J. Korean Math. Soc. **38** (2001), No. 3, pp. 523–559

A PRACTICAL THREE-DIMENSIONAL ESTIMATION TECHNIQUE FOR SPATIAL DISTRIBUTION OF GROUNDWATER CONTAMINANT CONCENTRATIONS*

RICHARD EWING, SUNGKWON KANG**,
JEONGOOK KIM, AND THOMAS B. STAUFFER

ABSTRACT. To predict the fate of groundwater contaminants, accurate spatially continuous information is needed. Because most field sampling of groundwater contaminants are not conducted spatially continuous manner, a special estimation technique is required to interpolate/extrapolate concentration distributions at unmeasured locations. A practical three-dimensional estimation method for *in situ* groundwater contaminant concentrations is introduced. It consists of two general steps: estimation of macroscopic transport process and kriging. Using field data and nonlinear optimization techniques, the macroscopic behavior of the contaminant plume is estimated. A spatial distribution of residuals is obtained by subtracting the macroscopic transport portion from field data, then kriging is applied to estimate residuals at unsampled locations. To reduce outlier effects on obtaining correlations between residual data which are needed for determining variogram models, the $\mathcal{R}_p$-*estimator* is introduced. The proposed estimation method is applied to a field data set.

Received February 24, 2000.

2000 Mathematics Subject Classification: 86A32, 65U05.

Key words and phrases: groundwater contaminants, geostatistics, $\mathcal{R}_p$-estimator, kriging.

*The paper has not been subject to USAF review, however, and accordingly does not necessarily reflect the reviews of the USAF. A part of this work was presented at the Conference on Control of Distributed Parameter and Stochastic Systems, Hangzhou, China, June 19-22, 1998. **The research of this author was supported by Korea Research Foundation Grant (KRF-99-041-D00075).

1. Introduction

Groundwater contamination is an important environmental issue. Researchers have conducted extensive field experiments to analyze geophysical, chemical and biological processes that control the fate and movement of groundwater contaminants (see, [12, 17, 23, 25, 26] for the Borden test site, Canada, [13, 15] for the Cape Cod site, Massachussetts, and [1, 4-6, 16, 22] for the Columbus site, Mississippi).

In order to describe and predict underlying physical, chemical and biological processes affecting chemical fate and transport, accurate spatially continuous information is needed. Because most field sampling of groundwater contaminants are not conducted spatially continuous manner, a special estimation technique is required to interpolate/extrapolate contaminant concentration distributions at unmeasured locations. These interpolations/extrapolations are complicated by uncertainties often associated with an unknown distribution of contaminant fluxes in space and time reflected in a complex velocity field within a heterogeneous aquifer.

Many geostatistical techniques have been developed for estimating geophysical/chemical parameters and groundwater contaminant concentrations (see, for example, [2, 7, 19, 27, 28] and references there in). But, application of these estimation methods to groundwater contaminant data often fail to obtain satisfactory results because methods are based on the *geostatistical intrinsic assumption* [14], and because field data behave irratically or contain "outliers." The geostatistical intrinsic hypothesis (or stationary assumption) is that spatial correlationship between data points depend only on the separation vector (modulus and direction) and not on the individual sample location. But, the global behavior of contaminant plume follows dynamical process governed by groundwater flow; consequently, the concentration of contaminant strongly depends on sample location. Questions concerning the locations or regions of high pollutant concentrations and contaminant plume dimensions are key issues in environmental concerns. The first obstacle, the intrinsic geostatistical hypothesis, can be overcome by extracting the *plume macroscopic behavior* from the field data. The macroscopic behavior is a large scale behavior of the plume and estimated by approximating the solution of the contaminant transport equation. The concept of macroscopic plume behavior is essentially similar to that of

drift or *trend* in geostatistics in the sense of nonstationarity. In geostatistics, the general profile of most regionalized variables is assumed to be stationary, and, hence, the slowly varying minor nonstationary components (drift or trend) observed in the field may be approximated by lower order polynomials. The greater portion of a goundwater contaminant plume exhibits nonstationary characteristics. Thus, a major component of the plume can be estimated from dynamical processes expressed in the governing transport equations. This macroscopic plume behavior can be measured in a large region. On the other hand, the drift should be estimated in a "small neighborhood" of the point where kriging to be performed. Additional problems arise when concentrations are estimated at locations where measurements are not available. For example, conventional semivariogram is too sensitive to obtain correct spatial correlations for data exhibiting a wide variance. These correlations are needed for determining variogram models in a kriging procedure. The log transformation commonly used to compress data variance contains a logical conflict between original data structure and application of kriging algorithm. Consequently, to make spatial interpolations of data exhibiting a large variance, there is a need to develop a new robust estimator. This paper introduces the $\mathcal{R}_p$-*estimator*. It is consistent and robust. These properties are taken from combining theoretical concepts that define the norm of Banach space L^p, $0 < p \leq 1$, [24] and the semivariogram in geostatistics. To obtain correlation between data points, geometric anisotropies such as *symmetry* and *ratio* between the three-dimensional coordinate system are also considered. The symmetry is inherent from the groundwater contaminant distribution while the ratio evolves from the geological structure of the aquifer material.

The general procedure of the estimation method is following: (1) Divide the field site into several subregions based on all available information. (2) In each subregion, macroscopic plume behaviors (or deterministic transport components) are estimated from the field data. These estimated values are subtracted from the field data to obtain residuals. (3) Based on complexities of spatial distribution of the residuals, divide each subregion into several small blocks. (4) In each block, calculate experimental variograms using the robust $\mathcal{R}_p$-estimator and determine mathematical variogram models. (5) Perform kriging to estimate residual at each desired location. (6) Finally, combine kriged residual values with the estimated macroscopic transport components.

The purposes of this paper are following: first, to provide a systematic methodology for estimating *in situ* groundwater contaminant concentrations, second, to introduce the $\mathcal{R}_p$-*estimator* for producing correlations between data points characterized by large variance, and, third, to generate a "complete" data base on a uniform grid. This method can be used generally to estimate space and time dependent geophysical, chemical and biological parameters; thus it may be useful to those developing numerical models for capturing the main feature of the groundwater contaminant distributions.

In Section 2, basic principles for estimating macroscopic plume bahavior are explained. A robust estimator is introduced in Section 3, and anisotropies such as symmetry, ratio and mathematical variogram models are discussed in Section 4. The punctual kriging method is explained in Section 5, and our proposed estimation method is applied to a field data set in Section 6.

2. Estimation of macroscopic plume behavior

2.1. Regionalization

As the first step for estimating the global plume behavior represented in the field data, the field site is divided into several subregions. The size of each subregion strongly depends on the geological structure of the field site and the global characteristics of data distributions. Each region showing distinctive distribution behavior is contained in a separated subregion. All available field data are visualized and analyzed. Also, any information related to the field site such as geological aquifer history are incorporated.

2.2. Estimation of macroscopic plume behavior

The distribution of groundwater contaminant concentrations may be approximately by a solution of the transport equation describing many complex processes. Each sampling set represents a single realization of such complex physical, chemical and biological activities. Spatial and temporal distribution of reactive and nonreactive tracers could be used to indirectly assess aquifer geophysical characteristics and groundwater transport processes. Nonreactive tracers (i.e., those which are nonsorbing and conservative) are particularly useful for characterizing dispersion and advection because they are not retained and as such provide

transport related information over larger aquifer scales than reactive tracers. In this step a *macroscopic behavior* of the plume is estimated from field data. This is essentially a similar process as detrending the data. The macroscopic behavior is a spatially continuous large scale behavior describing the main profile of the plume movement. This step is difficult because the parameters in the transport equations, such as, dispersion/diffusion and seepage velocity share a highly nonlinear interdependence in space and time. Also, the measured contaminant concentration themselves add uncertainties due to spatial variabilities and unequal analytical confidence intervals, i.e., higher confidence in larger concentration and less confidence at near background levels. In this step the criteria on how to choose basis functions and some specific approximation functions to estimate the macroscopic plume behavior are provided. These approximating functions are chosen based on the global behavior of the solute transport process in porous media. Since nonlinear optimization techniques are needed for estimating parameters appeared in the approximating functions, general principles for selecting the initial parameters are given. These principles prevent the failure of the convergence of data fitting iteration process and also reduce the computing time considerably.

