7 6 3

WATER RESOURCES RESEARCH, VOL. 27, NO. 4, PAGES 533–546, APRIL 1991

The Geostatistical Characteristics of the Borden Aquifer

ALLAN D. WOODBURY

Department of Geological Engineering, University of Manitoba, Winnipeg, Canada

E. A. SUDICKY

Waterloo Centre for Groundwater Research, University of Waterloo, Ontario, Canada

A complete reexamination of Sudicky's (1986) field experiment for the geostatistical characterization of hydraulic conductivity at the Borden aquifer in Ontario, Canada is performed. The sampled data reveal that a number of outliers (low ln (K) values) are present in the data base. These low values cause difficulties in both variogram estimation and determining population statistics. The analysis shows that assuming either a normal distribution or exponential distribution for log conductivity is appropriate. The classical, Cressie/Hawkins and squared median of the absolute deviations (SMAD) estimators are used to compute experimental variograms. None of these estimators provides completely satisfactory variograms for the Borden data with the exception of the classical estimator with outliers removed from the data set. Theoretical exponential variogram parameters are determined from nonlinear (NL) estimation. Differences are obtained between NL fits and those of Sudicky (1986). For the classical-screened estimated variogram, NL fits produce an ln (K) variance of 0.24, nugget of 0.07, and integral scales of 5.1 m horizontal and 0.21 m vertical along A-A'. For B-B' these values are 0.37, 0.11, 8.3 and 0.34. The fitted parameter set for B-B' data (horizontal and vertical) is statistically different than the parameter set determined for A-A'. We also evaluate a probabilistic form of Dagan's (1982, 1987) equations relating geostatistical parameters to a tracer cloud's spreading moments. The equations are evaluated using the parameter estimates and covariances determined from line A-A' as input, with a velocity equal to 9.0 cm/day. The results are compared with actual values determined from the field test, but evaluated by both Freyberg (1986) and Rajaram and Gelhar (1988). The geostatistical parameters developed from this study produce an excellent fit to both sets of calculated plume moments when combined with Dagan's stochastic theory for predicting the spread of a tracer cloud.

INTRODUCTION

Sudicky [1986] described the results of a sampling program in which a large number of hydraulic conductivity measurements were taken along two transects at the site of an elaborate tracer test performed in the Borden aquifer in Ontario, Canada. These measurements, combined with a detailed evaluation of the dispersion characteristics of the injected tracer cloud [*Freyberg*, 1986], provided a unique opportunity to examine the validity of modern stochastic theories of contaminant transport that have emerged over the past decade. Based on the field data and subsequent geostatistical inferences, *Sudicky* [1986] computed mean values, variances and integral scales for the underlying log conductivity distribution of the Borden aquifer. Then, by using these quantities as input to stochastic transport theories by *Dagan* [1982, 1987] and *Gelhar and Axness* [1983], the predicted field-scale flow and dispersion parameters were shown to be consistent with the observed evolution of the tracer plume as interpreted by *Freyberg* [1986]. The geostatistical interpretation performed by *Sudicky* [1986], however, did not elucidate the uncertainties associated with estimating critical parameters such as the log conductivity integral scales. Subsequent comments on *Sudicky*'s [1986] work can be found in the works by *Kemblowski* [1988], *White* [1988], *Molz and Güven* [1988] and *Sudicky* [1988]. Other discussions centering on the applicability of various competing theories to explain the observed tracer behavior can be found in the works by *Naff et al.* [1988, 1989], *Barry* [1990], *Dagan* [1989*a*, 1990], *Neuman and Zhang* [1990], and *Zhang and Neuman* [1990].

In view of the fact that estimating the parameters describing the spatial variations of hydraulic conductivity is a challenging problem, even when many data are available, our purpose is to systematically examine the Borden hydraulic conductivity data with particular emphasis on how various assumptions and modes of interpretation might affect the values of inferred parameters. In particular, we address and quantify the uncertainty in the values of the geostatistical parameters and demonstrate how this uncertainty manifests itself when predicting the spread of the Borden experimental plume in a probabilistic framework. This latter aspect is extremely important when using stochastic transport theory. Indeed, *Dagan* [1988] emphasized this need, and to our knowledge, our work represents the first time that the effects of parameter uncertainty in stochastic dispersion models has been quantified on the basis of field data. We also show the effects of using different variogram estimators on the sampled data, and how truncation of the experimental variograms influences inferred geostatistical parameters. Neither of these topics were addressed in the second author's original work [*Sudicky*, 1986]. Other issues discussed in this paper relate to the generality of the lognormal assumption for hydraulic conductivity which is commonly invoked in stochastic groundwater flow models and the fact that it is unlikely that one can unequivocally distinguish between various competing variogram models (i.e., spherical, exponential, Bessel, etc.) even when very large amounts of data are at hand. This, in our opinion, will have some influence on

Paper number 90WR02545.
0043-1397/91/90WR-02545$05.00

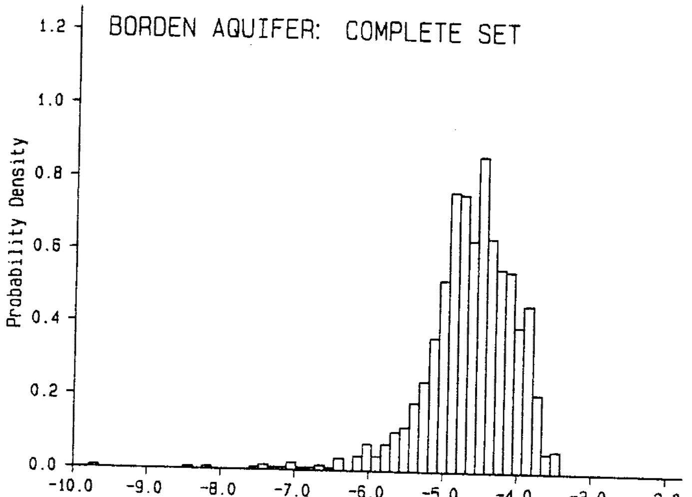

Fig. 1. Histogram of entire data base, Borden aquifer [*Sudicky*, 1986].

how stochastic flow and transport theories will evolve in the future. In summary, the following points are addressed in the analysis presented here: (1) the lognormality assumption of the sampled hydraulic conductivity data, (2) maximum lag values used in autocorrelation or variogram calculations and the effect of lag distance truncation on variogram fitting, (3) the method of computing variogram values, and linear versus nonlinear fitting procedures to infer geostatistical parameters, and (4) the influence of uncertainty in the values of the geostatistical parameters on plume spreading.

A logical approach to geostatistics has been suggested by *Armstrong* [1984] and these fundamental steps are taken in subsequent sections in this paper. These steps are (1) data, (2) experimental variogram, (3) variogram model, and (4) application:

1. The type of distribution and the mean and higher moments of the data are determined, and outliers identified.
2. The experimental variogram is computed. An appropriate estimator must be chosen, for example, classical [*Matheron*, 1963] or *Cressie and Hawkins* [1980].
3. Theoretical variogram parameters are sought from the experimental variogram. Important decisions are drift identification, and type of theoretical model, method of parametric estimation and length of experimental variogram record used.
4. The final step is in the application of the material established from steps 1–3. These applications could be, for example, kriging or stochastic models.

These points will be discussed in some detail and our final application of the geostatistical parameters for the data is the determination of, and a comparison with, the spatial variance tensor calculations from the natural gradient tracer experiment at the Borden aquifer.

The Lognormality of Hydraulic Conductivity

As is common in stochastic groundwater analysis, hydraulic conductivity values are assumed to follow a lognormal distribution and a logarithmic transformation is made such that the data in this transformed space are normally distributed [*Freeze*, 1975]. For example, with $Y = \ln(K)$, K being the hydraulic conductivity,

$$P(Y) = \frac{1}{(2\pi)^{1/2}\sigma} \exp\left[\frac{-(Y - \langle Y\rangle)^2}{2\sigma^2}\right] \quad (1)$$

defines a Gaussian probability density (pdf) describing the random variable Y where σ is the standard deviation of Y and $\langle Y\rangle$ is its mean. A chi-square goodness-of-fit test was performed by *Sudicky* [1986] based on a subset of data having a prescribed minimum separation between neighboring measurements such that independence could be assumed. These spacings were 0.25 m vertically and 2.0 m horizontally. *Sudicky* [1986] found that $P(Y)$ given by (1) was not unreasonable; however, a number of outliers are present in the full data set. Because of these outliers, an exponential distribution for Y is as likely a candidate for a pdf as a normal one. Figure 1 shows a frequency density histogram for all of the regularly spaced ln (K) data used to produce Figures 6 and 7 of *Sudicky* [1986], which show hydraulic conductivity profiles and cross-section contours of ln (K). It can be seen that a number of low Y values are present (i.e., $Y < -6.5$, $K < 1.5 \times 10^{-3}$ cm/s). As will be shown later, these low values cause difficulties in estimating both variogram and population statistics.

In order to examine sample statistics from a data set that is suspected of being correlated, an alternative to the selection procedure used by Sudicky based on a minimum spacing is to draw a sample from the population at random [e.g., *Smith*, 1981]. Figure 2 shows a histogram of the probability densities for a random sample comprising 100 points along with the continuous forms of the normal and exponential distributions. The exponential distribution is defined as

$$P(Y) = \frac{1}{2\sigma} \exp\left[\frac{-|Y - \langle Y\rangle|}{\sigma}\right] \quad (2)$$

Here, σ is the average deviation given by

$$\sigma = \frac{1}{N} \sum_{i=1}^{N} |Y - \langle Y\rangle| \quad (3)$$

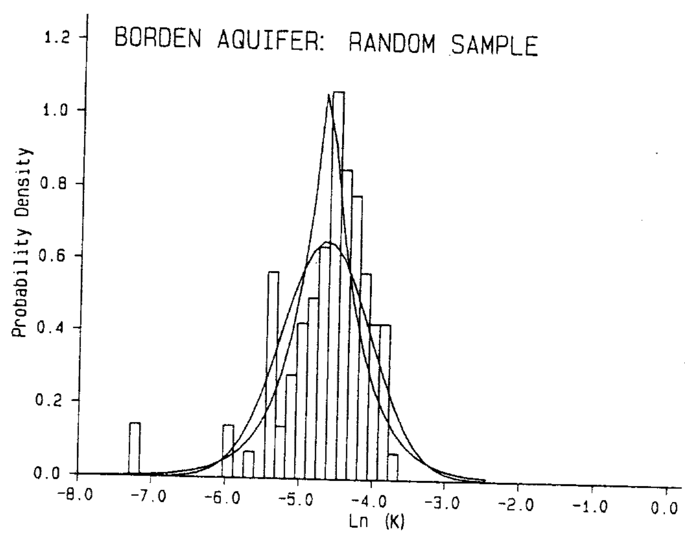

Fig. 2. Frequency density histogram and plotted theoretical curves for normal and exponential pdfs for a random sample of the entire data set.

