \begin{cases} \theta_1 d_{\phi0} + \theta_2 d_{\phi0} & \text{(point estimation)} \\ \theta_2 d_{\phi0} & \text{(block estimation)} \end{cases}$$

(26)

In (25), $\bar{d}_{\phi0}$ and $\bar{d}_{\phi0}$ are vectors of average distances from each block represented in the $f$ vector to the estimation point ($\bar{d}_{\phi0}$) or the estimation block ($\bar{d}_{\phi0}$). Similarly, $d_{\phi0}$ is the distance from measurement point 0 to the estimation point and $d_{\phi0}$ is the average distance from the measurement point to the estimation block.

Equations (22), (23), and (24) represent a set of $n + m + 1$ linear equations in the $n + m + 1$ unknowns

$$\mu_1, \ldots, \mu_n, \lambda_1, \ldots, \lambda_m, \nu$$

This set of equations can be readily solved at each point or block where the log-transmissivity is to be estimated. Note
that only the right sides change for each new estimation location. This allows great computational efficiency, since the matrix of left side constants needs to be factored only once. Note also that (20) provides an estimate of cokriging error.

4. Testing the Numerical Procedure

The described two-dimensional numerical application of the geostatistical approach to the inverse problem will be tested in this section through simulations. A log-transmissivity field and the attendant head field will be generated such that the log-transmissivities have known statistical properties and such that the heads would satisfy the pertinent partial differential equation subject to prescribed head boundary conditions and the given generated transmissivities. The numerical procedure will be tested to check its ability to reproduce the generated transmissivity field and to suggest the appropriate level of discretization needed to produce acceptable results. The numerical procedure that generates the measurement data will first be described.

The field of log-transmissivities is generated in two steps. First, the block-averaged log-transmissivities that will be employed in a finite-difference scheme to determine the hydraulic head are generated. Note that the nugget variability term drops out when averaging log-transmissivities over a finite block. Small-scale variability of log-transmissivity also has no effect on hydraulic head calculated with prescribed head boundary conditions. Consequently, in the generation of block-averaged log-transmissivities, small-scale variability must not be generated. The covariance model of block-averaged, log-transmissivities is reproduced by (1) with \( \theta_i \) set to zero and with \( d_{ij} \) taken as the average distance between blocks. The heads calculated with the block-averaged log-transmissivities are considered to be point measurements at the center of the finite-difference blocks.

In the second step, point measurements of log-transmissivities at block centerpoints are generated so that they have covariance given by (1) but conditional on the block-averaged values already generated. Note that the variability of point measurements is larger than the variability of block-averaged quantities. The data generation procedure for the block-averaged measurements is as follows.

The generalized covariance matrix of the log-transmissivities \( E[\theta^T] \) is first formed. Next, this log-transmissivity covariance matrix is transformed into a covariance matrix of log-transmissivity increments. If \( M \) log-transmissivities are to be generated, then the log-transmissivity increment \( X_i \) is defined as

\[
X_i = Y_i - Y_{\mu} = f_i - f_{\mu}
\]

This increment has the following properties:

\[
E[X_i] = 0
\]

and

\[
\text{Cov} (X_i, X_j) = E[f_i f_j] - E[f_i f_{\mu}] - E[f_j f_{\mu}] + E[f_{\mu} f_{\mu}]
\]

(26)

The log-transmissivity increment covariance matrix \( G \), defined by (26), is decomposed next into a lower triangular matrix \( V \), such that

\[
G = V V^T
\]

A vector of \( M - 1 \) normal random deviates, \( s \), is then generated so that the \( M - 1 \) log-transmissivity increments \( X \) can be computed as

\[
X = V s
\]

Evaluated as such, \( X \) has the correct first two moments:

\[
E[X] = V E[s] = 0
\]

and

\[
E[XX^T] = V E[ss^T] V^T = V V^T = G
\]

These increments, along with a single zero increment, are finally added to a constant (representing the \( M \)th log-transmissivity) to produce the \( M \) block log-transmissivity values needed.

The aquifer model used as a test case is square, with each side being 300 km long. The head is prescribed on the entire boundary. Figure 3 shows the expected head (or head drift) contours for this aquifer. Under the conditions set forth in this paper, the expected heads are essentially the heads that would result if the aquifer had a constant transmissivity. These head values, shown in Figure 3, are generated by using the USGS aquifer modeling computer program [Trescott et al., 1976].

Fig. 4. Fine discretization of test aquifer.
Figures 4, 5, and 6 show the three levels of discretization that will be used to investigate the dependence of block size on the numerical procedure. These three levels of discretization are designed so that the block centerpoints in the coarse and medium meshes coincide with block centerpoints in the fine mesh.