2.2.1. Approximating functions

Several types of basis functions are examined for approximation of the dynamic processes represented in the field data. For example, Gaussian features often observed in contaminant plumes can be approximated by linear combinations of exponential functions, which is a characteristic feature of solutions to parabolic partial differential equations. Using selected basis functions, the macroscopic plume behavior in each subregion is estimated from the field data. Basis functions are chosen using the following general criteria:

(1) "Simple" functions are preferred because they are easily evaluated. At the same time, a "low order" approximation should capture the main profile of the plume movement. Here, the order of the approximation refers to the number of basis functions needed for the approximation. After the primary features of the plume have been described, the residuals, derived from subtracting the macroscopic behavior from the field data, are calculated to obtain a new stationary random field; this new random field lends itself to existing geostatistical analysis techniques or newer more robust estimation procedures which then estimate residuals at unsampled locations. Ideally an entire plume can be cre-

ated over an entire site by adding the macroscopic transport portion to estimated residuals. Use of higher order approximation for the plume macroscopic behavior is undesirable, because the estimation procedure tries to fit the given data based on a least squares fit, so, eventually, at higher orders the fit narrows to a point by point basis. This means that as the number of basis functions increases the approximating function tries to capture the complex plume behavior caused by the geological heterogeneity and measurement artifacts instead of capturing the global behavior of the plume. Often in practice, unexpected or unwanted irregularities appear when high order polynomials or sinusoidal functions are used for basis functions. Major contributors to this irregularity are the field heterogeneity, the uneveness of sampling network, and the dependence of approximation or fitting algorithms on the intrinsic properties (such as frequencies in sinusoidal functions) of the selected basis functions. Therefore, to select basis functions for estimating the macroscopic plume behavior, both data fitting and global data trends must be considered.

(2) The approximation functions must capture the global plume behavior outside sampling network. In many practical applications, sampling networks often don't cover the entire extent of the contaminant plume; however, extrapolation of plume movement outside the sampling network is often desired.

(3) Basis functions should be "robust" in the sense that they are not very sensitive to unevenly spaced data points.

According to the selection criteria described above, specific basis functions are chosen based on the available field data and the process to be described (e.g., transport of a tracer or the spatial distribution of aquifer permeability). Focusing on the problem at hand, advective, dispersive/diffusive solute transport in porous media depends on space, time, and solute concentration. If, in addition, recharge, chemical, biological, and other reactive processes are considered, then the solute transport may be approximated [11], in Cartesian coordinates, using

$$\frac{\partial C}{\partial t} = \nabla \cdot (D \cdot \nabla C) - \nabla \cdot (VC) + f \tag{2.1}$$

with appropriate initial and boundary conditions, where $C(x, y, z; t)$ is the concentration of the solute, i.e., the mass of solute per unit volume of fluid, $D = D(x, y, z; t; C)$ is the dispersion tensor, $V = V(x, y, z; t; C)$ is the pore water velocity vector, $f = f(x, y, z; t; C)$ is a "forcing function" related to recharge, chemical, and biological activities. Note that

the coefficients D and V depend on space, time and concentration itself. Even though equation (2.1) is a nolinear partial differential equation, the solution, whenever it exists, can be approximated by dispersion and propagation processes. Chemical and biological activities can affect advection/dispersion and complicate the transport description; nevertheless, these dispersion/diffusion and propagation processes are inherent properties of the evolution system governed by a parabolic partial differential equation [20].

As a simple case, assume that the porous medium is homogeneous, isotropic, saturated, the flow is steady-state, and that there is no external source. Then the transport equation (2.1) can be simplified as

(2.2)
$$\frac{\partial C}{\partial t} = \left[D_x \frac{\partial^2 C}{\partial x^2} + D_y \frac{\partial^2 C}{\partial y^2} + D_z \frac{\partial^2 C}{\partial z^2}\right] - \left[\bar{v}_x \frac{\partial C}{\partial x} + \bar{v}_y \frac{\partial C}{\partial y} + \bar{v}_z \frac{\partial C}{\partial z}\right],$$

where D_x, D_y and D_z are dispersion coefficients in the x, y and z-directions, $\bar{v}_x$, $\bar{v}_y$ and $\bar{v}_z$ are the average linear pore water velocities in each coordinate direction defined by $\bar{v}_x = v_x/\phi$, $\bar{v}_y = v_y/\phi$, $\bar{v}_z = v_z/\phi$, in which v_x, v_y, and v_z are specific discharge components, and ϕ is the porosity of the medium.

If a contaminant is released instantaneously at the origin $(x, y, z) = (0, 0, 0)$, the mass distribution of the contaminant at time t is given by

(2.3)
$$C(x, y, z; t) = \frac{M}{8(\pi t)^{3/2}\phi\sqrt{D_x D_y D_z}} \exp\left(-\frac{(x - \bar{v}_x t)^2}{4D_x t} - \frac{(y - \bar{v}_y t)^2}{4D_y t} - \frac{(z - \bar{v}_z t)^2}{4D_z t}\right),$$

where M is the mass of contaminant introduced at the point source [11]. The averaged pore water velocities $\bar{v}_x$, $\bar{v}_y$ and $\bar{v}_z$ contribute movement of the center of mass of the contaminant plume (the propagation process). D_x, D_y, and D_z contribute to the longitudinal and transverse spreading of the plume around the plume centroid (the dispersion/diffusion process). The solution (2.3) of the ideal equation (2.2) is a simple representation of these two processes throughout two parameter sets $V = (\bar{v}_x, \bar{v}_y, \bar{v}_z)$ and $D = (D_x, D_y, D_z)$.

To approximate the gross distribution of contaminant concentrations in space, we propose the following linear combination of exponential

functions.

$$F(x,y,z;a,b,c) = \sum_{i=1}^{m} c_i \exp\left(-\left(\frac{x-a_i^x}{b_i^x}\right)^2 - \left(\frac{y-a_i^y}{b_i^y}\right)^2 - \left(\frac{z-a_i^z}{b_i^z}\right)^2\right), \tag{2.4}$$

where m is the number of basis functions, a_i^x, a_i^y, a_i^z, b_i^x, b_i^y, b_i^z, and c_i, $1 \le i \le m$, are parameters to be determined.

In the same context as transport equations (2.1) and (2.2), the parameter set $a = \{(a_i^x, a_i^y, a_i^z) : 1 \le i \le m\}$ represents the propagation or advection process, the set $b = \{(b_i^x, b_i^y, b_i^z) : 1 \le i \le m\}$ represents the dispersion/diffusion process, and the set $c = \{c_i : 1 \le i \le m\}$ is related to the magnitude of the source load.

2.2.2. Initial parameters

Using equation (2.4) to describe the macroscopic plume behavior will require use of a nonlinear optimization technique to estimate parameter sets a, b and c. The Levenberg-Marquart method (see, for example, [18,21]) is commonly used for estimating the parameters. This method is slow, but not sensitive to complex data distribution such as field data. For other nonlinear optimization techniques such as the conjugated gradient method, the quasi-Newton method, or the hybrid methods, see [3, 9, 19].