Each of the above three formulas will be applied to calculate an experimental variogram describing the Borden aquifer data.

The estimator of the autocorrelation function used by *Sudicky* [1986] in his equation (6) is a biased estimator [*Jenkins and Watts*, 1968, pp. 174–175; *Smith*, 1981] because of the normalization with respect to $1/n$ instead of $1/n(h)$. This estimator tends to reduce the mean square error and smooth variations at large lags although it is asymptotically unbiased as $n \rightarrow \infty$ [*Jenkins and Watts*, 1968]. Because $n(h)$ is relatively large for the lag distances used by *Sudicky* [1986], he found both the biased and unbiased estimator yielded similar results. The implication of using this autocorrelation estimator is that the integral scales estimated from such calculations may appear artificially shorter than unbiased estimators.

DRIFT ESTIMATION

Prior to presenting the results of the variogram calculations, let us examine the possibility of a drift or trend in the ln (K) data. An apparent anomaly in the large-time plume behavior of the Borden tracer experiment was noted by *Freyberg* [1986], and *Molz and Güven* [1988] suspected this behavior to be the result of some form of nonstationarity being present in the hydraulic conductivity field. *Sudicky* [1988], in response to the comments raised by *Molz and Güven* [1988], suggested this behavior was the result of a nonstationarity in the flow boundary conditions rather than the hydraulic conductivity. We examine the possibility of drift in the following sections by statistical methods.

It is common in practical geostatistical applications to only use (or recognize) that part of the variogram in which $|h| < L/2$, with L being the length of the transect sampled, such that, say, $n(h) > 30$ pairs [*Journel and Huijbregts*, 1978, p. 194], and where a trend or "drift" is not suspected (P. K. Kitanidis, 1989, personal communication). Figures 2, 6, 7 and 9 of *Sudicky* [1986] show the location, dimensions, and partial summary statistics of the original survey. The transects are 19 m long for A-A′ and 13 m long for B-B′. Vertical depths for both transects are 1.75 m. We point out that in Figure 10 of *Sudicky* [1986], only results for data pairs having lags in the range of 0–10 m are presented for the horizontal direction and 0–0.5 m in the vertical direction partially in accordance with the above mentioned criteria.

Figure 9 of *Sudicky* [1986] shows the sample means for each core plotted against horizontal distance for both the A-A′ and B-B′ directions. *Sudicky* [1986] originally argued that there was little evidence for a trend in log conductivity with depth and this argument was supported by several fully penetrating exploratory cores obtained nearby [*Mackay et al.*, 1986]. It is, however, possible to perform statistical tests to confirm or repudiate if a trend indeed exists in the mean vertical ln (K) profiles. The philosophy behind the tests is this; first, we assume that each core consisting of vertical data is viewed as a realization of a stochastic process. Second, we adopt the null hypothesis or the hypothesis that the observed data are drawn from a specified population. Third, we apply a χ^2 test to confirm the null hypothesis. Such a test can be defined as [e.g., *Tarantola*, 1987, p. 213]

$$\chi_Y^2 = (\mathbf{Y} - \langle \mathbf{Y} \rangle)^T [\mathbf{C}]^{-1} (\mathbf{Y} - \langle \mathbf{Y} \rangle) \quad (7)$$

If χ_Y^2 exceeds a value for a specified degree of freedom and confidence level, then the null hypothesis is rejected and, therefore, the observed values to not belong to the correlated population thought to be applicable. However, it should be noted that the test is sensitive to outliers in the data as there is an implicit assumption of normality in the formalism. It is assumed here that the data $\mathbf{Y}$ are measured (or determined) without error. For the vertical data, the parameters in (7) are a constant mean value vector of $\langle \mathbf{Y} \rangle = -4.63$ and a covariance matrix $\mathbf{C}$ generated from an exponential covariance function,

$$C_{ij} = \sigma_Y^2 \exp\left(\frac{-h_{ij}}{\lambda_z}\right) + \sigma_0^2 \delta_{ij} \quad (8)$$

where σ_Y^2 is the variance, σ_0^2 is the "nugget", and $\delta_{ij} = 1$ if $i = j$ and zero otherwise. Also, h_{ij} is the vertical distance between points i, j and λ_z is the vertical integral scale. In the tests we use a ln (K) variance of 0.28 m^2, a nugget of 0.10 m^2 and an integral scale of 0.12 m. These are Sudicky's original values and are assumed correct initially. The test (7) is applied on each set of vertical cores on each transect. There are 20 cores on line A-A′ and 13 cores on B-B′. Each core contains 36 values of ln (K), separated at 0.05 m.

The results indicate that all vertical cores passed the above chi-square test using *Sudicky*'s [1986] original parameters. Hence the data can be statistically represented as a multivariate Gaussian pdf with an exponential correlation structure. It has been assumed here that Sudicky's original population parameters are correct in the analysis.

The analysis is repeated for new population parameter values determined in subsequent sections (see Table 4). In this test we first screen outliers from each line (values of log K lower than -6.5). All cores pass the test. We then repeat the test without screening these low values, and eight cores on section A-A′ and two cores on B-B′ fail the test. Hence, it is concluded that the outliers do influence the test results and if we are not careful in accounting for their presence the exponential model might be rejected on some cores. However, by effectively rejecting their presence, tests indicate that the assumed exponential form of the covariance function can be safely assumed. Note that this covariance does not admit a trend in the data. This subject is discussed at greater length in a subsequent section.

EXPERIMENTAL VARIOGRAM CALCULATIONS

Figure 3 shows experimental variograms for section A-A′ in the vertical direction. Line A is the classical estimate, line B is the Cressie/Hawkins estimate and line C is classical estimate with outliers removed. Here, ln (K) data less than -6.5 are removed from the data set, based on the Chauvenet criterion [*Neville and Kennedy*, 1964]. The remaining two solid lines are 95% confidence intervals about line C. These intervals should be viewed as qualitative as they are based on power-variogram models [*Journel and Huijbregts*, 1978, p. 193]. Line D is the SMAD estimate. As shown in Figure 3, lines A, B and D are affected by the outliers. Note the difference in variogram values for all lags for lines A and C. Line B, the Cressie/Hawkins estimate, is somewhat lower than line A and is apparently less affected by outliers at shorter lags. Line D is the SMAD estimate. Note that the

The plotted probability density functions are based on the sample statistics provided in Table 1*a*. The sampled data are combined into 20 groups. Note in Figure 2 that the sample data appear to be more peaked than the normal pdf. Because the presence of outliers (Table 1*b*, Figure 1) suggests that an exponential pdf may also produce a satisfactory fit to the data, a chi-square test was carried out to determine the goodness-of-fit of each of the two distributions.

Table 2 summarizes the test results for the frequency histogram shown in Figure 2 and using the sample statistics given in Table 1*b*. The computed probabilities in Table 2 represent the probability that the null hypothesis, or the hypothesis that there is no statistical difference between sample and theoretical quantities, is satisfied. The lowest probability of assuming the null hypothesis is correct is 8.26% for the exponential distribution. Therefore, both distributions satisfy the data.

The inability to unambiguously distinguish between the various competing pdfs to describe hydraulic conductivity is disturbing considering the sample size and the intense field efforts required to collect such data. Our finding is not without precedent. *Smith* [1981], for example, concluded that the hydraulic conductivity of the Quadra sand at a site near Vancouver, British Columbia, could be described equally well by either a normal or a lognormal pdf.

The analysis demonstrates that assuming either a normal distribution or exponential distribution for log conductivity at the Borden site is appropriate. Thus, it seems that the standard practice in groundwater hydrology of assuming that spatial variations in hydraulic conductivity follow a lognormal distribution is subject to debate and that stochastic theories constrained by such an assumption may be restricted in their general applicability.

The presence of "tails" in a presumed normal (i.e., $Y = \ln(K)$) pdf can have significant consequences on the inferred contaminant transport characteristics of the subsurface. In the case of the Borden aquifer where the so-called outliers comprised several low hydraulic values, these low values could restrict vertical motion, thereby diminishing vertical mixing processes. On the other hand, the neglect of a few high hydraulic conductivity lenses mistakenly viewed as outliers in a standard parametric statistical analysis can introduce a bias that obscures the possibility of rapid interconnected transport pathways in an aquifer. These sentiments are reflected in the non-Gaussian approach to uncertainty in spatial data suggested by *Journel* [1990]. Another consequence of several extreme values of hydraulic conductivity in a collection of measurements is their impact on variogram estimation.

TABLE 1*b*. Borden: Randomly Sampled Data Set Sample Statistics for Screened Data Set

Parameter	Estimate
Mean	−4.62
Median	−4.60
Average deviation	0.412
Standard deviation	0.511
Variance	0.261
Squared average deviation	0.170

Sample size of 100 values randomly sampled out of a population of 1188. Log conductivity values less than −6.5 removed from data set.