The first test of the numerical procedure is to determine the effect of discretization error in the measurement covariance matrix formulation. This discretization error is directly associated with the particular numerical method used to approximate (8). Regardless of the method used, the effects of discretization should be investigated. The error occurs in the head covariance and head log-transmissivity cross-covariance terms. To detect possible discretization error, these covariances are evaluated at block centerpoints (i.e., no interpolation is required) in the course and medium meshes. Comparisons are then made with values at coincident points in the fine mesh. Monte Carlo simulations have shown that, as long as \( \theta_h \) is small enough for the linearizing assumption to hold, the fine mesh produced error-free covariances. If \( i \) and \( j \) are indices of a pair of coincident block centerpoints, then the following ratios are evaluated for each pair \( (i,j) \):

\[
\begin{align*}
    r_1(i,j) &= \frac{\text{Cov}(\phi_i, \phi_j)}{\text{Cov}(\phi_i, \phi_j)} \quad \text{coarse mesh} \\
    r_2(i,j) &= \frac{\text{Cov}(\phi_i, \phi_j)}{\text{Cov}(\phi_i, \phi_j)} \quad \text{fine mesh} \\
    r_3(i,j) &= \frac{\text{Cov}(\phi_i, \phi_j)}{\text{Cov}(\phi_i, \phi_j)} \quad \text{medium mesh} \\
    r_4(i,j) &= \frac{\text{Cov}(\phi_i, \phi_j)}{\text{Cov}(\phi_i, \phi_j)} \quad \text{fine mesh}
\end{align*}
\]

The mean and standard deviation of the four ratios are given in Table 1. The results show that, for a given value of \( \theta_h \), the coarser the mesh the more \( Q_{ff} \) and \( Q_{ff} \) are underestimated. The reason is that, as the size of the block over which the log-transmissivity is averaged is increased, more variability is filtered out. In the inverse problem, where head variability is given and parameters are to be determined, the coarser the mesh the more \( |\theta_h| \) is overestimated. However, there is little discretization error in moving from the fine to the medium mesh. There does appear to be significant error, although in the \( \text{Cov}(\phi_i, \phi_j) \) term, in the move from the fine mesh to the coarse mesh.

To detect error in the measurement covariance matrix resulting from the interpolation scheme represented by (13), a similar procedure is used. In this test the covariance values are evaluated at points that coincide with block centerpoints in the fine mesh but which are not block centerpoints in the medium and coarse mesh. The same ratios \( r_1, r_2, r_3, \) and \( r_4 \) are evaluated as before, and the results are presented in Table 2. These ratios are smaller than those in Table 1, since these numbers actually show the combined effect of interpolation and discretization error. Note that the more uncorrelated the head values at the nodes surrounding the point measurement, the more the interpolation scheme causes the underestimation of the variance of the head measurement. The combined effect

<table>
<thead>
<tr>
<th>Variable</th>
<th>Sample Size</th>
<th>Sample Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>( r_1 )</td>
<td>10</td>
<td>0.72</td>
<td>0.04</td>
</tr>
<tr>
<td>( r_2 )</td>
<td>91</td>
<td>0.93</td>
<td>0.03</td>
</tr>
<tr>
<td>( r_3 )</td>
<td>52</td>
<td>0.59</td>
<td>0.13</td>
</tr>
<tr>
<td>( r_4 )</td>
<td>169</td>
<td>0.94</td>
<td>0.02</td>
</tr>
</tbody>
</table>

Values the result of discretization errors alone.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Sample Size</th>
<th>Sample Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>( r_1 )</td>
<td>91</td>
<td>0.41</td>
<td>0.12</td>
</tr>
<tr>
<td>( r_2 )</td>
<td>91</td>
<td>0.77</td>
<td>0.10</td>
</tr>
<tr>
<td>( r_3 )</td>
<td>169</td>
<td>0.41</td>
<td>0.12</td>
</tr>
<tr>
<td>( r_4 )</td>
<td>169</td>
<td>0.78</td>
<td>0.10</td>
</tr>
<tr>
<td>( r_5 )</td>
<td>136</td>
<td>0.82</td>
<td>0.04</td>
</tr>
<tr>
<td>( r_6 )</td>
<td>256</td>
<td>0.82</td>
<td>0.03</td>
</tr>
</tbody>
</table>

Values the result of combined discretization and interpolation error.

*No measurements within one block of boundary.
Each of the parameter estimates falls within approximately 2 (estimated) standard deviations of the value used in the data generation. The parameter estimation error variances are quite appreciable, but this is to be expected, considering the number of measurements used [see also Kitanidis, 1983]. While it is important to obtain the best estimates of structural parameters, values of the field parameter (log-transmissivity in this case) predicted via cokriging are not very sensitive to the values of the geostatistical parameter used.