Most nonlinear or linear optimization algorithms for parameter estimation depend strongly on the initial input parameters, because performance indices (or cost functions) are, in general, nonlinear in parameter variables, and may have many local minima. Thus, it is important to select initial parameters which are close to the global minimum. The accuracy and efficiency of numerical approximation depend on the choice of algorithm, basis functions, initial parameters, and, of course, computing machines used. The following basic principles on how to select the initial parameters for a given nonlinear optimization algorithm can be considered.

(1) Choose the initial parameters so that the first basis function captures the main dominated profile, the second basis function covers the second dominated part, and so on. This principle reduces the number of basis functions as well as the computational time for the approximation with the given desired accuracy.

(2) Provide *a priori* information observed from the data distribution to the nonlinear optimization algorithm through the initial input data. This principle prevents the failure of the convergence of data fitting iteration process (optimization process) and also reduces the processing time considerably.

3. $\mathcal{R}_p$-estimator

The residuals are obtained by subtracting the macroscopic transport portion from field data. The experimental semivariogram or variogram is used to describe the pattern of spatial correlation displayed by the residuals. A mathematical model is fitted to this experimental variogram, and this model is used in kriging to estimate the residuals at unmeasured locations. In this section, a robust estimator is introduced. This estimator is designed for producing the spatial correlation (experimental variogram) displayed by the data which contains large random components and outliers. Thus, this estimator can be applied directly to other parameters such as hydraulic conductivities and soil sorption coefficients which often showing erratic behavior.

Due to physical and chemical heterogeneities and other uncertainties, most field data reflects a large random component. The conventional semivariogram is too sensitive to obtain correct spatial correlations for data exhibiting a wide variance. The log transformation commonly used to compress data variance contains a logical conflict between original data structure and application of kriging algorithm as described in Section 5. Consequently, to make spatial interpolations of data showing a large variance, there is a need to develop a systematic robust estimator. This paper introduces the "*$\mathcal{R}_p$-estimator*," where $\mathcal{R}$ stands for "robust" and $p > 0$ indicates the order of robustness. For $0 < p \leq 1$, the estimator is robust; whereas for $p > 1$ the estimator becomes sensitive to apparent outliers. The estimator satisfies the following properties:

(1) The estimator is consistent such that spatial correlations among the original data are preserved under linear transformation. Thus, the original data structure, estimation of correlations, and kriging procedures are consistent.

(2) The estimator is robust. It reduces outlier effects on estimated correlations between data points.

(3) The estimator is systematic. Depending on the distribution of the data, the order p of robustness can be adjusted systematically.

Let $Z(\mathbf{x})$ be a regionalized function on a domain Ω in three dimiensional space, and $Z(\mathbf{x}_i)$ be the realization of the function $Z(\mathbf{x})$ at $\mathbf{x}_i = (x_i, y_i, z_i) \in \Omega$, $i = 1, 2, \cdots, n$. Let $p > 0$ be a positive real number. For any vector $\mathbf{h} = (h_x, h_y, h_z)$, we define the $\mathcal{R}_p$-*estimator* as

$$\mathcal{R}_p(\mathbf{h}) = \left[\frac{1}{n(\mathbf{h})} \sum_{i=1}^{n(\mathbf{h})} |Z(\mathbf{x}_i) - Z(\mathbf{x}_i + \mathbf{h})|^p \right]^{\frac{1}{p}}, \tag{3.1}$$

where $n(\mathbf{h})$ is the number of data pairs separated by the vector $\mathbf{h}$.

For any positive integer k, and for any positive real number $p > 0$, the following inequalities hold:

$$\begin{aligned} a_1^p + a_2^p + \cdots + a_k^p &\le (a_1 + a_2 + \cdots + a_k)^p, \quad p \ge 1, \\ a_1^p + a_2^p + \cdots + a_k^p &\ge (a_1 + a_2 + \cdots + a_k)^p, \quad 0 < p \le 1, \end{aligned} \tag{3.2}$$

where $a_i \ge 0$, $i = 1, 2, \cdots, k$. Thus, the function

$$(a_1, a_2, \cdots, a_k) \mapsto \frac{a_1^p + a_2^p + \cdots + a_k^p}{k} \tag{3.3}$$

is a convex function for $p \ge 1$ and a concave function for $0 < p \le 1$, where the symbol $\mapsto$ represents the mapping of an element of $\mathbf{R}^k$ to the right hand side of equation (3.3) and $\mathbf{R}^k$ is the k-dimensional Euclidean vector space. As $p > 0$ approaches 0, robust effects are increased and the estimator $\mathcal{R}_p$ defined by equation (3.1) reduces effects of outliers for $0 < p \le 1$.

Moreover, it is easy to see that, for any positive constant $c > 0$,

$$\begin{aligned} &\left[\frac{1}{n(\mathbf{h})} \sum_{i=1}^{n(\mathbf{h})} |cZ(\mathbf{x}_i) - cZ(\mathbf{x}_i + \mathbf{h})|^p \right]^{\frac{1}{p}} \\ =& c \left[\frac{1}{n(\mathbf{h})} \sum_{i=1}^{n(\mathbf{h})} |Z(\mathbf{x}_i) - Z(\mathbf{x}_i + \mathbf{h})|^p \right]^{\frac{1}{p}}, \end{aligned} \tag{3.4}$$

and, hence, the estimator $\mathcal{R}_p$ preserves any scaling factor in the original data. Therefore, if the original data set has a large range of values, then the range can be scaled by multiplying by a fixed positive constant without destroying the correlation structure found within the original data. The Cressie-Hawkins robust estimator [8]

$$\gamma_{ch}(\mathbf{h}) = \frac{1}{2} \frac{\left[\frac{1}{n(\mathbf{h})} \sum_{i=1}^{n(\mathbf{h})} |Z(\mathbf{x}_i) - Z(\mathbf{x}_i + \mathbf{h})|^{\frac{1}{2}}\right]^4}{[0.457 + 0.494/n(\mathbf{h})]} \tag{3.5}$$

which is commonly used in practice, the squared median of the absolute deviations estimator [10]

$$\gamma_{smad}(\mathbf{h}) = 2.198 \times [median\, |Z(\mathbf{x}_i) - Z(\mathbf{x}_i + \mathbf{h})|]^2, \tag{3.6}$$

and the conventional semivariogram [14]

$$\gamma(\mathbf{h}) = \frac{1}{2\, n(\mathbf{h})} \sum_{i=1}^{n(\mathbf{h})} |Z(\mathbf{x}_i) - Z(\mathbf{x}_i + \mathbf{h})|^2 \tag{3.7}$$

are essentially similar to the $\mathcal{R}_p$-estimator with $p = 1/2$, $p = 1$, and $p = 2$, respectively. However, the semivariogram γ is not robust. The influence of outliers on the semivariogram γ increases by the square $|Z(\mathbf{x}) - Z(\mathbf{x}+\mathbf{h})|^2$ as the difference $|Z(\mathbf{x}) - Z(\mathbf{x}+\mathbf{h})|$ increases. The Cressie-Hawkins estimator γ_{ch} and the squared median of the absolute deviations estimator γ_{smad} are not robust enough so that they do not produce correct correlation between data points showing erratic behaviors which are commonly observed in field data. Moreover, they do not preserve scaling factors and are not systematic.

4. Symmetry, ratio and variogram models

Symmetry is one of typical characteristics of groundwater contaminant distributions. Since the contaminant follows the groundwater flow path, for example, in y-direction, its distribution has certain symmetric behaviors along the transverse directions, x- and z-directions, due to dispersion and diffusion processes. Primary transport processes are exhibited the spatial distributions of groundwater plume constituents. These primary processes can be estimated deterministically and then

extracted from field data as described in Section 2. The residuals subsequently obtained still produce a symmetric distribution.