EXPERIMENTAL VARIOGRAM ESTIMATION

Rather than present experimental autocorrelations of $\ln(K)$ as a function of separation distance, experimental semivariograms are adopted here to avoid normalizing second-order moments of the data by an assumed variance. Because there is much debate about which variogram estimator is most appropriate, particularly if the data are corrupted by noise or outliers [*Armstrong*, 1984], several estimators are used in the analysis presented below. The "classical" experimental semivariogram estimator, $\gamma^*(h)$, hereinafter called the variogram, is [*Matheron*, 1963]

$$\gamma^*(h) = \frac{1}{2n(h)} \sum_{i=1}^{n(h)} [Y(\mathbf{x}_i) - Y(\mathbf{x}_i + h)]^2 \qquad (4)$$

where $n(h)$ is the number of data pairs separated by lag distance h, and $Y(\mathbf{x}_i)$ are measured values of log conductivity at coordinates $\mathbf{x}_i$. The above equation has been shown by *Omre* [1984] to be an optimal estimator of the variogram provided that the pairs $Y(\mathbf{x}_i)$ and $Y(\mathbf{x}_i + h)$ are bivariate normal and independent. The Cressie-Hawkins estimator is [*Cressie and Hawkins*, 1980]

$$\gamma^*(h) = \frac{1}{2}\left[\frac{1}{n(h)} \sum_{i=1}^{n(h)} |Y(x_i) - Y(x_i + h)|^{1/2}\right]^4 \Big/ [0.457 + 0.494/n(h)] \qquad (5)$$

This estimator is considered "robust" in that it reduces the effect of outliers in the data. Another robust estimator is the squared median of the absolute deviations (SMAD) estimator given by [*Dowd*, 1984]

$$\gamma^*(h) = 2.198 \times [\text{median}|Y(\mathbf{x}_i) - Y(\mathbf{x}_i + h)|]^2 \qquad (6)$$

TABLE 1*a*. Borden: Randomly Sampled Data Set Sample Statistics

Parameter	Estimate
Mean	−4.68
Median	−4.63
Average deviation	0.464
Standard deviation	0.597
Variance	0.356
Squared average deviation	0.215

Sample size of 100 values randomly sampled out of a population of 1188.

TABLE 2. Borden: Randomly Sampled Data Set χ^2 Tests on Normal/Exponential Hypotheses

Statistic	Value
χ^2 (normal)	13.81
Probability (null)	16.5%
χ^2 (exponential)	20.53
Probability (null)	8.26%

Chi square tests carried out on observed and expected frequencies. Low probability values here indicate that the null hypothesis is incorrect, i.e., that the observed class frequencies are not represented by the chosen function.

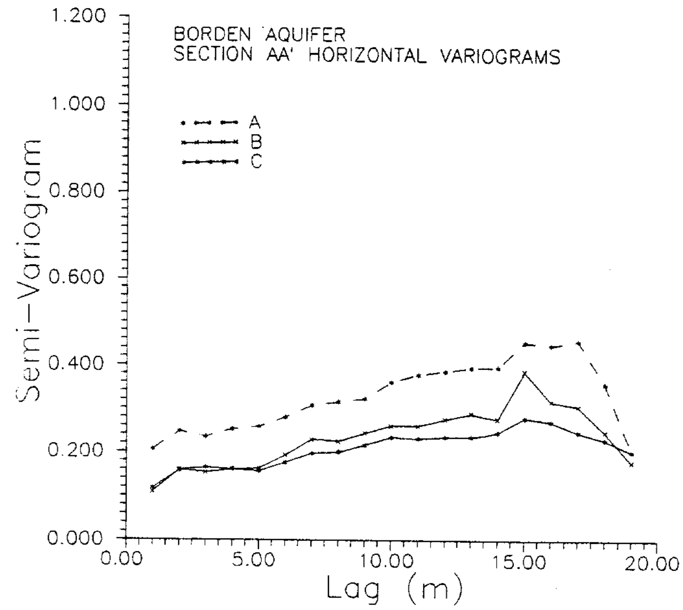

Fig. 5. A-A' horizontal variogram. A is the "classical" variogram; B, Cressie/Hawkins; C, classical variogram for data set with outliers removed.

various discrimination and cross-validation tests on competing models (for example, the Akaike information criterion (AIC) [see *Russo and Jury*, 1987]). This criterion is:

$$AIC = 2(\mathcal{L} + k) \tag{9}$$

where k is the number of independent parameters in the solution and $\mathcal{L}$ is the negative log likelihood function [e.g., *Hoeksema and Kitanidis*, 1985]:

$$\mathcal{L} = \frac{1}{2}[N \ln(2\pi) + N \ln(\sigma^2) + \chi^2] \tag{10}$$

While we do not present any AIC or cross-validation tests for competing models, we do carry out standard tests such as examining the mean square reduced errors and mean errors for the exponential model. These tests are used in a subsequent section on inversion results.

Given an assumed structural model and the experimental variogram, the model parameters need to be determined by some inversion procedure. In performing this inversion process, a theoretical variogram model $\gamma(h_j)$ will be adopted that is equivalent to *Sudicky*'s [1986] covariance function. The form of $\gamma(h_j)$ is

$$\gamma(h_j) = \hat{\sigma}_Y^2 - (\hat{\sigma}_Y^2 - \sigma_0^2) \exp(-|h_j|/\lambda_j) \tag{11}$$

where the j subscript refers to a direction (i.e., $j = x, y, z$), λ_j is the integral scale of ln (K) in the jth direction, σ_0^2 is the so-called "nugget effect" and $\hat{\sigma}_Y^2$ is the variance of the process. Although (11) is a one-dimensional variogram model, it can in principle be fit to the variogram data estimated along each core transect, both vertically and horizontally, in order to determine the needed geostatistical parameters. Applied in this way, however, the influence of, say, variogram data in the horizontal y direction along B-B' on the determination of λ_x in the horizontal A-A' direction cannot be accounted for. Nevertheless, we will first fit the one-dimensional form (11) to the variogram data in order to be consistent with *Sudicky* [1986]. Subsequent to this exercise, and for comparison purposes, we will later fit a two-dimensional anisotropic variogram model [i.e., *Dagan*, 1989*b*, p. 162],

$$\gamma(h_{ij}) = \hat{\sigma}_Y^2 - (\hat{\sigma}_Y^2 - \sigma_0^2) \exp\left[-\left(\frac{h_{ij_h}^2}{\lambda_h^2} + \frac{h_{ij_z}^2}{\lambda_z^2}\right)^{1/2}\right] \tag{12}$$

to ln (K) data along the two coordinate directions on each transect in order to simultaneously determine all of the geostatistical parameters.

Sudicky [1986] found that $\sigma_0^2 = 0.10$, $\hat{\sigma}_Y^2 = 0.38$, $\lambda_x = \lambda_y = 2.8$ m and $\lambda_z = 0.12$ m by first applying a logarithmic transformation to the estimated autocorrelations and then performing a linear least squares fit using data truncated at approximately one-half the maximum lag distance.

The use of a logarithmic transformation is convenient as a linearization procedure when fitting exponential functions to data and is appropriate when the model being fitted is intrinsically linear in its parameters [*Menke*, 1984; *Draper and Smith*, 1966]. The method works as follows. Suppose a data model relationship of the form

$$\gamma^*(x) = a \exp(bx) \tag{13}$$

exists. By letting $\gamma'(x) = \ln(\gamma^*(x))$ and $a' = \ln(a)$ then

$$\gamma'(x) = a' + bx \tag{14}$$

which can be solved by standard linear regression techniques. To justify least squares in the instance, the following two assumptions must be made [*Draper and Smith*, 1966] (1) the variable x is without error, and (2) the residuals $[\gamma'(x) - (a' + bx)]$ are independent random variables following a Gaussian distribution and their variance is independent of lag distance (homoscedastic). Of course, the latter assumption is not needed to apply least squares as a model data fitting tool; however, if the Gaussian assumption is not correct, then the computed model results will be seriously affected by outliers in the data. In this situation other robust

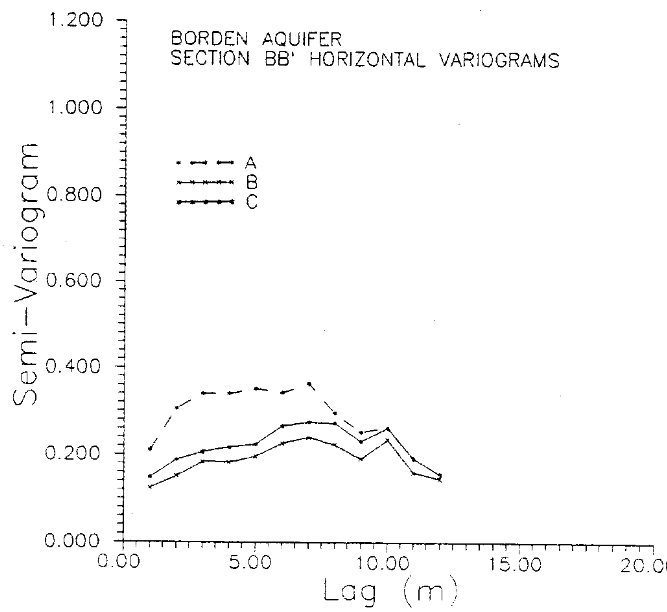

Fig. 6. B-B' horizontal variogram. A is the "classical" variogram; B, Cressie/Hawkins; C, classical variogram for data set with outliers removed.

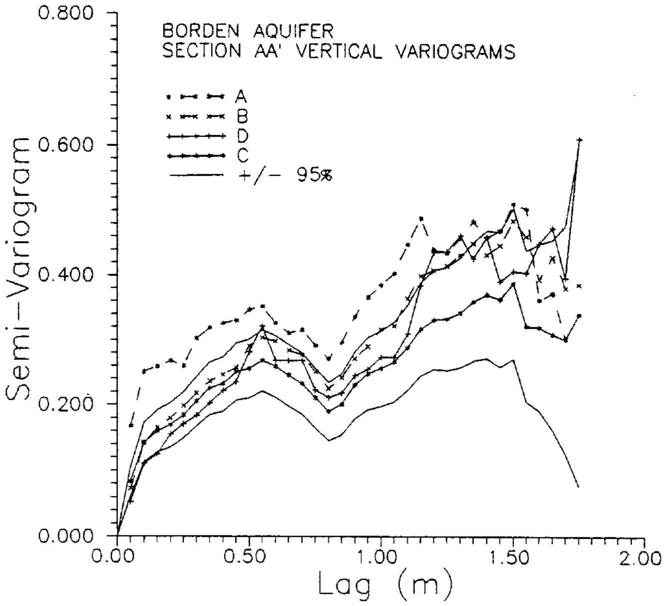

Fig. 3. Section A-A' vertical variograms. A is the "classical" variogram; B, Cressie/Hawkins; C, classical variogram for data set with outliers removed; D, SMAD estimated variogram.

variogram value for the large lag numbers for line C is about 0.10 lower than the classical estimate. Here the uncorrelated variance comes closer to the sample statistics for the screened data (Table 1*b*). All estimators do not replicate each other at large lags. Hence they all appear to be unreliable at lags greater than *L*/2.