Comparing the results of the two different mesh sizes shows that $\theta_3$ is smaller for the fine mesh (and $\theta_1$ is consequently larger). This is a characteristic observed in many other runs and is at least partly due to the discretization and interpolation error. It is obvious that, unless the grid is very fine, the estimates of the geostatistical parameters are affected somewhat by the discretization level. However, simulation results indicate that satisfactory block-averaged log-transmissivity estimates are obtained when cokriging is performed at the same discretization level at which the geostatistical parameters were estimated.

To check the validity of the assumed geostatistical structure, the estimated parameters and the measurement vectors $y_{\text{mod}}$ are used to generate two normalized measurement vectors: $y_1$ and $y_2$. The first normalized measurement vector $y_1$ is evaluated as

$$y_{1,i} = \frac{x_{\text{mod},i} - E[x_{\text{mod},i}]}{\sqrt{\text{Var}(x_{\text{mod},i})}}$$

(27)

The values of $\text{Var}(x_{\text{mod},i})$ are the diagonal elements of $Q_{\text{mod}}$ and are readily determined once $\theta_1$ and $\theta_3$ have been estimated. The second vector $y_2$ is an independent normal measurement vector (the elements of $y_1$ are not independent). This vector is evaluated as

$$y_2 = C^{-1}[x_{\text{mod}} - E[x_{\text{mod}}]]$$

(28)

where

$$CC^T = Q_{\text{mod}}$$

[see Kitanidis and Vomvoris, 1983].

The first measurement vector, $y_1$, may be used to detect outlying data. Normalized measurement errors whose values are much different than the rest may indicate the presence of outliers or the violation of modeling assumptions.

The second vector of residuals, $y_2$, is used to test the validity of the assumed geostatistical structure. Three appropriate tests for these residuals are $(1 - 2 < y_{2,i} < 2$ for about 95% of the

<table>
<thead>
<tr>
<th>Location, (\text{km})</th>
<th>$x$</th>
<th>$y$</th>
<th>Head $\phi$, m</th>
<th>Log-Transmissivity $Y$</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>100</td>
<td></td>
<td>737.5</td>
<td>7.095</td>
</tr>
<tr>
<td>75</td>
<td>250</td>
<td></td>
<td>658.6</td>
<td>6.550</td>
</tr>
<tr>
<td>100</td>
<td>175</td>
<td></td>
<td>726.5</td>
<td>9.214</td>
</tr>
<tr>
<td>125</td>
<td>75</td>
<td></td>
<td>786.9</td>
<td>9.277</td>
</tr>
<tr>
<td>175</td>
<td>250</td>
<td></td>
<td>702.6</td>
<td>6.934</td>
</tr>
<tr>
<td>225</td>
<td>50</td>
<td></td>
<td>804.2</td>
<td>6.382</td>
</tr>
<tr>
<td>250</td>
<td>125</td>
<td></td>
<td>779.9</td>
<td>3.484</td>
</tr>
<tr>
<td>250</td>
<td>200</td>
<td></td>
<td>750.2</td>
<td>7.283</td>
</tr>
<tr>
<td>175</td>
<td>150</td>
<td></td>
<td>751.4</td>
<td>7.137</td>
</tr>
<tr>
<td>200</td>
<td>100</td>
<td></td>
<td>784.4</td>
<td>7.104</td>
</tr>
<tr>
<td>200</td>
<td>200</td>
<td></td>
<td>735.6</td>
<td>8.895</td>
</tr>
<tr>
<td>100</td>
<td>125</td>
<td></td>
<td>745.5</td>
<td>7.755</td>
</tr>
</tbody>
</table>
values, \( 2 < y_{2,i}y_{2,j} < 2 \) for about 95% of the values \((i \neq j)\), and (3)

\[
\sum_{i=1}^{N} y_{2,i}^2 \sim \chi^2(N)
\]

In the third test, \( N \) is the number of elements in \( y_2 \). However, note that if the parameter estimates are maximum likelihood, \( P_{\theta\theta} \) matrix in (2) is zero, and only \( \theta_1 \) and \( \theta_2 \) are to be estimated, then it has been shown by Kitanidis [1983] that

\[
\sum_{i=1}^{N} y_{2,i}^2
\]

is fitted equal to \( N \).

In the two cases presented here there are no outliers in the \( y_1 \) vector, and all three tests hold for the \( y_2 \) vectors. Therefore, the geostatistical structure is assumed to be valid. Once the parameters have been estimated and geostatistical structure is considered to be valid, cokriging can be used to obtain estimates of log-transmissivity anywhere in the field. The results for block estimates are shown as log-transmissivity contour maps in Figures 8 and 9.