Dispersion/diffusion ratio between different directions are often observed. This ratio is incorporated with the geological structure of the aquifer material and shows similarity along the same aquifer. To estimate this ratio, choose a data range which can represent the main dispersion/diffusion profile of the data distribution in the region considered. The distance covered by the selected data range is measured in each direction. The ratio is obtained by comparing these distances. If the ratio is, for example, $x : y : z = 2 : 4 : 1$ in meter scale, any two points separated approximately $2m$ distance in x-direction, $4m$ distance in y-direction, and $1m$ in z-direction are considered to have the same correlation in the average sense. Symmetry and ratio are geometric anisotropies [14] resulted from the interaction between contaminant, groundwater flow, and aquifer materials.

The next step is to fit the experimental variogram by a mathematical model. The spherical model

$$\gamma(h) = \begin{cases} c0 + \beta\left(\frac{3h}{2\alpha} - \frac{h^3}{2\alpha^3}\right), & h \le \alpha, \\ c0 + \beta, & h > \alpha, \end{cases} \tag{4.1}$$

and the exponential model

$$\gamma(h) = c0 + \beta\left(1 - exp(-\frac{h}{\alpha})\right), \tag{4.2}$$

are commonly used, where $h = |\mathbf{h}|$ is the radius of the vector $\mathbf{h}$, and the parameters $c0$, the nugget effect, $c0 + \beta$, the sill value, and α, the distance h at which variogram reaches sill value, are to be determined by fitting a model to the experimental variogram. Other models such as linear or logarithmmic models can be found in [14].

5. Kriging

Kriging is to estimate the variables at unmeasured locations. It uses the mathematical model variograms fit to experimental variograms. Many kriging methods are available. Among them, the universal kriging and the punctual kriging are simple and can be easily implemented. In

this section, the punctual kriging is explained when the experimental variograms are obtained by the $\mathcal{R}_p$-estimator. The application of the estimator to the universal kriging can be done in a similar way. The preservation of scaling factors and the consistency of the estimator are explained. Logical conflict of using the log transformation to compress data variance is also explained.

For each i, $i = 1, 2, \cdots, m$, let $Z(\mathbf{x}_i)$ be a given value at location $\mathbf{x}_i = (x_i, y_i, z_i)$ that is selected for kriging. For a given location $\mathbf{x}_o = (x_o, y_o, z_o)$, assume that the value $Z(\mathbf{x}_o)$ at $\mathbf{x}_o$ can be approximated by a linear sum of known values $Z(\mathbf{x}_i)$, $i = 1, \cdots, m$. Let

$$Z(\mathbf{x}_o) = \sum_{i=1}^{m} \omega_i Z(\mathbf{x}_i), \tag{5.1}$$

where $\omega_i \geq 0$, $i = 1, \cdots, m$, are weights to be determined by the following kriging system:

$$\begin{aligned} \sum_{j=1}^{m} \omega_j \mathcal{R}_p(h_{ij}) + \lambda &= \mathcal{R}_p(h_{io}), \quad 1 \leq i \leq m, \\ \sum_{i=1}^{m} \omega_i &= 1, \end{aligned} \tag{5.2}$$

where $\mathcal{R}_p(h_{ij})$ is the correlation value estimated by the $\mathcal{R}_p$-estimator at lag h_{ij}, the subscript p is the order of robustness, h_{ij} is the "correlation lag" between two points $\mathbf{x}_i$ and $\mathbf{x}_j$, λ is the Lagrange multiplier, and $\sum_{i=1}^{m} \omega_i = 1$ is the optimality condition. Here, by the correlation lag, we mean the distance between two points in which the symmetry and ratio described in Section 4 are accounted. If the semivariogram $\gamma(h)$ (see, equation (3.7)) is used for variogram estimation, then $\mathcal{R}_p(h_{ij})$ is raplaced by $\gamma(h_{ij})$ in the kriging equations (5.2). Note that this kriging system (5.2) is "optimal" in the sense that the method produces the exact (original) value at the sampled location. However, the kriging system is optimal only inside the sampling network (convex) domain. Thus, the estimation procedure for points outside the sampling domain must consider macroscopic properties such as trend, drift, etc., of the original data structure together with the kriging system because the optimality condition in equation (5.2) is no longer valid outside the (convex) domain, as long as the approximation (5.1) is a linear combination of sample values. This optimality constraint described above comes

from linear optimization theory and, hence, is independent of the choice of variogram; $\mathcal{R}_p$-estimator, semivariogram γ in equation (3.7), or any other estimators.

With regard to the scaling factor, recall that, from Section 3 (see, equation (3.4)), for each $p > 0$,

$$\mathcal{R}_p(ch) = c\mathcal{R}_p(h) \tag{5.3}$$

for any $c > 0$ and lag h. Thus, the scaling factor $c > 0$ of the original data set is preserved in the correlation estimation step. Moreover, for any constant $c > 0$, the following two kriging systems:

$$\begin{aligned} &\sum_{j=1}^{m} \omega_j (c\mathcal{R}_p(h_{ij})) + \lambda = c\mathcal{R}_p(h_{io}), \quad 1 \le i \le m, \\ &\sum_{i=1}^{m} \omega_i = 1, \end{aligned} \tag{5.4}$$

and

$$\begin{aligned} &\sum_{j=1}^{m} \omega_j \mathcal{R}_p(h_{ij}) + \lambda/c = \mathcal{R}_p(h_{io}), \quad 1 \le i \le m, \\ &\sum_{i=1}^{m} \omega_i = 1 \end{aligned} \tag{5.5}$$

are equivalent. That is, if the system (5.4) has a unique solution ω_i, $i = 1, 2, \cdots, m$, and λ, then the solution of the other system (5.5) is ω_i, $i = 1, 2, \cdots, m$, and λ/c, and conversely. In order words, the weights ω_i, $i = 1, \cdots, m$, determined by kriging system (5.4) or (5.5) are independent of the scaling factor $c > 0$. The only change is on the Lagrange multiplier λ which is needed for estimating the kriging variance. This property derived from equations (5.3)-(5.5) shows that the estimation procedure is consistent in the respect that preserves the scaling factor.

The kriging system (5.2) can be written in the following matrix form:

$$\mathbf{A}\mathbf{w} = \mathbf{b}, \tag{5.6}$$

where

$$\mathbf{A} = \begin{bmatrix} \mathcal{R}_p(h_{11}) & \mathcal{R}_p(h_{12}) & \cdots & \mathcal{R}_p(h_{1m}) & 1 \\ \mathcal{R}_p(h_{21}) & \mathcal{R}_p(h_{22}) & \cdots & \mathcal{R}_p(h_{2m}) & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathcal{R}_p(h_{m1}) & \mathcal{R}_p(h_{m2}) & \cdots & \mathcal{R}_p(h_{mm}) & 1 \\ 1 & 1 & \cdots & 1 & 0 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} \mathcal{R}_p(h_{1o}) \\ \mathcal{R}_p(h_{2o}) \\ \vdots \\ \mathcal{R}_p(h_{mo}) \\ 1 \end{bmatrix}$$

and $\mathbf{w} = [\omega_1 \quad \omega_2 \quad \ldots \quad \omega_m \quad \lambda]^T$.

In order for the system (5.6) to have a unique solution $\mathbf{w}$, the matrix $\mathbf{A}$ must be regular, i.e., invertible. Data scaling is needed, especially, when a given data set ranges too widely. Symmetry may cause singularity of the kriging matrix $\mathbf{A}$. If two samples, say, $Z(\mathbf{x}_i)$ and $Z(\mathbf{x}_j)$, that are selected for kriging happen to be located at exactly symmetric positions, the corresponding two (i-th and j-th) rows in the kriging matrix $\mathbf{A}$ become identical, and hence the matrix $\mathbf{A}$ becomes singular. Taking the average of two samples and counting it as a single sample with double weights, $2\omega_i$, can prevent this singularity of the kriging matrix. In this case, the size of the matrix $\mathbf{A}$ becomes $m \times m$ instead of $(m+1) \times (m+1)$.