Figure 4 shows experimental variograms for section B-B' in the vertical direction. The classical estimator is shown as line A, the Cressie/Hawkins estimator as line B, and the classical-screened estimator as line C. Again, the two solid lines without symbols are 95% confidence intervals about line C. Line A shows that outliers have an impact on the variogram at all lags. The Cressie-Hawkins (B) and C lines are quite close for about the first half of the variogram, considering the estimation errors (solid lines). Note that past 1.0 m, all variograms show large variability past about one half of their length.

Based on the above visual arguments and other considerations, we choose to adopt the classical estimator with outliers removed (for example, line C), as our preferred estimator. Note that various authors such as *Armstrong* [1984] and *Omre* [1984] recommend this approach (i.e., removing outliers) after careful examination of the data set and identification of the reasons for the outlier behavior. In our case, since a real trend or heterogeneity in the data is ruled out (see section on drift estimation) and the outliers in the data set are the result of small lenticular lenses of low hydraulic conductivity material, this approach seems justified. We believe that similar results to those obtained in subsequent sections can be achieved by using the Cressie-Hawkins estimator over about one half of the variogram length.

Figures 5 and 6 show variograms in the horizontal directions for section A-A' and B-B', respectively. The large differences between variogram estimates denoted by lines A and C for all lags illustrate the consequences of a few extreme values on variogram estimation.

It is concluded that all of the variogram estimators show erratic behavior at large lags, consistent with well-known observations [*Journel and Huijbregts*, 1978]. This presents a difficulty in determining exactly how much of the variogram record should be used in fitting a theoretical variogram model to the data. We discuss the resolution of this problem in a subsequent section on inversion results.

METHOD OF PARAMETRIC ESTIMATION

After having calculated experimental variograms, the next task is to model the experimental data with a theoretical function. In geostatistics, several functions are commonly used as variogram models. These functions include the exponential, spherical, power, linear, Gaussian and logarithmic models [see *Journel and Huijbregts*, 1978, pp. 164–171]. The functional form of a variogram (or covariance) model is likely to have some influence on the value of bulk hydraulic conductivities and macrodispersivities predicted by stochastic-analytic transport theory [e.g., *Matheron and de Marsily*, 1980; *Gelhar and Axness*, 1983]. Of all the available functions, the exponential model is postulated to be a leading candidate for the Borden aquifer based on visual inspection and physical-mechanical arguments. An exponential model is often assumed by researchers in stochastic hydrology [e.g., *Hoeksema and Kitanidis*, 1985; *Dagan*, 1989*b*, p. 169]. Some theoretical work [*Agterberg*, 1970] shows that a continuous random variable in three-dimensional space has an exponential autocorrelation function if it is subject to a property analogous to the Markov property in time series analysis. Such a model is thought to be applicable for fluvial processes. An exponential function was also used by *Sudicky* [1986] to determine model parameters describing the autocorrelation structure of the Borden aquifer and satisfactorily predicts (with Dagan's model) the observed plume spreading. Also, in the section on drift estimation, the vertical data are shown to be consistent with realizations drawn from a stochastic process with an exponential covariance. In addition to these arguments, one can also apply

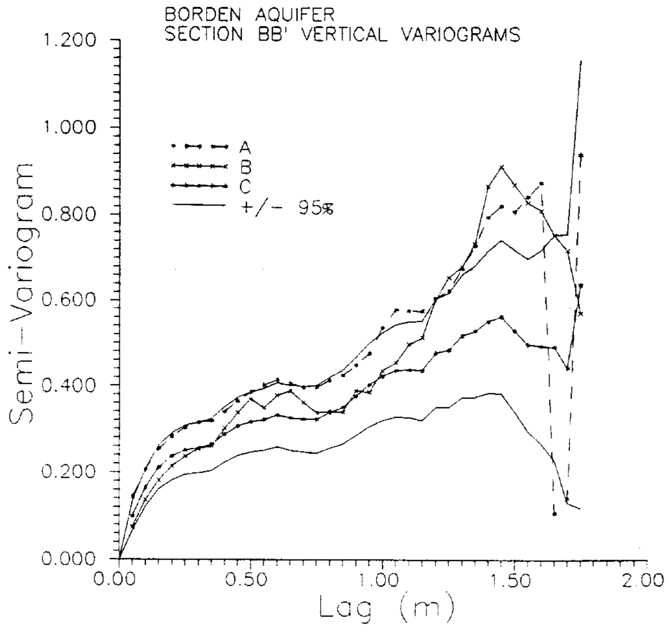

Fig. 4. Section B-B' vertical variograms. A is the "classical" variogram; B, Cressie/Hawkins; C, classical variogram for data set with outliers removed.

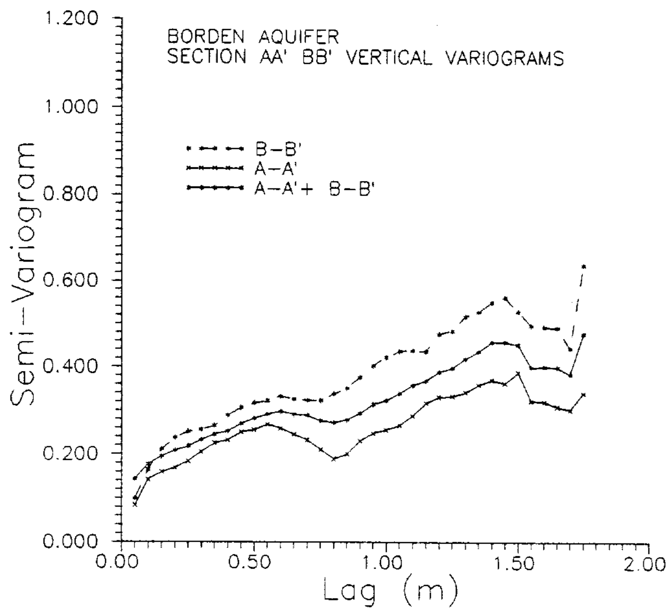

Fig. 7. Experimental vertical variograms from lines A-A′ and B-B′. Variograms computed using classical estimator with outliers removed.

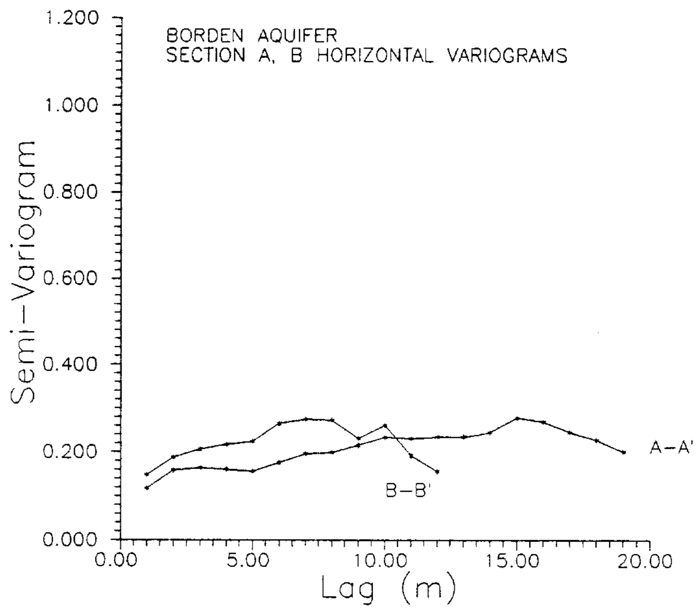

Fig. 8. Experimental horizontal variograms from lines A-A′ and B-B′. Variograms computed using classical estimator with outliers removed.

These values are found to be close to one and zero, respectively, for lines A-A′ and B-B′. Therefore, the selected model is theoretically consistent.

Note that in each case, there is a difference between the two variograms and, hence, the parameter sets determined from each line. Random samples of ln (K) taken from each line instead of the complete set as in Table 1*a* show a variance of 0.21 for A-A′ and 0.34 for B-B′. These values are close to the sill values on each variogram (see Figures 3 and 4 and Table 4). A statistical variance ratio (F) test indicates this difference is significant at the 99% confidence level.

The difference in parameter sets is further verified by performing a Hotelling's T^2 test [*Johnson and Wichern*, 1988, p. 239] on the difference of mean parameter vectors and covariances from each line. The T^2 test is similar to the χ^2 test mentioned earlier except a pooled covariance matrix is first formed. The test determines if the difference between the two mean vectors is statistically different. The test was run on the two parameter sets and the result is positive past the 99% confidence level. Therefore there is a statistical difference in geostatistical parameters from data taken from lines A-A′ and B-B′.

Recall that the horizontal directions spanned by the two transects are orthogonal; hence it is possible that anisotropies exist in this plane. However, the vertical directions spanned by the two transects are coincident and therefore integral scales and variances inferred for the vertical direction, based on two transects, should be the same. This result is disturbing to the authors, and it implies that perhaps one or all of the following conditions are responsible for the differences: (1) there were differences in experimental procedures when measuring K along A-A′ as opposed to B-B′; (2) the weak stationarity assumption regarding the variogram calculation is not true, i.e., the covariance depends on spatial location; and (3) not enough data was collected over enough integral scales on one transect compared to another. Since identical personnel and equipment was used in collecting the data set on both A-A′ and B-B′ originally, it is difficult to consider the first point as a viable explanation. The second point is a possibility since the particular depositional model of the Borden aquifer is that of an old shoreline [*Bolha*, 1986]. It is possible that lines oriented in different directions cause samples at similar depths to intersect different layers. In this case a simple form for the geometric anisotropy and/or stationarity may not be applicable. The third point seems applicable since line A-A′ is longer (19 m) than line B-B′ (12 m), although variances tend to become larger for integrations performed over more integral scales [*Journel and Huijbregts*, 1978]. In any event, line A-A′ is oriented in the direction of mean groundwater flow and therefore is the longitudinal direction of the tracer plume. The implication of this result is that a universal

TABLE 3. Optimal Variogram Lengths

Parameter	A-A′$_H$	B-B′$_H$	SUD	A-A′$_V$	B-B′$_V$	SUD
$\hat{\sigma}_Y^2$	0.422	0.267	0.38	0.245	0.354	0.38
σ_0^2	0.118	0.097	0.10	0.027	0.063	0.10
λ	7.15	2.79	2.8	0.160	0.252	0.12
AIC	−97	−50		−98	−101	
Length, m	17.	10.		1.1	0.90	

H or *V* subscript refers to horizontal or vertical direction, $\hat{\sigma}_Y$ is the variance, σ_0^2 is the "nugget" and λ is the integral scale. SUD refers to values obtained by *Sudicky* [1986].

norms are better suited [*Woodbury et al.*, 1987]. It should be noted that, for the purpose of variogram fitting, the residuals are actually χ^2 distributed and the fitting procedure involves minimizing fourth-order increments of the original data [*Cressie*, 1984]. If the Gaussian assumption is made in (14), then the distribution of the data in the original parameterization (13) must be non-Gaussian at each lag distance. In other words, if the variogram estimates are normally distributed at each lag distance, then the logarithmic transformation has the effect that the original data $\gamma^*(x)$ are measured with an apparent accuracy that increases with x. A reduction in the accuracy of variogram estimates with increasing lag distance is more likely because the number of data pairs decreases with increasing lag distance. Therefore, the appropriate choice of parameter estimation here must be of a nonlinear type. As will be shown, differences are obtained between the nonlinear fits obtained here and those of *Sudicky* [1986] although some of the differences in parameter values can be attributed to Sudicky's use of the biased form of an autocorrelation estimator.