The methodology can next be tested to see if the original head field can be recovered. The estimated transmissivities are applied to the USGS groundwater modeling program to estimate the original head field. Figures 10 and 11 show the results. In each figure the original head field is shown with dashed lines. The results show that in both cases the original head field is adequately recovered. Even though the transmissivity maps look somewhat different, the important features (important relative to the head field) are adequately represented.
the three runs already described the head estimation error sample mean and variance are as follows:

Run 1 Mean = $-1.04\, \text{m}$ Var = $13.9\, \text{m}^2$ MSE = $15.0\, \text{m}^2$
Run 2 Mean = $-0.433\, \text{m}$ Var = $12.6\, \text{m}^2$ MSE = $12.8\, \text{m}^2$
Run 3 Mean = $12.1\, \text{m}$ Var = $76.1\, \text{m}^2$ MSE = $222.5\, \text{m}^2$

The addition of head data in the geostatistical parameter estimation procedure is seen here to reduce the MSE in the head estimates by a factor of about 17.

The developed approach is also very useful in obtaining best estimates and conditional simulations of the hydraulic head field [see Delhomme, 1979]. For purposes of illustration the best estimates of the hydraulic head field were obtained through cokriging. (The cokriging equations for the hydraulic head are not included in this paper.) The results of cokriging for the fine-mesh case are

Mean = $0.441\, \text{m}$ Var = $8.81\, \text{m}^2$ MSE = $9.00\, \text{m}^2$

indicating further reduction in the MSE of estimation of hydraulic heads. The results of cokriging reproduce the head measurements but produce a field that is smoother than the actual field. To obtain the range of variation that would exist, conditional simulations of the log-transmissivity and hydraulic head fields, given all available data, can be easily obtained.

A relevant question is whether the geostatistical approach does indeed solve the inverse problem in the sense of ensuring that when the estimates of transmissivities are introduced into the discrete form of the flow equations the computed heads are close to their observed values. The sum of squares of the errors (measured head minus reproduced) for the three runs is as follows

Run 1 $\text{SSE}_h = 98.3\, \text{m}^2$
Run 2 $\text{SSE}_h = 63.8\, \text{m}^2$
Run 3 $\text{SSE}_h = 3015\, \text{m}^2$

Thus the MSE of estimation of block-averaged log-transmissivities from the solution of the geostatistical inverse problem (fine mesh) is approximately half the corresponding MSE from kriging of log-transmissivity measurements only. This is a significant improvement, considering that the MSE obtained through simple kriging is already quite small. The medium-mesh results are only slightly better than the results of kriging. Note, however, that run 3 produced block estimates at a mesh twice as fine as run 1. (When fewer transmissivity measurements were utilized, the solution of the inverse problem using the medium mesh gave log-transmissivity estimates that were significantly better than the results of simple kriging but not as good as the results of the inverse problem using the fine mesh.)

The results of run 3 are also used to recover the original head field. This is shown in Figure 13. Here it is seen that the results of the first two runs provide a much better prediction of the head field. Since the original head field is known, the head estimation error $e_h$ can be evaluated at each block.
Thus the log-transmissivity obtained through the geostatistical solution to the inverse problem reproduced the head measurements much more closely (as measured by the sum of squares of the differences) than the log-transmissivities obtained from kriging. However, how closely the computed heads must reproduce the measurements is determined in the geostatistical approach when the geostatistical structure is identified and its parameters estimated. (In our model it depends on the values of $\theta, \sigma^2$, and $\text{Var}(h_0)$.) In practice the sum of squares of the fitting errors is also somewhat affected by linearization and discretization-interpolation errors.

5. Case Study

The described methodology is next applied to estimate the transmissivity field in the Jordan aquifer of Iowa [Horick and Steinthal, 1978]. The Jordan aquifer is part of the Cambrian-Ordovician aquifer system, which consists of the Cambrian Jordan sandstone as the principal water-yielding unit and the overlying Ordovician Oneota dolomite and Root Valley sandstone. The Jordan aquifer is present under most of the state of Iowa and is one of the most dependable of water supplies for large capacity wells in Iowa. It is confined everywhere, except in a small region near the northeastern corner of the state. The data used in this study was first compiled and screened by the Iowa USGS. The finite-difference model used to obtain average heads was also developed by the Iowa USGS.

Figure 14 shows the discretized model used to construct the measurement covariance matrix. The block dimension is 20 miles. This figure also shows the location of the measurements. The following screening criteria were used to reduce the original data to one containing 31 head and 56 log-transmissivity measurements:

1. Only pre-1940 head measurements were used to avoid transient conditions caused by accelerated rates of pumping after this period.
2. If multiple head measurements occur at a single location, only the earliest is used.
3. Head measurements near the boundaries were avoided.