We now discuss the reduction of the data set range by the log transformation which is commonly done instead of using a scaling factor. Assume that all elements in the data set $\{ Z(\mathbf{x}_i) : i = 1, 2, \cdots, m\}$ are positive so that the log transformation is well-defined. The log transformation is a nonlinear transformation which causes the transformed data set $\{ \log(Z(\mathbf{x}_i)) : i = 1, 2, \cdots, m\}$ to loose any given linear structure that was in the untransformed data. Moreover, if a kriging method that is based on linear approximation, $\sum_{i=1}^{m} \omega_i \log(Z(\mathbf{x}_i))$, is used for finding the "optimal weights" $\omega_i \geq 0$, these ω_is are nothing more than power indices, i.e.,

$$\sum_{i=1} \omega_i \log(Z(\mathbf{x}_i)) = \sum_{i=1} \log(Z(\mathbf{x}_i)^{\omega_i}) = \log\left(\prod_{i=1}^{m} Z(\mathbf{x}_i)^{\omega_i} \right), \tag{5.7}$$

where the symbol $\prod$ presents a product of all components on the right hand side of $\prod$. Therefore, when log transformed data are used, obtained are the power indicies from the kriging algorithm rather than the weights on $Z(\mathbf{x}_i)$. In addition, if the log transformation is inverted by an exponential function (with or without including the variance at

each sampling position), a hidden assumption has been made: *The value $Z(\mathbf{x}_o)$ at any unmeasured location $\mathbf{x}_o$ can be approximated by the product of a certain power of each sample data $Z(\mathbf{x}_i)$*, which is obviously suspicious. This occurs due to the logical error of using two different functional actions, nonlinear and linear, independently during the estimation prodecure. The log transformation is nonlinear and the kriging procedure is based on linear theory. One possibile method for keeping consistency of functional actions is to use nonlinear weights such as exponential weights in the kriging algorithm. But, this nonlinear weight correction returns the transformed data structure back into the original data structure. Therefore, any advantages gained by taking the log transformation would disappear. It is important to maintain consistency throughout the entire estimation procedure. Also, the correlation between data must be obtained from the original data structure without any destruction of correlation structure. The estimation method in this paper maintains this structure and is consistent, and the $\mathcal{R}_p$-estimator accounts for outliers effects coming from the randomness or complexity of geological characteristics. In the estimation of groundwater contaminants, high concentration regions have more attention than low concentration regions. But, by compressing the high concentration data, major information will be disregarded and the kriging procedure puts more focus on low concentration data. Therefore, changing data distribution structure by nonlinear transformation to fit the statistical assumptions which are hardly satisfied by the groundwater contaminant field data needs a special caution.

6. Application

Site and experiment description

The MADE-2 site is located at Columbus Air Force Base in northeastern Mississippi (see Figure 1). The shallow unconfined aquifer which immediately underlies the site consists of an alluvial terrace deposit averaging approximately 11 m in thickness. The aquifer is composed of poorly- to well-sorted sandy gravel and gravelly sand with variable silt and clay content. Sediments are generally unconsolidated, and occur as irregular horizontal or nearly horizontal lenses and layers. Marine sediments belonging to the Eutaw Formation and consisting of clays, silts, and fine-grained sands form an aquitard beneath this alluvial aquifer.

An earlier macrodispersion experiment, referred to as MADE-1, in-

volving bromide and three fluorobenzoate tracers, was conducted between October 1986 and June 1988 at the same site (see [1, 4, 6, 22]). More recently, a large-scale natural gradient tracer experiment(MADE-2) was conducted to acquire detailed data on the transport of a conservative tracer (tritium) and four reactive organic compounds (benzene, naphthalene, ^{14}C labeled p-xylene and o-dichlorobenzene). The goal of this experiment was to develop through direct field observation a better understanding of the physical, chemical, and biological processes affecting transport of dissolved contaminants in aquifers. Also, this experiment was to supplement the existing information produced from the previous field experiment MADE-1. Hydrologic properties of the Columbus aquifer and a more detailed description of the MADE site were given in [6].

The MADE-2 experiment was started on June 26, 1990 with a two-day pulse injection of 9.7 m^3 of tracer solution into the saturated zone approximately at $(x, y, z) = (0, 0, 57.8)$. The y-axis is the assumed groundwater flow direction and the x-axis crosses y-axis. The z-coordinate indicates the elevation above mean sea level. During the injection period, the maximum increase in hydraulic head in the nearby injection point was 0.45 m. Tracer concentration distributions were subsequently monitored at one to three month intervals over a period of 15 months in three dimensions using an extensive network of saturated zone multilevel samplers.

The extensive tritium plume data (2267 samples) collected on May 21, 1991, 328 days after injection during the MADE-2 experiment were used in this paper. After extracting macroscopic plume transport portion from the field data, geostatistical techniques and the $\mathcal{R}_p$-estimator are applied to the residuals. The use of linear geostatistics requires that the regionalized variable be additive, i.e., that linear combinations of the values of the variable have the same meaning [14]. In groundwater contaminant applications, this means that the contaminant concentrations obtained from groundwater samples must be multiplied by the porosity of the aquifer at the point where the sample was taken [7]. However, the porosity of the MADE-2 site aquifer was not measured at every sampled location. Instead, the porosity was determined by gravimetric and volumetric analyses of 84 minimally disturbed soil cores. The mean and standard deviation of the porosity measurements were 0.31 and 0.08, respectively [6]. The constant value of 0.31 was assumed for the aquifer porosity. Since the $\mathcal{R}_p$-estimator preserves any data scaling factor, the

tritium data concentration (as they are) were used in this paper instead of using the multiplied data by 0.31.

Regionalization

Figures 1 and 2 show the location and the plan view (xy-coordinate) of the test site in meter scale. Figures 2-4 show the entire sampling network in the xy-, yz- and xz-coordinate plane, respectively. In Figures 2-4, dots ($\cdot$) indicate sample locations with concentration less than 5 pCi/mL, and circles with dots inside ($\odot$) represent samples whose concentration are greater than or equal to 5 pCi/mL and less than 10 pCi/mL. The solid circles ($\bullet$) illustrate the samples with concentration greater than or equal to 10 pCi/mL. Measured activities from the background appeared to be 1-3 pCi/mL. Samples greater than 4 pCi/mL are regarded as the injected triated water plume. The approximate lateral boundary of the surface expression of an old geological stream channel (meander) is shown by dotted lines in Figures 1-2. This meander runs from the southwest to the northeast direction with approximately 60^o slope along the x-axis and the apparent center of the channel passes about the point $(x, y) = (0, 175)$. The groundwater flow follows the y-axis from the injection point (approximately $(x, y, z) = (0, 0, 57.8)$) to approximately $y = 125$ m and then appears to follow the meander channel. Thus, in the meander area, the coordinate system was corrected to follow the flow direction in the analysis. Figure 3 shows the presence of two horizontal aquifer layers. The upper aquifer above the approximate depth of $z = 58.25$ m in the injection area is going slightly downward until $y = 80$ m, and then more steeply to $y = 175$ m. Thus, the injection occurred just below the aquifer layer boundary. The center of plume has moved from the injection point approximately to $(x, y, z) = (-1.6, 3.6, 58.8)$ for the period of 328 days. Based on Figures 2-4 and the other site information[5], the test site was divided into five continuously overlapped regions: Region 1 (near field, -10 m $< y <$ 25 m), Region 2 (17 m $< y <$ 41 m), Region 3 (middle field, 25 m $< y <$ 125 m), Region 4 (100 m $< y <$ 150 m), and Region 5 (far field, meander area, 125 m $< y <$ 300 m). The summary of regionalization is shown in Table 1.