Nonlinear Least Squares Estimation

In this study the experimental and theoretical variograms are fitted by two separate methods. The model parameters are the variance, integral scale and "nugget." The inversion approaches utilized here consist of a constrained simplex (CS) approach [*Woodbury et al.*, 1987] and the Levenberg-Marquardt (LM) method [*Press et al.*, 1986, pp. 523–528]. Both methods minimize the χ^2 misfit between observed and calculated quantities. The CS scheme easily handles hard constraints, does not require partial derivatives of the objective function with respect to the parameters, and can be employed to minimize complicated nonlinear systems. The LM scheme, being gradient based, is much faster but less robust than the CS scheme. These methods are used independently by both authors as a check on final solutions. In this way, uniqueness of the solutions is rigorously examined. Results from *Sudicky*'s [1986] analysis are used as a first trial point in the CS scheme.

The covariance of the parameter estimates is a useful end product of any inversion. However, because the inversion is nonlinear the a posteriori distribution for the parameters is potentially non-Gaussian. The reason for the non-Gaussian nature of the functional surface has been examined theoretically by *Tarantola and Valette* [1982]. In the nonlinear case the covariance of the estimated model parameters may be difficult to interpret especially in terms of confidence intervals. An additional difficulty in the nonlinear problem associated with this paper is that the parameter values have hard upper and lower constraints imposed that cannot be violated in the inversion. The implication is that the a posteriori frequency distribution of the parameters may be of a truncated nature, particularly if the expected values of the parameters lie near hard constraint boundaries. In spite of these difficulties covariances from nonlinear inversions are often estimated as for the linear case, as shown by *Menke* [1984, p. 152]. These caveats aside, we make the assumption that the a posteriori pdf of the parameter estimates is Gaussian and can be computed in the same way as the linear inversion:

$$\mathbf{C}_m \approx [\mathbf{J}^T[\mathbf{C}_d]^{-1}\mathbf{J}]^{-1} = \sigma_d^2[\mathbf{J}^T\mathbf{J}]^{-1} \tag{15}$$

where $\mathbf{C}_m$ and $\mathbf{C}_d$ are the model and data covariances. $\mathbf{C}_d$ is estimated from the data residuals and is equal to $s_d^2\mathbf{I}$, where s_d^2 is an unbiased estimator of σ_d^2 and is equal to

$$s_d^2 = \frac{1}{(N-M)} \sum_{j=1}^{N} [\gamma^*(h_j) - \gamma(h_j)]^2 \tag{16}$$

M is the number of parameters, and $\mathbf{J}$ is the Jacobian or sensitivity matrix, equal to $\partial\gamma(h_j)/\partial m_k$, where m_k are the parameters. The derivatives are evaluated at m_{opt}. Statistical inferences based on $\mathbf{C}_m$ are assumed to be valid for a normal distribution. Therefore, the a posteriori pdf of the model parameters can be directly written as

$$f(\mathbf{m}) \approx [(2\pi)^{nm}|\mathbf{C}_m|]^{-1/2} \cdot \exp\left[-\tfrac{1}{2}(\mathbf{m} - \langle\mathbf{m}\rangle)^T\mathbf{C}_m^{-1}(\mathbf{m} - \langle\mathbf{m}\rangle)\right] \tag{17}$$

where $nm = 5$, $\mathbf{m} = (\hat{\sigma}_y^2, \sigma_0^2, \lambda_x, \lambda_y, \lambda_z)$ for the general case and $\langle\mathbf{m}\rangle$ denotes expected values (values determined from the inverse).

The residuals

$$\gamma^*(h_j) - \gamma(h_j) = \eta_j \tag{18}$$

are checked for deviations from a normal distribution using a Kolmogorov-Smirnov test [*Press et al.*, 1986, p. 470].

Inversion Results

The variogram parameters are estimated as follows. First, horizontal and vertical variograms for each transect are calculated with the classical-screened estimator. Second, model parameters are obtained from the nonlinear (NL) methods on each line, horizontally and vertically. The procedure in each case is to obtain parameter sets for each line and direction by deleting one data point off the end of the data set and examining the parameters for stability and for the minimum value of the AIC information criterion. This criterion is equivalent to selecting the minimum mean square error model. In this way, the optimal length of variogram is established and the parameter set is stable. Third, and finally, once the optimal lengths of each line are determined, a simultaneous inversion is carried out in which the vertical and horizontal integral scales are determined for each line.

A-A′ and B-B′ vertical classical-screened variograms are plotted on Figure 7. As shown, the B-B′ variogram is significantly higher than that for A-A′. The middle line was obtained by "pooling" the A-A′ and B-B′ data and then computing the variogram. The horizontal variograms are plotted on Figure 8. Note that if one ignores the data past 1.0 m vertically ($L/2$), the variance for B-B′ appears higher than A-A′. This suspicion is verified by the analysis below.

Three model parameters are sought for each line in turn. These values are the variance, nugget and integral scales. Examining each variogram on each line in turn and monitoring the AIC values produces the optimal length for each variogram. This information is summarized in Table 3. Next, simultaneous fits for the variance, nugget, and horizontal and vertical integral scales are carried out on A-A′ and B-B′ separately. Table 4 shows the results of these inversions. We also validate the fitted exponential model on each line by computing mean errors and mean squared reduced errors.

second-order spatial moment tensor of a tracer cloud about its center of mass position, L, and with X_{ij} also being a function of the vector of geostatistical parameters $\mathbf{m} = (m_1, m_2, m_3, \cdots, m_k)$, then

$$\langle X_{ij}(L)\rangle = \int X_{ij}(L|\mathbf{m})f(\mathbf{m})\, d\mathbf{m} \qquad (20)$$

where $f(\mathbf{m})$ is the joint pdf of the parameter estimates. For the case of k parameters the above integral is k-dimensional. Similarly, the variance of $X_{ij}(L)$, $\sigma^2_{X_{ij}}(L)$ is given by

$$\sigma^2_{X_{ij}}(L) = \int X^2_{ij}(L|\mathbf{m})f(\mathbf{m})\, d\mathbf{m} - \langle X_{ij}(L)\rangle^2 \qquad (21)$$

Note that a prediction of plume spreading based on (20) can be quite different from that given by substituting $\langle \mathbf{m}\rangle$ without parameter uncertainty into an expression for X_{ij} because of nonlinear interactions.

Sudicky [1986] and *Freyberg* [1986] found reasonable success in deterministic application of theoretical formulae described by *Dagan* [1987] for prediction of the longitudinal X_{11}, and horizontal transverse X_{22} moments of the Borden tracer plume. Averaging ln (K) in the vertical direction over an integral scale λ_z results in a reduction of the variance of the averaged ln (K) from σ_Y^2 to $0.74\sigma_Y^2$. The equations for $X_{11}(L)$ and $X_{22}(L)$ for a horizontally isotropic medium are [see *Dagan*, 1987; *Sudicky*, 1986, equations (14*a*) and (14*b*)]

$$X_{11}(t) = 0.74(2\sigma_Y^2\lambda^2)\left\{\frac{3}{4} - \frac{3}{2}E + \frac{L}{\lambda} + \frac{3}{2}\left[\text{Ei}\,(-L/\lambda) - \ln\,(L/\lambda) + \frac{\lambda}{L}\exp\,(-L/\lambda)\left(1 + \frac{\lambda}{L}\right) - \frac{\lambda^2}{L^2}\right]\right\} \qquad (22)$$

$$X_{22}(t) = 0.74(2\sigma_Y^2\lambda^2)\left\{\frac{3}{2}\left[\frac{\lambda^2}{L^2} - \frac{\lambda}{L}\left(1 + \frac{\lambda}{L}\right)\exp\,(-L/\lambda)\right] - \frac{1}{2}[\text{Ei}\,(-L/\lambda) - \ln\,(L/\lambda)] - \frac{3}{4} + \frac{1}{2}E\right\} \qquad (23)$$

where $E = 0.577\cdots$ is the Euler constant, $\lambda = \lambda_x = \lambda_y$ is the integral scale of Y, $L = Ut$, U is the average groundwater velocity, and E$i\,(-x)$ is the exponential integral. Although U will in general be subject to uncertainty, we will assume in this analysis that U is precisely known and its value is given by

$$U = K_G J_1/n \qquad (24)$$

where J_1 is the hydraulic gradient, n is the porosity and K_G is the geometric mean hydraulic conductivity. Using values presented by *Sudicky* [1986], $U = 0.086$ m/day (see also *Naff et al.* [1988]). This value is very close to the velocity estimated by *Freyberg* [1986] on the basis of the rate of translation of the tracer plume. It should be recalled that the ln (K) variance, σ_Y^2 appearing in (22) and (23) excludes σ_0^2 because the nugget does not contribute to the dispersion process [*Dagan*, 1989*b*]. Note that the relationship of the variance computed from the variogram fitting and the process variance used in the evaluation of (22) and (23) is

$$\sigma_Y^2 = \hat{\sigma}_Y^2 - \sigma_0^2 \qquad (25)$$

In order to perform the necessary integrations in (20) and (21), the form of the joint parameter pdf must be specified. We assume the normal form given by (17). Given a mean groundwater velocity $U = 0.086$ m/day and σ_Y^2, and assuming the vertical data are averaged, then $k = 2$ with $\mathbf{m} = (\sigma_Y^2, \lambda_h)$. $\mathbf{C}_m$ becomes a 2 × 2 covariance matrix; λ_z is not used in this model. The contraction of the general $k \times k$ parameter covariance matrix obtained from variogram fitting to a 2 × 2 matrix is carried out by standard methods [e.g., *Kreyszig*, 1967, pp. 782–783] (see also the appendix). Experience has shown that numerical integration of (20) and (21) in two-dimensional parameter space for a Gaussian pdf can be performed accurately and efficiently using Gauss-Legendre quadrature with about 20 Gauss points.