The first run used only the 56 log-transmissivity measurements. Parameter estimation yielded

$$\theta_1 = 0.300 \quad \theta_2 = -0.00389/\text{mile}$$

$$\text{Cov}(\theta_1, \theta_2) = \begin{bmatrix} 0.815 (10)^{-2} \\ 0.113 (10)^{-3} & 0.553 (10)^{-5} \end{bmatrix}$$

Ninety-six percent of the normalized residuals and 94% of the cross products of the normalized residuals were in the interval $[-2, 2]$. The sum of squares of the normalized residuals of the fitted model was 55 (as expected, see Kitamidis [1983]). Further examination of the residuals indicated that the fitted linear model represents the spatial variability of the log-transmissivities of the Jordan aquifer quite adequately. These parameters were then used for the estimation through kriging of block-averaged log-transmissivities and then transmissivities throughout the aquifer, which are shown in Figure 15. (For communication purposes the transmissivities shown are the conditional mean calculated as $\exp(\hat{Y} + \text{Var}(\hat{Y})/2)$, where $\hat{Y}$ represents the cokriging variance of log-transmissivity).

The next run was application of the geostatistical solution...
to the inverse problem by using the entire set of 31 head and 56 log-transmissivity measurements. To represent uncertainty about the constant boundary heads used in the numerical model, boundary head variance was set equal to 200 ft$^2$, with integral scale $l = 100$ miles. It was assumed that there is no accretion, pumping, or leakage. The head variance $\sigma_b^2$ (equation (19)) was treated as a parameter $\theta_1$ to be estimated from the data. The results of parameter estimation are

$$\theta_1 = 0.234 \quad \theta_2 = -0.00755 \text{/mile} \quad \theta_3 = 1212 \text{ ft}^2$$

$$\text{Cov} (\theta_1, \theta_2, \theta_3) = \begin{bmatrix}
0.824(10)^{-2} \\
0.188(10)^{-3} \\
0.318
\end{bmatrix} \begin{bmatrix}
0.120(10)^{-4} \\
0.479(10)^{-1} \\
0.114(10)^{-6}
\end{bmatrix}$$

Note that parameter estimates $\theta_1$ and $\theta_2$ have not changed drastically from the values obtained when only log-transmissivities were used. Parameter $\theta_3$ is representative of how closely the numerical model reproduces the piezometric head measurements. Thus it is affected by head measurement error as well as inaccuracy in modeling assumptions (assumed input and boundary conditions) and discretization, interpolation, and roundoff error. Examination of the normalized residuals revealed that 94% of the normalized residuals and their cross products were located in the interval $[-2, 2]$, and the sum of the squares of the residuals is 86.4. However, two of the head measurements appeared to have fitted normalized residuals that were much larger than all the other residuals. These outliers have excessive influence on the estimation of $\theta_3$.

The log-transmissivity field was next obtained through cokriging. Figure 16 is a contour map of block-averaged transmissivity estimates. The large value of $\theta_3$ prompted the cokriging procedure to put relatively little weight on head measurements.

Sensitivity analysis revealed that the excessive value of $\theta_3$ is related mainly to the assumption that there is no inflow into the aquifer. In reality there appeared to be significant leakage into the Jordan aquifer, even though little is known about its spatial distribution. To account for leakage into the aquifer, (3) would require an additional term that would not allow elimination of the mean log-transmissivity $F$ from (8). Thus the following procedure was followed to improve the results.

The best estimates of leakage coefficients and heads on the other side of confining layers derived by USGS hydrologists was used to calculate the expected head field (corresponding to the average transmissivity obtained from the measurements). Since leakage was assumed concentrated over a small area on central western Iowa, and given the general uncertainty associated with leakage coefficients and head differences, the effect of leakage on head covariance matrices was neglected. More exact treatment of the effects of leakage is possible but is beyond the scope of this paper.

Thus the next run used the same measurements as the previous, except that the calculated expected head field accounted for leakage. The constant value of transmissivity used was the average of transmissivity measurements. The estimates of the structural parameters are

$$\theta_1 = 0.263 \quad \theta_2 = -0.00523 \text{/mile} \quad \theta_3 = 902 \text{ ft}^2$$

$$\text{Cov} (\theta_1, \theta_2, \theta_3) = \begin{bmatrix}
0.730(10)^{-2} \\
0.124(10)^{-3} \\
0.270
\end{bmatrix} \begin{bmatrix}
0.718(10)^{-5} \\
0.331(10)^{-1} \\
0.632(10)^{-5}
\end{bmatrix}$$
Ninety-five percent of the fitted normalized residuals, as well as their cross products, were between −2 and 2. The sum of the normalized residuals was 86.2.