Estimation of macroscopic plume behavior

The primary deterministic processes in each region were estimated using the approximating function defined by equation (2.4). For this analysis $m = 2$ was chosen in equation (2.4) to demonstrate the feasibility of our estimation method. Basic methodologies and principles

for determining the initial parameters are similar for each region. However, Region 5 (meander area) required special attention in our study. For each coordinate direction, the first basis function captures the main profile of the concentration distribution such as height, width and the symmetric axis of the exponential function, whereas the second basis function in combination with the first function approximates the observed concentration distribution. For a_1^x, a_1^y and a_1^z, initial parameters were chosen as the expected center of the plume in each direction. The initial guesses for the parameters b_1^x, b_1^y and b_1^z were selected to adjust the width of the concentration distribution for given initial parameter sets $a_1 = (a_1^x, a_1^y, a_1^z)$. The initial data for the second class of parameters $a_2 = (a_2^x, a_2^y, a_2^z)$ and $b_2 = (b_2^x, b_2^y, b_2^z)$ were chosen such that the combination of two functions capture the field data distribution. The parameter set $c = (c_1, c_2)$ is related to the magnitude of concentration. The Levenberg-Marquart method[21] was used for estimating parameters in each region.

(1) Region 1 (near field):

This region contains the approximate injection point (0,0,57.8). The parameter estimation for Region 1 is more difficult than that for the other regions because the concentration in this region was controlled by the apparent boundary between the two layers, the injection period, the local mounding, and the radial flow from the injection well. The estimated parameters for each region are given in Table 2. Table 2 also shows the differences in the mean and variance values of the original field data and the residuals. In Table 2, mean(residuals) indicates the mean value of the residuals obtained from the field data by subtracting the estimated values in each region, and σ^2(residuals) is the variance of the residuals.

Figure 5 shows camparison between the original concentration distributions (+) and the estimated concentration (o) of the global plume behavior at the same sampled locations. The total number of samples used in this region was 775. Figure 6 shows the residuals in the x-direction. The residuals still have certain distribution patterns which probably come from characteristics of aquifer material and prediction limitations associated with a $m = 2$ approximation of plume macroscopic bahavior. It was also observed, from the distribution of residuals, that randomness of the residuals increased significantly compared with the original distribution.

(2) Region 2 - Region 4:

Region 2 connects Region 1 (near field) and Region 3 (middle field). The magnitude of concentration changes significantly in y-direction. The first basis function was chosen for fitting the trend from the near field, and the second was chosen for capturing the trend from the middle field. Region 4 connects the middle field and the meander area. The upper aquifer above the approximate depth of $z = 58$ m in the near field is going slightly downward until $y = 80$ m, and then more steeply to $y = 175$ m (Figure 3). The first basis function was chosen to capture this steep slope, and the second basis function was set to fit the main profile of the data distribution in the meander area.

(3) Region 5 (far field):

This region appears to be an old ox bow river bed that is buried (Figures 1-3). In this region, the transform of the coordinate system was needed to improve the accuracy of the approximation (equation (2.4)), and to account for the symmetry and the ratio along the groundwater flow path. Note that the y-axis was chosen as the groundwater flow path and the x-axis was the transverse direction with respect to the y-axis. Since the groundwater appears to flow along the river bed, the y-axis should be set in the river bed direction. The river bed direction, width and its depth can be found easily by changing the coordinate system and observing the data distribution along the transformed axes. It was observed that the angle between the river direction (north-east) with the x-axis was approximately 60^o and that the center of the river bed passed through the point $(x, y) = (0, 175)$ (Figures 1-3). The corrected coordinate system $(\hat{x}, \hat{y}, \hat{z})$ appears to be

$$\begin{aligned} \hat{x} &= x \cos\alpha + (y - y_o) \sin\alpha, \\ \hat{y} &= -x \sin\alpha + (y - y_o) \cos\alpha + y_o, \\ \hat{z} &= z \end{aligned} \tag{6.1}$$

with $\alpha = -(30/180)\pi$ and $y_o = 175$, where (x, y, z) is the original coordinate system. No coordinate transform was made in the z-direction since the water table over this region, 125 m $< y <$ 250 m, is approximately parallel to the sea level (see Figure 3). Recall that the data were obtained from water samples in the saturated zone. The transformed coordinate system $(\hat{x}, \hat{y}, \hat{z})$ was used for parameter estimation.

Variogram

To obtain experimental variograms which are needed in kriging procedure, each region was divided into several blocks depending on the distribution structure of the residuals. Entire five regions were divided into twelve blocks with adjacent blocks overlapped. Since methodologies to determine symmetry, ratio and correlation are similar for all blocks, only Block 3 in Region 1 was chosen for our discussion. The application strategies presented can be extended to any general field situation. Block 3 consists of area surrounded by -70 m $< x <$ 70 m, 5 m $< y <$ 15 m, and 53 m $< z <$ 64 m. This block is in Region 1 and located at just behind the center of plume (approximately, $y = 3.6$ m) along the groundwater flow direction. Bock 3 showed the most complicated data distribution structure among twelve blocks. The total number of samples in this block was 306.

Symmetry

Even though the primary symmetric portion was extracted from the field data, the residual deterministic components produced a symmetric distribution. Symmetric behaviors were found in the $x-$ and z-directions. No clear symmetric structure was found in the y-direction. The symmetry axis in the y-direction was located approximately at $y = 2.1$ m, outside (the left hand side of) Block 3 (see Table 3). Note, from Table 3, that there is a difference between the location of plume center ($y = 3.6$ m) and the symmetry axis ($y = 2.1$ m) in y-direction. The center of plume was estimated from original plume field data and the symmetry axis was measured from the residuals. An averaged variance was considered by applying by the formula $(1/n(I)) \sum_{i=1}^{n(I)} |Z_i - Z_j|$ to each interval, where the intervals were continuously overlapped and covered whole x-coordinate range, $n(I) = k(k-1)/2$ was the number of sample pairs in the interval I and k was the number of samples. Forty and fifteen intervals were used to obtain the averaged variances in the x- and the z-direction, respectively. Since the symmetry axis is independent of variance estimator used, other variance estimators such as $(1/n(I)) \sum_{i=1}^{n(I)} |Z_i - Z_j|^2$ or $((1/n(I)) \sum_{i=1}^{n(I)} |Z_i - Z_j|^p)^{1/p}$, etc., may be used to determine the symmetry axis. However, it is desirable to use the variance estimator which has a similar structure to the correlation estimator. The symmetric axes were, approximately, $x = -1.5$ m and $z = 58.5$ m. To cover the uncertainty of precise symmetric axes, the averaging method over small neighborhood of each lag was used to estimate correlation between the data points.

Ratio

To account for anisotropy effecting correlation, we considered a ratio between the three directions, x, y and z. The ratio is a realization of the advection and dispersion/diffusion processes in conjunction with the aquifer material. This ratio can be obtained by considering either the plume distribution or the residuals. Based on the residual distributions in the y- and z-direction, the ratio for Block 3 was determined. But, the methodology for determining the ratio can be generally applied to obtain the correlation between data distribution and the aquifer material properties. From the plume distribution in each region (from Region 1 to Region 5), the ratio shows consistancy throughout the entire test site.