Figures 9 and 10 show the expected rates of plume spreading in the longitudinal and transverse directions, $\langle X_{11}\rangle$ and $\langle X_{22}\rangle$, respectively. Also shown are the envelopes given by $\langle X_{ii}\rangle \pm 2\sigma_{X_{ii}}$. Note that these confidence limits assume that the a posteriori variances are from a Gaussian process. Because $L = Ut$, the predictions are alternatively expressed as a function of the plume residence time t and the values have been augmented by $X_{11}(t = 0)$ and $X_{22}(t = 0)$ using field data from *Freyberg* [1986]. These predictions are based on the preferred variogram model arrived at using only the A-A′ data ($\hat{\sigma}_Y^2 = 0.244$, $\sigma_0^2 = 0.072$, $\lambda_h = 5.14$, $\lambda_z = 0.209$) for reasons given earlier. The reader is reminded that A-A′ is oriented along the direction of flow in the aquifer. Also provided in Figures 9 and 10 are field estimates of X_{11} and X_{22} based on calculations performed by *Freyberg* [1986] and *Rajaram and Gelhar* [1988] for both the components of the chloride and bromide tracers in the direction of flow. The field data are clustered about $\langle X_{ii}\rangle$ at all levels of time and essentially all the data fall within the $\langle X_{ii}\rangle \pm 2\sigma_{X_{ii}}$ uncertainty envelopes. Results from these simulations are summarized in Tables 7 and 8. Numbers given in these tables represent the mean square errors between the calculated moments given by various authors [*Rajaram and Gelhar*, 1988, Table 5; *Freyberg*, 1986, Table 6]. Data for time equals 1038 days are ignored in this comparison as they are assumed anomalous (the reader is referred to discussions by *Freyberg* [1986] and *Rajaram and Gelhar* [1988] on the nature of the plume spreading at large time values). As shown in these tables, the geostatistical data provided by this paper give results that have generally smaller mean square errors than previous calculations [*Sudicky*, 1986]. The agreement between theory and field estimates of plume spreading is rather remarkable considering the numerous assumptions implicit in the analysis and the fact that no calibration has been attempted. The authors worked independently on this phase, Woodbury calculating variogram parameters and Sudicky calculating moment statistics.

It is important to recognize that the expressions used for $\langle X_{ii}\rangle$ are approximate in the sense that they reflect two-dimensional transport in a stationary vertically integrated ln (K) field and vertical tracer movement is restricted on account of primarily horizontal groundwater flow or presumed thin "impervious" beds [e.g., *Dagan*, 1988, 1989*a*]. Note also that (22) and (23) predict ensemble mean behavior and the observed moments are in fact only one realization. At later times the plumes should have effectively spread over enough integral scales to reduce this error.

An attempt to use the three-dimensional theory presented by *Dagan* [1988], with σ_Y^2, λ_h and λ_z uncertain, leads to an

TABLE 4. Simultaneous Inversion Results

Parameter	A-A′	s.d.	B-B′	s.d.	SUD
$\hat{\sigma}_Y^2$	0.244	0.008	0.366	0.019	0.38
σ_0^2	0.072	0.021	0.111	0.16	0.10
λ_h	5.14	1.17	8.33	1.68	2.8
λ_v	0.209	0.050	0.336	0.075	0.12

Here, s.d. refers to standard deviation, $\hat{\sigma}_Y^2$ is the variance, σ_0^2 is the "nugget" and λ is the integral scale. SUD refers to values obtained by *Sudicky* [1986].

TABLE 6. Synthetic Data Sets, Second Case

Parameter	Horizontal	s.d.	True
$\hat{\sigma}_Y^2$	0.333	0.007	0.380
σ_0^2	0.000*	0.035	0.000
λ	3.01	0.046	2.80

Field generated from random process of $\hat{\sigma}_Y^2 = 0.38$, $\sigma_0^2 = 0$, and $\lambda = 2.80$. See text for details.
*An a priori bound on this particular parameter was reached.

assumption about weak stationarity even for relatively homogeneous aquifers cannot be taken for granted.

Determination of Integral Scales

Part of the difficulty in using variograms to infer integral scales is that of spatial aliasing. For example, if the sample spacing is long compared to the range or correlation length then determination of a short integral scale becomes impossible [*Russo and Jury*, 1987]. To investigate if spatial aliasing is a problem with respect to the Borden data ,à series of numerical experiments is carried out. In a first simulation, a hydraulic conductivity field is generated over a regular grid of 0.05-m square blocks, 7 m in length in one direction and 0.5 m in the other, for a total of 1400 values. Hydraulic conductivity values are generated from a lognormal distribution (hydraulic conductivity in centimeters per second) with a mean log conductivity equal to −4.63 and a variance of 0.38. An exponential covariance function is adopted,

$$P_{ij} = \sigma_Y^2 \exp\left[-\left(\frac{h_{ij_h}^2}{\lambda_h^2} + \frac{h_{ij_z}^2}{\lambda_z^2}\right)^{1/2}\right] \tag{19}$$

where $h_{h,z}$ are the distances between points i, j in different directions and the λ are the integral scales. Since the covariances are defined as point quantities, one· would normally integrate the point quantities that correspond to, say, a mesh of finite elements, over the area of the element to produce an areally averaged covariance for the cell. Here, since a regular mesh is chosen, point covariance values corresponding to distances from the center of each cell are generated and applied as an average value for that cell.

A synthetic data set is sampled to resemble Sudicky's data set, i.e., 1 m separations horizontally and 0.05 m vertically, except that $\lambda_h = \lambda_z = 0.12$ m. Variograms are computed for two transects through this data base, with a basic lag of 0.05 m vertically and 1.0 m horizontally. The experimental variogram data are then used as an input to the NL codes and inverted to obtain the model parameters back again. The results are summarized in Table 5. Here the underlying "true" parameters for the vertical models are in good agreement with the estimated quantities from the NL procedure. The horizontal problem is much worse. Note large standard deviations for all parameters, indicating that estimating small integral scales from coarsely sample data is quite difficult.

The above experiment is repeated, only in this case the original field generated is anisotropic, $\lambda_h = 2.8$ m and $\lambda_z = 0.12$ m. The sample is generated on a grid 1.0 m horizontal and 0.05 m vertical. Variograms are computed for two transects through this data base, and the experimental variogram data are then used as an input to the NL codes and inverted to obtain the model parameters. The results are summarized on Table 6. Here the underlying "true" parameters for the horizontal model are in good agreement with the estimated quantities from the NL procedure. In Sudicky's original work, the horizontal spacing was 1.0 m horizontally and 0.050 m vertically. Based on the above analyses and the computed integral scales from the inversion, we conclude that the original sampling density was probably adequate.

Prediction of Tracer Spreading Under Parameter Uncertainty

We have shown that it is generally not feasible to precisely determine values for the variance and integral scales of ln (K), even when the aquifer has been sampled using a fairly dense array of measurement points. Thus, the geostatistical parameters describing the spatial variability of hydraulic conductivity must in themselves be regarded as random variables according to some joint pdf. Nevertheless, it is possible to quantify the expected spreading rate of a tracer plume and the uncertainty in the spreading rate provided that information is at hand concerning the magnitude of the uncertainty of the geostatistical parameters, possible interactions between parameters, and the form of their joint pdf [*Dagan*, 1988]. The parameter covariance matrix $\mathbf{C}_m$ determined from the nonlinear variogram fitting exercise allows us to estimate the effect of uncertainty in describing the geostatistical structure of the Borden aquifer on the spreading of an injected tracer. Rather than perform a Monte Carlo analysis, we present an approach suggested by *Dagan* [1988]. With $\langle X_{ij}(L)\rangle$ denoting the expected values of the

TABLE 5. Synthetic Data Sets, First Case

Parameter	Vertical	s.d.	True	Horizontal	s.d.	True
$\hat{\sigma}_Y^2$	0.386	0.084	0.380	0.011	15.5	0.380
σ_0^2	0.000*	0.099	0.000	0.806	0.493	0.000
λ	0.105	0.042	0.120	11.8	296.	0.120

Field generated from random process of $\hat{\sigma}_Y^2 = 0.38$, $\sigma_0^2 = 0$, and $\lambda = 0.12$. See text for details.
*An a priori bound on this particular parameter was reached.

TABLE 7. Mean Square Errors for Data From *Rajaram and Gelhar* [1988]

Simulation	Bromide X_{11}	Bromide X_{22}	Chloride X_{11}	Chloride X_{22}
This study	7.55	0.209	18.5	0.586
Sudicky [1986]	12.4	0.569	12.4	1.54

The above values refer to the mean square errors of the observed minus the computed values. The simulation data sets are from this study ($\hat{\sigma}_Y^2 = 0.244$, $\sigma_0^2 = 0.072$, $\lambda_h = 5.14$) and from *Sudicky* [1986] ($\hat{\sigma}_Y^2 = 0.38$, $\sigma_0^2 = 0.10$, $\lambda_h = 2.80$). Note $X_{11}(0) = 1.92$ and $X_{22}(0) = 2.31$.

plume in a heterogeneous aquifer. The implication here is that large-scale effective flow and transport parameters such as bulk hydraulic conductivity and macrodispersivity cannot be regarded as single-valued deterministic entities. Thus, even large-scale simulations of mass transport in groundwater based on effective macroscale parameters must be treated in a stochastic framework. Our perceived view is that of a two-tiered level of uncertainty; one due to an inability to precisely describe point-to-point variations in the local properties of a geologic medium and the other due to an inability to uniquely describe its statistical properties.