The same two head measurements that highly influenced the results of the second run appeared to affect the results of this run to an inordinately high degree. These two measurements are in disagreement with the other head and transmissivity measurements and the assumed model. It was decided to repeat the calculation without these two measurements. Run 4, which is the same as run 3, except that it used two head measurements less, yielded the following parameter estimates

$$
\theta_1 = 0.256 \quad \theta_2 = -0.00538 \text{ mile} \quad \theta_3 = 427 \text{ ft}^2
$$

with estimation error covariance matrix

$$
\text{Cov} (\theta_1, \theta_2, \theta_3) = 
\begin{bmatrix}
0.699(10)^{-2} & 0.116(10)^{-3} & 0.699(10)^{-5} \\
0.116(10)^{-3} & 0.193 & 0.227(10)^{-1} \\
0.699(10)^{-5} & 0.227(10)^{-1} & 0.166(10)^5
\end{bmatrix}
$$

The residuals passed the tests, and the presence of outliers was not obvious. Note that $\theta_3$ corresponds to a standard error of about 20 ft compared with a maximum head drop throughout the aquifer of about 600 ft. However, the large value of $\theta_3$ reduces the usefulness of head measurements in the identification of small-scale variability in the transmissivity field (Figure 17). The 95% confidence limits are given in Figures 18 and 19. It is possible that the estimation accuracy may improve further through more accurate accounting of leakage.

Note that the small value of $\theta_2$ indicates that linearized analysis is appropriate for the solution of the inverse problem in the case study. Sensitivity analysis showed that the assumed boundary head variance and integral scale have only minor effects on the results of parameter estimation for a large range of values near the values used. Therefore, the assumed value used should be chosen carefully to reflect the modeler's uncertainty in the boundary conditions, but a slight error will have only a small effect on the results. The effect of discretization and interpolation error is negligible.

As long as the number of measurements is not too large, most of the computer storage and time requirements are associated with the generation of the covariance matrix of the data using a numerical model (equations (10) through (19)). The computer storage requirements at this step increase with the square of the number of nodes of the numerical model (126 in the case of the Jordan aquifer). However, our simulation results indicate that the use of very fine grids to calculate the data covariance matrix may not be warranted. The computer storage requirements for the rest of the procedures is determined by the number of measurements and the number of parameters. Thus one complete run with the Jordan aquifer model, including covariance matrix generation, parameter estimation, testing of residuals, and cokriging to obtain the log-transmissivity estimates at 126 blocks, used only 2400 s of CPU time on one of the University of Iowa's PRIME 750 minicomputers. (For comparison, computing on a main-frame computer can be about 10 times faster.) This is considered a satisfactory time for the solution of an inverse problem of such magnitude and scope. Further improvements in the computer storage and time requirements could be achieved through use of implicit schemes (which would avoid the inversion of matrix $A$ of equation (10)) and sparse-matrix techniques in the calculation of the data covariance matrix.
Fig. 17. Predicted transmissivity (ft²/d) contour map for Jordan aquifer run 4. Leakage is accounted for only in the calculation of the average head field. Estimated parameters are $\theta_1 = 0.256$, $\theta_2 = -0.00538$ and $\theta_3 = 427$ ft².

Fig. 18. Lower limit of the 95% confidence interval of transmissivities (ft²/d).
6. SUMMARY AND CONCLUSIONS

This paper has presented a methodology for the estimation of transmissivity from head and transmissivity measurements in the two-dimensional steady flow case following the geostatistical approach to inverse modeling [Kitanidis and Vomoidis, 1983]. The differential equation relating hydraulic head to log-transmissivity perturbations for a two-dimensional confined aquifer with given head along the boundary can be linearized, provided the large-scale variability of the log-transmissivity field is not very pronounced. The geostatistical structure of the log-transmissivity field may be conveniently described through a linear generalized covariance function $\theta_1\delta + \theta_2d$ where $\delta$ is Kronecker's delta, $d$ is the distance between points, $\theta_1$ is the variance of small-scale or nugget variability, and $\theta_2$ is the variance of large-scale or interblock variability. Then the covariance matrix of the head and log-transmissivity measurements can be expressed as a linear function of the geostatistical parameters with coefficient matrices that are numerically calculated only once.

Having constructed the covariance matrix of the measurement vector, maximum-likelihood estimates of the geostatistical, as well as additional, parameters can be readily obtained on the basis of the actual data. The validity of the fitted model is tested, and if the model is accepted as adequate, the log-transmissivity at any point or over any block in the field can be predicted through cokriging.