We considered the distribution of residuals whose absolute values are greater than 20 ($\times 20$). Here, the symbol $\times 20$ represents the scaling factor of the data. Other numbers such as 10 ($\times 20$) or 15 ($\times 20$) may be used to estimate the ratio. But, this number should be selected so that it represents the main profile of the data distribution in the block considered. The residuals whose absolute values are greater than 20 ($\times 20$) were distributed from approximately $x = -5\,m$ to $x = 5\,m$. The corresponding y- and z-ranges were from $y = 5$ m to $y = 15$ m and from $z = -56.5$ m to $z = 61.5$ m, respectively. Since the x and z directions have symmetric distribution behaviors, the ratio is approximately $x : y : z = 10/2 : 10 : 5/2 = 2 : 4 : 1$. This ratio varied slightly in each region (from Region 1 to Region 5) and in each block (from Block 1 to Block 12). However, overall the simple ratio $x : y : z = 2 : 4 : 1$ was generally true ($\hat{x} : \hat{y} : \hat{z} = 2 : 4 : 1$ for Blocks 10-12) so this ratio was set for estimating the correlation between data points. Here, $(\hat{x}, \hat{y}, \hat{z})$ is the transformed coordinate system (equation (6.1)) to incorporate the flow direction in the meander area. From this ratio, any two points separated by approximately 2 m distance in x-direction, 4 m distance in y-direction, and 1 m distance in z-direction are considered to have the same correlation in the average sense.

Figure 7 shows the improvement of an experimental semivariogram by accounting for symmetry and ratio. For this comparison, the conventional semivariogram:

$$\gamma(h) = \frac{1}{2\,n(h)} \sum_{i=1}^{n(h)} |Z_i - Z_j|^2 \tag{6.2}$$

was used, where $n(h)$ is the number of pairs (Z_i, Z_j) separated by lag h. Semivariograms were estimated by taking the average value over the

interval $[h-1, h+1]$ for the first case, and the corresponding intervals for other cases. All three cases were plotted on the same scale. The first semivariogram (dotted line with circles) was obtained by equation (6.1) without any consideration of symmetry and ratio. It is interesting to see that the correlation between lag (h) and semivariogram value $\gamma(h)$ has an inverse relationship. Distribution symmetry and outlier effects produce the inverse relation which is typical whenever the spatial structure of groundwater contaminant data is estimated with conventional semivariograms. For a second semivariogram (dashdotted line with cross), the symmetry ($x = -1.5$ m and $z = 58.5$ m) were incorporated. The semivariogram for capturing the global trend is improved. The last semivariogram (solid line with point) incorporate both symmetry and ratio ($x : y : z = 2 : 4 : 1$). Compared to the second semivariogram using symmetry alone, the final semivariogram accounts for both symmetry and ratio and produced improved correlations, i.e., the range of variance between estimated values decreased.

The kriging procedure usually uses several (typically, 12-20) sample values to estimate the value at the unsampled location. Thus, what is really needed is a correlation in near the point or within half the range of the lag distance in the semivariogram. From Figure 7, even though a global semivariogram is obtained by considering symmetries and ratio, the semivariogram in the range of $0 \leq h \leq 15$ shows no distinguishable correlation. This behavior may be due to the influence of outliers. From equation (6.1), it can be seen that the influence of outliers on the semivariogram increases by the square ($|Z_i - Z_j|^2$) as the difference $|Z_i - Z_j|$ increases. Therefore, to reduce outlier's influence on correlation, a robust estimator is required.

Robustness

Figure 8 illustrates the effectiveness of $\mathcal{R}_p$-estimator. In all correlations, the symmetries and ratio were taken into consideration. For the first correlaton (+), the robust index $p = 1$ was chosen. The global pattern of correlation is similar to the semivariogram (solid line) in Figure 7. But, the variogram shows improvement near the origin. The robust indices p for the second (x) and the third (o) correlations in Figure 8 were $\frac{1}{2}$ and $\frac{1}{8}$, respectively. As p approaches 0, the correlation near the origin improves. The solid line in Figure 8 is a typical fitted mathematical model for the third variogram (o). The following exponential model was used to fit the correlation.

$$\gamma(\alpha, \beta, c0; h) = c0 + c\,(\beta(1 - \exp(-h/\alpha)). \tag{6.3}$$

The fitted parameters were $c = 20$, $\alpha = 1.1$, $\beta = 1.9$ and $c0 = 0.95$, where c is the scaling factor. The parameters were chosen so that the model (exponential) captured the variogram within the range $0 \leq h \leq 6$, because the variogram in this range will be used for the kriging procedure. Symmetries, ratio, and fitted model parameters for variograms for other blocks were summarized in Table 3.

In Table 3, d is the lag average and $(\hat{x}, \hat{y}, \hat{z})$ is the transformed coordinate system (equation (6.1)) for the meander area. Note that there is an apparent symmetry $\hat{y} = 183\,m$ in $\hat{y}$-direction for Blocks 10-12. The robust indices p for Blocks 1-5 in the range of -10 m $< y <$ 40 m are much smaller than those of other blocks in the middle or far field. Note that $p = 1/8$ for Block 3. Recall that Block 3 has the most complex data distribution structure among Blocks 1-12. Figure 9 shows fitted model variograms for Blocks 1-12. To compare each variogram, the nugget effect $c \times c0$ in Table 3 was subtracted from the estimated value $\mathcal{R}_p(h)$ for each block.

Kriging

Using the exponential variogram models (equation (6.3)) fitted to experimental variograms obtained by the $\mathcal{R}_p$-estimator, punctual kriging was used to estimate residuals at unmeasured locations. Anisotropies such as symmetry and ratio were considered. The symmetry may cause sigularity of the kriging matrix. If, for example, two samples selected for kriging happen to be located at exactly symmetric positions, only one sample was selected for kriging. The averaged value of those two samples was assigned for the value of this sample. In kriging procedure, the weight for this sample was doubled to take account for the eliminated sample due to symmetry. At each estimating point, 25-35 samples were used for kriging.

7. Conclusion

The spatial or temporal distribution of groundwater contaminant concentrations has different behavior in macroscale as well as microscale compared to that of other geostatistical subjects such as the elevation of the ground surface or changes in grade within an ore body. Unlike most other geostatical subjects described by stationary process, the trajectory of the contaminant plume follows complex dynamical processes which may be described approximately by a transport equation.

Application of geostatistical methods to the groundwater contaminant concentrations has a difficulty because they do not obey the stationary or semistationary assumption. In this paper, a practical, systematic, three-dimensional estimation method for *in situ* groundwater contaminant concentrations was introduced. After extracting the macroscopic plume behavior from field data, geostatistical estimation methods were applied to residuals. Methodologies and principles on how to choose the approximating function and the initial parameters for estimating the global plume behavior were explained. A robust $\mathcal{R}_p$-estimator was introduced. This estimator reduces outlier effects systematically and produces correlations between data points. These correlations are used for determining variogram models which needed for kriging a spatially distributed parameter at unsampled location.

In our application, after extracting macroscopic plume behavior from the tritium data, the randomness of the residuals was increased significantly compared with that of the original data set. The $\mathcal{R}_p$-estimator reduced outlier effects on obtaining correlation between the residuals, and produced good experimental variograms. The estimated tritium concentrations were calculated by the following formula: Total estimated value = Estimated value of macroscopic transport portion + Residual value estimated by kriging. Kriging was performed throughout all twelve blocks (Table 3) in 2 m, 4 m, and 1 m intervals in the x-, y-, and z-directions. The total data set obtained ranges from -70 m to 70 m, from -10 m to 300 m, and 53 m to 64 m in the x-, y-, and z-coordinates. Figures 10 and 11 are contour plots of the tritiated water concentration distributions for the xy-cut plane at $z = 59\,m$ and for the yz-cut plane at $x = 0\,m$, respectively. Contours were chosen for concentrations $C = 5$, 10, 20, 50, 500, 2000 pCi/mL. Concentrations greater than 4 pCi/mL were regarded as the injected tritiated water plume. Compared with the *in situ* concentration distribution data (Figures 2-3), Figures 10-11 represent total estimated values on uniform grids. Figure 12 shows a perspective view of the isosurface with concentration $C = 5$ pCi/mL of the tritiated water plume.