CONCLUSIONS

Our purpose is to systematically reexamine the Borden hydraulic conductivity data with particular emphasis on how various assumptions and modes of interpretation might affect the values of inferred geostatistical parameters. Our emphasis is on the determination of, and a comparison with, the spatial variance tensor calculations from the natural gradient tracer experiment at the Borden aquifer. After a complete reexamination of *Sudicky*'s [1986] field experiment the following results are obtained.

The lognormal assumption of the sampled hydraulic conductivity data is examined. A closer look at the sampled data reveals that a number of outliers (low ln (K) values) are present in the data base. As is shown, these low values cause difficulties in estimating variograms. The analysis shows that assuming either a normal distribution or exponential distribution for log conductivity is equally appropriate.

The corresponding estimator of the related autocorrelation used by *Sudicky* [1986] is a biased estimator because of the normalization of $1/n$ as opposed to $1/n(h)$. The estimator used by Sudicky tends to reduce mean square error and smooth variations at large lags. In this study, variograms are used as opposed to autocorrelations. The classical, Cressie/Hawkins and SMAD estimators are used. The classical estimator is shown to be unreliable due to the presence of outliers in the data base. Screening outliers from the data base and then using the classical estimator provides a reliable method for determining the experimental variogram. The Cressie/Hawkins estimator produces a variogram closest to the screened estimate at short lags. All estimates appear unreliable past about one half of the total length of the variogram.

In order to determine parameter values for a chosen exponential model, *Sudicky* [1986] performed a logarithmic transformation to the data values (autocorrelations in his case) and a fit was carried out by linear least squares using truncated data sets. However, the use of a logarithmic transformation is not correct in this instance. The appropriate choice of parameter estimation here must of a nonlinear (NL) type.

The NL adjustments produce noticeable differences compared to Sudicky's parameters. For the classical-screened estimated variogram, NL fits produce a variance of 0.24, nugget of 0.07, and integral scales for 5.1 horizontal and 0.21 vertical on transect A-A′. For transect B-B′ these values are 0.37, 0.11, 8.3 and 0.34. The values of variance on each line are close to their sampled values. The fitted parameter set for B-B′ data (horizontal and vertical) is statistically different than the parameter set determined for A-A′.

We evaluate a probabilistic form of *Dagan*'s [1982, 1987] equations relating geostatistical parameters to a tracer cloud's spreading moments. The equations are evaluated using the parameter estimates and covariances determined from line A-A′ as input, with a velocity of 9.0 cm/day. The results are compared with actual values determined from the field test, but evaluated by both *Freyberg* [1986] and *Rajaram and Gelhar* [1988]. The geostatistical parameters developed from this study produce an excellent fit to both sets of calculated moments. Variations about the expected values are well within the error envelope calculated from (21). For example, the mean square error for the bromide data for X_{11} for *Rajaram and Gelhar*'s [1988] parameter set is 7.55 as opposed to 12.4 for Sudicky's original parameter estimate. The geostatistical parameters satisfactorily predict the actual spread of a tracer cloud when combined with Dagan's stochastic theory of dispersion.

TABLE 8. Mean Square Errors for Data From *Freyberg* [1986]

Simulation	Bromide X_{11}	Bromide X_{22}	Chloride X_{11}	Chloride X_{22}
This study	6.47	0.292	6.79	0.698
Sudicky [1986]	14.1	0.669	13.2	0.159
Freyberg [1986]	7.39	1.09	8.01	0.135

The above values refer to the mean square errors of the observed minus the computed values. The simulation data sets are from this study ($\hat{\sigma}_Y^2 = 0.244$, $\sigma_0^2 = 0.072$, $\lambda_h = 5.14$), *Sudicky* [1986] $\hat{\sigma}_Y^2 = 0.38$, $\sigma_0^2 = 0.10$, $\lambda_h = 2.80$), and *Freyberg* [1986] (($\sigma_Y^2 = 0.24$, $\lambda_h = 2.70$). Note $X_{11}(0) = 1.80$ and $X_{22}(0) = 2.60$.

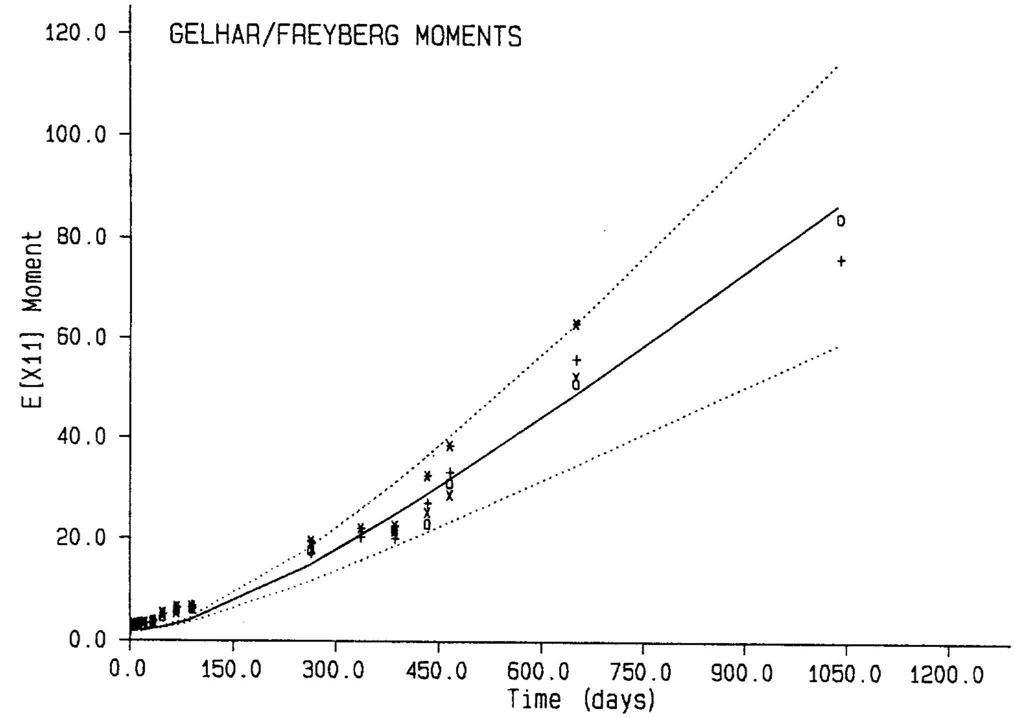

Fig. 9. Borden plume moments X_{11} from various investigators. Pluses denote bromide, and asterisks denote chloride from *Rajaram and Gelhar* [1988]; circles denote bromide, and crosses denote chloride from *Freyberg* [1986]. Solid line is computed from geostatistical parameters developed from line A-A′, incorporating uncertainty in moment calculations. Dashed lines are the upper and lower confidence intervals at 95%.

overprediction of the longitudinal spread and an underestimation of the horizontal and vertical transverse spread. While a treatment that accounts for the fluctuating nature of the flow field at the Borden site due to seasonal hydraulic gradient variations could possibly resolve this disagreement [see *Naff et al.*, 1989], the needed water level data have not yet been collected and this issue cannot therefore be resolved until they become available.

Our calculations further demonstrate (Figures 9 and 10) that even relatively small uncertainties in the values of the geostatistical parameters can lead to considerable uncertainties in the prediction of the bulk rate of spreading of a solute

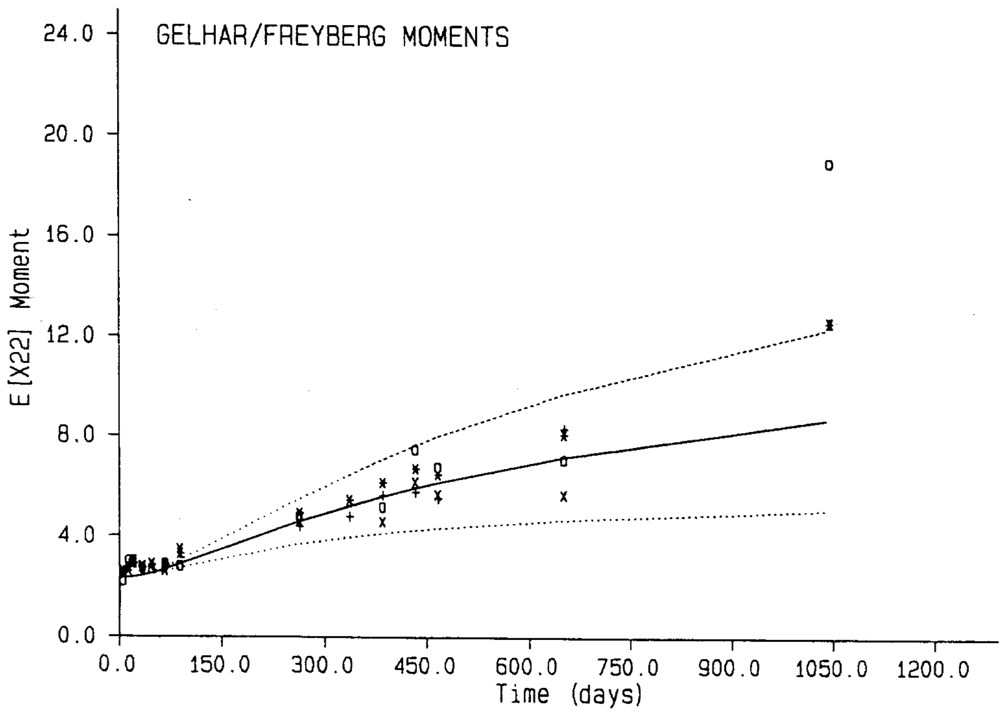

Fig. 10. Borden plume moments X_{22} from various investigators. Symbols and lines have the same interpretation as in Figure 9.

Rajaram, H., and L. W. Gelhar, Field application of stochastic models of dispersive contaminant transport, *Rep. CR-813359-01*, Robert S. Kerr Environ. Res. Lab, U.S. Environ. Prot. Agency, Ada, Okla., 1988.

Russo, D., and W. A. Jury, A theoretical study of the estimation of the correlation scale in spatially variable fields, 1, Stationary fields, *Water Resour. Res.*, *23*(7), 1257–1268, 1987.