The procedure was applied to a test case in which the actual values of log-transmissivity and head were generated with known values of the geostatistical parameters. The test aquifer is square with 300-km sides and constant head boundaries. To investigate the effect of discretization (or block-averaging) and interpolation (head measurements not coinciding with grid nodes) error, three levels of discretization were investigated (fine: $\Delta x = 25$ km; medium: $\Delta x = 50$ km; coarse: $\Delta x = 100$ km). The results of simulations showed that the coarse grid does not give useful results because it effectively averages out most of the large-scale log-transmissivity variability. The medium level of discretization gave reasonable results, but not as good as the fine level of discretization. An important result of the simulation study is that adequate results can be obtained with block dimensions as large as 1/12 of the aquifer dimension. This is a fairly coarse level of discretization for most numerical groundwater flow models.

Use of head measurements was shown to improve the log-transmissivity estimates, although the degree of improvement may vary widely from case to case. When these estimates were used in the numerical flow model, the head measurements were closely reproduced. Best estimates of the head field were obtained through cokriging on both head and log-transmissivity data.

The developed methodology was applied for the estimation of transmissivities of the Jordan aquifer in Iowa. This case study illustrated that the geostatistical approach can be a useful tool in the iterative process of model development. Model validation serves to diagnose possible conflicts between modeling assumptions and available data and to suggest ways in which these conflicts can be resolved. In the case of the Jordan aquifer, accounting for leakage into the aquifer and removal of two outliers appeared to improve the consistency among modeling assumptions and measurements. The results also illustrate that the usefulness of piezometric head measurements in improving log-transmissivity estimates in regional aquifer studies depends critically on the accuracy of these measurements as well as the accuracy of the assumed model.
piwecwise continuous, the integral is defined in Lebesques' sense. By virtue of Green's identities, (A1) is equivalent to (3) in the text but, unlike (3), does not involve transmissivity derivatives. Let
\[ Y = \ln T = F + f_1 + f_2 \]  
(A2)
where \( F \) is a constant and \( f_1 \) and \( f_2 \) are two zero-mean independent random functions; \( f_1 \) is a "nugget-effect" or "white-noise" process:
\[ E[f_1(x_1)f_1(x_2)] = \begin{cases} \theta_1 & \text{if } x_1 = x_2 \\ 0 & \text{otherwise} \end{cases} \]  
(A3)
where \( \theta_1 \geq 0 \) and \( f_2 \) is a continuous intrinsic function. For example, its generalized covariance function may be given by the linear model and denoted by
\[ E[f_2(x_1)f_2(x_2)] = \theta_2|x_1 - x_2| \]  
(A4)
where \( \theta_2 \leq 0 \). Let
\[ w_1(x) = \exp(f_1(x)) \]  
(A5)
\[ w_2(x) = \exp(f_2(x)) \]  
(A6)
and
\[ w_1(x) = \tilde{w}_1 + w_1'(x) \]  
(A7)
\[ w_2(x) = \tilde{w}_2 + w_2'(x) \]  
(A8)
where
\[ \tilde{w}_1 = E[w_1(x)] \]  
(A9)
\[ \tilde{w}_2 = E[w_2(x)] \]  
(A10)
TABLE 4. Values of the Coefficients Used in the Numerical Integration Scheme to Determine the Average Distance Between Two Square Areas

<table>
<thead>
<tr>
<th></th>
<th>( \lambda_i )</th>
<th>( \alpha_i )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.2710</td>
<td>0.0571</td>
</tr>
<tr>
<td>2</td>
<td>0.4069</td>
<td>0.2768</td>
</tr>
<tr>
<td>3</td>
<td>0.2597</td>
<td>0.5836</td>
</tr>
<tr>
<td>4</td>
<td>0.06235</td>
<td>0.86025</td>
</tr>
</tbody>
</table>

Since \( f_1 \) is a white-noise process, its exponential transformation is also a white-noise process. That is, \( w_1 \cdot \phi(x) \) is zero-mean white noise. Similarly, since \( f_2 \) is a continuous stochastic process, its exponential transformation is a continuous process.

Substituting in (A1)

\[
\begin{align*}
\hat{w}_1 \hat{w}_2 \int_0 \frac{\partial \phi}{\partial n} dS + \int_0 w_2(S) \frac{\partial \phi}{\partial n} dS + \hat{w}_2 \int_0 w_1(S) \frac{\partial \phi}{\partial n} dS \\
+ \int_0 w_1(S) w_2(S) \frac{\partial \phi}{\partial n} dS = 0 \quad (A11)
\end{align*}
\]

The relationship can be simplified [Titchmarsh, 1939] by taking into account that, for every finite length \( S \) and any continuous deterministic or stochastic function \( g \), the (Lebesques) integral

\[
\int_0 w_1(S) g(S) dS = 0 \quad (A12)
\]

Since our objective is to calculate the head with Dirichlet boundary conditions, we may restrict our attention to solution for which \( \frac{\partial \phi}{\partial n} \) is a piecewise continuous function. Hence integrals that involve \( w_1(S) \) are dropped, and after cancelling the constant \( \hat{w}_1 \), the equation becomes

\[
\int_0 w_2(S) \frac{\partial \phi}{\partial n} dS = 0 \quad (A13)
\]

or

\[
\int_0 \exp(f_2(S)) \frac{\partial \psi}{\partial n} dS = 0 \quad (A14)
\]

We proceed now with the small-perturbation approach. Let

\[
H = \psi + h \quad (A14)
\]

where \( H = E[\psi] \). Dropping the subscript 2 and substituting (A14) in (A13) gives

\[
\int_0 \exp(f) \left( \frac{\partial H}{\partial n} + \frac{\partial h}{\partial n} \right) dS = 0 \quad (A15)
\]

We assume that \( f \) and \( h \) are small perturbations, and we order in powers of fluctuations. Noting that \( \exp(f) = 1 + f + O(f^2) \). At zero order,

\[
\frac{\partial H}{\partial n} dS = 0 \quad (A16)
\]

which is equivalent to (7) in the text. At first order,

\[
\frac{\partial H}{\partial n} dS = 0 \quad (A17)
\]

One may readily verify, by using one of Green's identities, that if \( \frac{\partial}{\partial x} \) and \( \frac{\partial^2 h}{\partial x^2} \) exist, then (A17) is equivalent to (8) in the text. The advantage of the integral analysis is that it avoids using these derivatives, which may not exist in a mean square sense.

The typical first-order term is of the order of \( \int f_{1} |\partial H/\partial n| \), while the typical second-order term is of the order of \( \int f_{2} |\partial H/\partial n| \). Thus the small-perturbation or linearized approach holds for small \( |f| \). For the model of equation (A4) this implies that \( |\partial h/\partial L| \), where \( L \) is the maximum aquifer length, is small.

APPENDIX B

The average distance from point \( k \) to the square finite difference block \( l \) is designated as \( d_{kl} \) and is given by the relation

\[
d_{kl} = \frac{1}{\Delta x^2} \int_{x_k}^{x_l} \int_{y_k}^{y_l} d(x, y) dA
\]

\[
= \frac{1}{\Delta x^2} \left[ \frac{1}{3} \sum_{i=1}^{3} (x_i d_{3} - x_{i+1} d_{2}) + \frac{1}{6} \sum_{i=1}^{3} (y_i d_{3} - y_{i+1} d_{2}) \right]
\]

where \( \Delta x \) is the block dimension and \( x_k, x_{k+1}, x_{k+2}, y_k, y_{k+1}, \) and \( y_{k+2} \) are distances defined in Figure 20. Upon integration, the following result is obtained:

\[
d_{kl} = \frac{1}{3} \sum_{i=1}^{3} (x_i d_{3} - x_{i+1} d_{2}) + \frac{1}{6} \sum_{i=1}^{3} (y_i d_{3} - y_{i+1} d_{2})
\]

\[
\int_{x_k}^{x_l} \int_{y_k}^{y_l} \sqrt{x^2 + y^2} \, dx \, dy
\]

The average distance from the square finite-difference block \( k \) to the square block \( l \) is designated as \( d_{kl} \). The value of \( d_{kl} \) is estimated by using a four-point Gauss numerical integration as presented in Journal and Huijbers [1978]. Using the notation of this paper, the relationship for \( d_{kl} \) is expressed as

\[
\hat{d}_{kl} = \sum_{i=1}^{4} \sum_{j=1}^{4} \lambda_{ij} \left[ (x_0 + e_1 x \Delta x)^2 + (Y_0 + e_2 y \Delta x)^2 \right]^{1/2}
\]

where \( x_0 \) and \( Y_0 \) are defined in Figure 21 and the coefficients \( \lambda \) and \( e \) are given in Table 4.

Acknowledgments. The material is based on work supported by the National Science Foundation under grant CME-8106577. The Jordan aquifer study was supported in part by funds provided by the U.S. Department of the Interior through the Iowa Water Resources Research Institute (Project CT 31702). The assistance of Mike Burkart and Bob Lane of the Iowa District of the USGS is gratefully acknowledged. The support of the Graduate College of the University of Iowa is also appreciated.

REFERENCES


R. J. Hoeckema and P. K. Kitanidis, Institute of Hydraulic Research, The University of Iowa, Iowa City, IA 52242.

(Received August 25, 1983; revised February 21, 1984; accepted March 8, 1984.)