Smith, J. L., Spatial variability of flow parameters in a stratified sand, *Math. Geol.*, *13*(1), 1–21, 1981.

Sudicky, E. A., A natural gradient experiment on solute transport in a sand aquifer: Spatial variability of hydraulic conductivity and its role in the dispersion process, *Water Resour. Res.*, *22*(13), 2069–2082, 1986.

Sudicky, E. A., Reply, *Water Resour. Res.*, *24*(7), 1211–1216, 1988.

Tarantola, A., *Inverse Problem Theory*, 613 pp., Elsevier, Amsterdam, 1987.

Tarantola, A., and B. Valette, Generalized nonlinear inverse problems using the least squares criterion, *Rev. Geophys.*, *20*(2), 219–232, 1982.

White, I., Comment on "A natural gradient experiment on solute transport in a sand aquifer: Spatial variability of hydraulic conductivity and its role in the dispersion process" by E. A. Sudicky, *Water Resour. Res.*, *24*(6), 892–894, 1988.

Woodbury, A. D., L. Smith, and W. S. Dunbar, Simultaneous inversion of hydrogeologic and thermal data, 1, Theory and application using hydraulic head data, *Water Resour. Res.*, *23*(8), 1586–1606, 1987.

Zhang, Y.-K., and S. P. Neuman, A quasi-linear theory of non-Fickian and Fickian subsurface dispersion, 2, Application to anisotropic media and the Borden site, *Water Resour. Res.*, *26*(5), 903–913, 1990.

E. A. Sudicky, Waterloo Centre for Groundwater Research, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1.

A. D. Woodbury, Department of Geological Engineering, University of Manitoba, Winnipeg, Manitoba, Canada R3T 2N2.

(Received May 4, 1990;
revised November 13, 1990;
accepted November 20, 1990.)

APPENDIX

This appendix describes how the contraction of a 4 × 4 parameter covariance matrix obtained from variogram fitting to a 2 × 2 matrix is carried out. First, note that

$$\mathbf{m} = (\hat{\sigma}_Y^2, \sigma_0^2, \lambda_h, \lambda_z)$$

However, *Dagan*'s [1987] model does not require λ_z. Hence, the first contraction of $\mathbf{C}_m$ involves removing principal and cross terms related to λ_z. Since we assume that $f(\mathbf{m})$ in (17) is Gaussian, any subset of $f(\mathbf{m})$ is also Gaussian (see *Johnson and Wichern* [1988, p. 129] for proof). The expected value of λ_z is also removed from $\langle \mathbf{m} \rangle$, and $\mathbf{C}_m$ is now 3 × 3.

Second, in *Dagan*'s [1987] model, σ_Y^2 excludes σ_0^2 because the nugget does not contribute to the dispersion process. The required relationship is

$$\sigma_Y^2 = \hat{\sigma}_Y^2 - \sigma_0^2$$

Hence, $\mathbf{C}_m$ must further be contracted with $\hat{\sigma}_Y^2 - \sigma_0^2$ terms. Suppose $\mathbf{x} = (x_1, x_2, x_3, \cdots)$ and we now desire the joint pdf of $f([x_1 - x_2], x_3, x_4, \cdots)$. Here we rely on the following relationships between, say, x_1, x_2, x_3. On the off-diagonals of $\mathbf{C}$,

$$\text{Cov}([x_1 - x_2], x_3) = \text{Cov}(x_1, x_3) - \text{Cov}(x_2, x_3)$$

On the diagonals,

$$\text{Cov}([x_1 - x_2], [x_1 - x_2]) = \text{Cov}(x_1, x_1) - 2\,\text{Cov}\, x_1, x_2) + \text{Cov}(x_2, x_2)$$

Letting $x_1 = \hat{\sigma}_Y^2$ and $x_2 = \sigma_0^2$ in the above achieves the desired result and $\mathbf{C}_m$ is 2 × 2.

Acknowledgments. This research was funded by grants from the Natural Sciences and Engineering Research Council of Canada (NSERC) to A.D. Woodbury and E.A. Sudicky. The authors would like to thank Peter Kitanidis and Gedeon Dagan for their helpful suggestions.

REFERENCES

Agterberg, F. P., Autocorrelation functions in geology, in *Geostatistics*, edited by D. F. Merriam, pp. 113–142, Plenum, New York, 1970.

Armstrong, M., Improving the estimation and modeling of the variogram, Geostatistical for Natural Resources Characterization, Part 1, *NATO ASI Ser., Ser. C*, *122*, 1–20, 1984.

Barry, D. A., Comment on "Time-dependent macrodispersion for solute transport in anisotropic heterogeneous aquifers" by G. Dagan, *Water Resour. Res.*, *26*(2), 357–359, 1990.

Bolha, J., A sedimentological investigation of a progradation foreshore sequence: C. F. B. Borden, M.Sc. thesis, Dept. of Earth Sci., Univ. of Waterloo, Waterloo, Ont., Canada, 1986.

Cressie, N., Towards resistant geostatistics, Geostatistics for Natural Resources Characterization, Part 1, *NATO ASI Ser., Ser. C*, *122*, 21–44, 1984.

Cressie, N., and D. Hawkins, Robust estimation of the variogram, *Math. Geol.*, *12*(2), 115–125, 1980.

Dagan, G., Stochastic modeling of groundwater flow by unconditional and conditional probabilities, 2, The solute transport, *Water Resour. Res.*, *18*(4), 835–848, 1982.

Dagan, G., Theory of solute transport by groundwater, *Annu. Rev. Fluid Mech.*, *19*, 183–215, 1987.

Dagan, G., Time-dependent macrodispersion for solute transport in anisotropic heterogeneous aquifers, *Water Resour. Res.*, *24*(9), 1491–1500, 1988.

Dagan, G., Comment on "A note on the recent natural gradient experiment at the Borden site" by R. L. Naff, T.-C. Jim Yeh, and M. W. Kemblowski, *Water Resour. Res.*, *25*(12), 2521–2522, 1989*a*.

Dagan, G., *Flow and Transport in Porous Formations*, 465 pp., Springer-Verlag, New York, 1989*b*.

Dagan, G., Reply, *Water Resour. Res.*, *26*(2), 360, 1990.

Dowd, P. A., The variogram and kriging: Robust and resistant estimators, Geostatistics for Natural Resources Characterization, Part 1, *NATO ASI Ser., Ser. C*, *122*, 91–107, 1984.

Draper, N., and H. Smith, *Applied Regression Analysis*, John Wiley, New York, 1966.

Freeze, R. A., A stochastic-conceptual analysis of one-dimensional groundwater flow in nonuniform homogeneous media, *Water Resour. Res.*, *11*(5), 725–742, 1975.

Freyberg, D. L., A natural gradient experiment on solute transport in a sand aquifer, 2, Spatial moments and the advection and dispersion of nonreactive tracers, *Water Resour. Res.*, *22*(13), 2031–2046, 1986.

Gelhar, L. W., and C. L. Axness, Three-dimensional stochastic analysis of macrodispersion in aquifers, *Water Resour. Res.*, *19*(1), 161–180, 1983.

Hoeksema, R. J., and P. K. Kitanidis, Comparison of Gaussian conditional mean and kriging estimation in the geostatistical solution to the inverse problem, *Water Resour. Res.*, *21*(6), 825–836, 1985.

Jenkins, G. M., and D. G. Watts, *Spectral Analysis and Its Applications*, Holden-Day, San Francisco, Calif., 1968.

Johnson, R. A., and D. W. Wichern, *Applied Multivariate Statistical Analysis*, 2nd ed., 607 pp., Prentice-Hall, Englewood Cliffs, N. J., 1988.

Journel, A. G., Imaging of spatial uncertainty: A non-Gaussian approach, in *Geostatistical, Senitivity, and Uncertainty for Ground-Water Flow and Radionuclide Transport Modeling*, edited by B. E. Buxton, pp. 585–599, Battelle, Columbus, Ohio, 1990.

Journel, A. G., and C. J. Huijbregts, *Mining Geostatistics*, Academic, San Diego, Calif., 1978.

Kemblowski, M. W., Comment on "A natural gradient experiment on solute transport in a sand aquifer: Spatial variability of hydraulic conductivity and its role in the dispersion process" by E. A. Sudicky, *Water Resour. Res.*, *24*(2), 315–317, 1988.

Kreyszig, E., *Advanced Engineering Mathematics*, 2nd ed., 898 pp., John Wiley, New York, 1967.

Mackay, D. M., D. L. Freyberg, P. V. Roberts, and J. A. Cherry, A natural gradient experiment on solute transport in a sand aquifer, 1, Approach and overview of plume movement, *Water Resour. Res.*, *22*(13), 2017–2030, 1986.

Matheron, G., Principles of geostatistics, *Econ. Geol.*, *58*, 1246–1266, 1963.

Matheron, G., and G. de Marsily, Is transport in porous media always diffusive? A counterexample, *Water Resour. Res.*, *16*(5), 901–917, 1980.

Menke, W., *Geophysical Data Analysis: Discrete Inverse Theory*, Academic, San Diego, Calif., 1984.

Molz, F. J., and O. Güven, Comment on "A natural gradient experiment on solute transport in a sand aquifer: Spatial variability of hydraulic conductivity and its role in the dispersion process" by E. A. Sudicky, *Water Resour. Res.*, *24*(7), 1209–1210, 1988.

Naff, R. L., T.-C. J. Yeh, and M. W. Kemblowski, A note on the recent natural gradient experiment at the Borden site, *Water Resour. Res.*, *24*(12), 2099–2104, 1988.

Naff, R. L., T.-C. J. Yeh, and M. W. Kemblowski, Reply, *Water Resour. Res.*, *25*(12), 2523, 1989.

Neuman, S. P., and Y.-K. Zhang, A quasi-linear theory of non-Fickian and Fickian subsurface dispersion, 1, Theoretical analysis with application to isotropic media, *Water Resour. Res.*, *26*(5), 887–902, 1990.

Neville, A. M., and J. B. Kennedy, *Basic Statistical Methods*, 325 pp., International Textbooks, Scranton, Pa., 1964.

Omre, H., The variogram and its estimation, Geostatistics for Natural Resources Characterization, Part 1, *NATO ASI Ser., Ser. C*, *122*, 107–126, 1984.

Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes: The Art of Scientific Computing*, Cambridge University Press, New York, 1986.