TABLE 1. Regionalization

	Reg.1(y)	Reg.2(y)	Reg.3(y)	Reg.4(y)	Reg.5(y)
range(m) $-70 < x < 70$ $53 < z < 64$	$[-10, 25]$	$[17, 41]$	$[25, 125]$	$[100, 150]$	$[125, 300]$

TABLE 2. Parameter estimation

	Region 1	Region 2	Region 3	Region 4	Region 5
num. of samples	775	291	729	233	763
mean(field data)	136.90	14.59	8.51	4.80	5.36
σ^2(field data)	183854	2164.1	326.16	39.19	31.91
estimated					
a_1^x	-1.5777	-1.0388	. -6.7366	0.5244	$a_1^{\hat{x}}$=3.0551
a_1^y	3.5673	-46.3318	26.4202	115.3410	$a_1^{\hat{y}}$=174.1736
a_1^z	58.7535	59.2096	60.2719	58.9507	$a_1^{\hat{z}}$=57.2405
b_1^x	3.1662	6.8355	2.3997	6.6973	$b_1^{\hat{x}}$=21.0780
b_1^y	3.1932	68.4465	20.1242	18.9481	$b_1^{\hat{y}}$=36.7760
b_1^z	0.7509	1.6122	0.7954	1.9855	$b_1^{\hat{z}}$=3.7705
c_1	2564.51	212.0565	132.5352	39.7180	c_1=19/5050
a_2^x	-0.7880	0.8700	1.0305	-12.9028	$a_2^{\hat{x}}$=4.7229
a_2^y	9.6721	-44.5587	65.7696	212.6386	$a_2^{\hat{y}}$=281.6550
a_2^z	58.7961	58.5933	59.1561	57.8191	$a_2^{\hat{z}}$=58.6649
b_2^x	3.0609	10.8718	7.9127	52.2654	$b_2^{\hat{x}}$=72.5429
b_2^y	7.1115	48.4521	83.3313	88.9769	$b_2^{\hat{y}}$=214.5971
b_2^z	1.9222	4.4246	2.4166	3.7290	$b_2^{\hat{z}}$=6.4690
c_2	1690.38	97.0468	44.5492	15.1114	c_2=4.9264
mean(residuals)	11.17	0.7252	0.83	0.12	0.01
σ^2(residuals)	63649	1750.5	174.66	10.98	16.67

TABLE 3. Symmetry, ratio$(x : y : z)$, average(d) and variogram$(\mathcal{R}_p(h) = c_0 + c(\beta(1 - e^{-h/\alpha}))$

block (sampels)	range $-70 < x < 70$ $53 < z < 64$	sym. axis	ratio	p, d	variogram
Block 1 (224)	$-10 < y < 3$	$x = 0.0$ $y = 2.1$	2:4:1	$p = 1/2$ $d = 0.2$	$c_0 = 0.5$, $c = 20$ $\alpha = 2.2$, $\beta = 2.5$
Block 2 (313)	$0 < y < 10$	$x = -0.9$ $y = 2.1$ $z = 58.6$	2:4:1	$p = 1/6$ $d = 0.2$	$c_0 = 2.5$, $c = 20$ $\alpha = 1.1$, $\beta = 2.2$
Block 3 (306)	$5 < y < 15$	$x = -1.5$ $z = 59.0$	2:4:1	$p = 1/8$ $d = 0.2$	$c_0 = 0.95$, $c = 20$ $\alpha = 1.1$, $\beta = 1.9$
Block 4 (224)	$10 < y < 25$	$x = -0.6$ $z = 59.0$	2:4:1	$p = 1/4$ $d = 0.2$	$c_0 = 0.45$, $c = 20$ $\alpha = 1.1$, $\beta = 1.2$
Block 5 (224)	$17 < y < 41$	$x = 0.0$ $z = 59.0$	2:4:1	$p = 1/4$ $d = 0.2$	$c_0 = 3.8$, $c = 1$ $\alpha = 2.2$, $\beta = 10.5$
Block 6 (293)	$25 < y < 55$	$x = 0.0$ $z = 59.5$	2:4:1	$p = 1/2$ $d = 0.2$	$c_0 = 4.5$, $c = 1$ $\alpha = 15.0$, $\beta = 16.0$
Block 7 (319)	$50 < y < 90$	$x = 1.2$ $z = 59.4$	2:4:1	$p = 1/4$ $d = 0.2$	$c_0 = 3.0$, $c = 1$ $\alpha = 5.0$, $\beta = 4.0$
Block 8 (214)	$80 < y < 125$	$x = 2.0$ $z = 59.0$	2:4:1	$p = 1/2$ $d = 0.2$	$c_0 = 1.75$, $c = 1$ $\alpha = 10.0$, $\beta = 6.0$
Block 9 (233)	$100 < y < 150$	$x = 2.0$ $z = 58.0$	2:4:1	$p = 1/2$ $d = 0.4$	$c_0 = 1.2$, $c = 1$ $\alpha = 5.2$, $\beta = 2.0$
Block 10 (351)	$125 < y < 200$	$\hat{x} = -1$ $\hat{y} = 183$ $\hat{z} = 58.5$	$(\hat{x} : \hat{y} : \hat{z})$ 2:4:1	$p = 1/2$ $d = 0.2$	$c_0 = 2.5$, $c = 1$ $\alpha = 3.0$, $\beta = 2.6$
Block 11 (263)	$175 < y < 225$	$\hat{x} = -1$ $\hat{y} = 183$	$(\hat{x} : \hat{y} : \hat{z})$ 2:4:1	$p = 1/2$ $d = 0.2$	$c_0 = 2.0$, $c = 1$ $\alpha = 3.5$, $\beta = 2.1$
Block 12 (400)	$200 < y < 300$	$\hat{x} = -1$ $\hat{y} = 183$	$(\hat{x} : \hat{y} : \hat{z})$ 2:4:1	$p = 1/2$ $d = 0.4$	$c_0 = 1.2$, $c = 1$ $\alpha = 5.0$, $\beta = 1.2$

List of figures

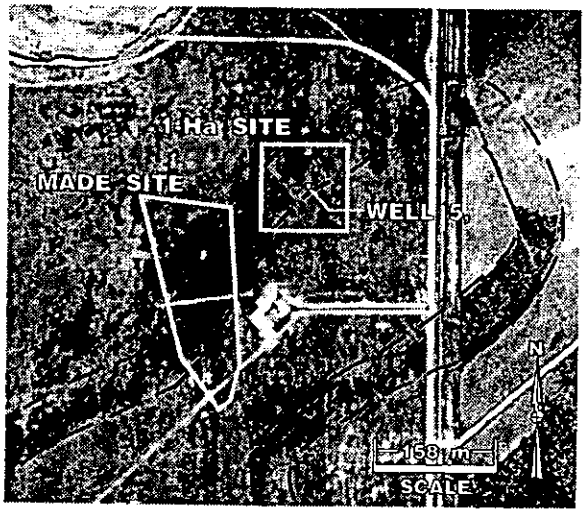

FIGURE 1

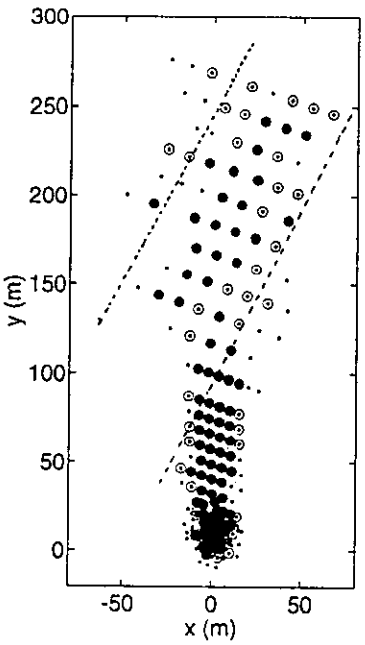

FIGURE 2

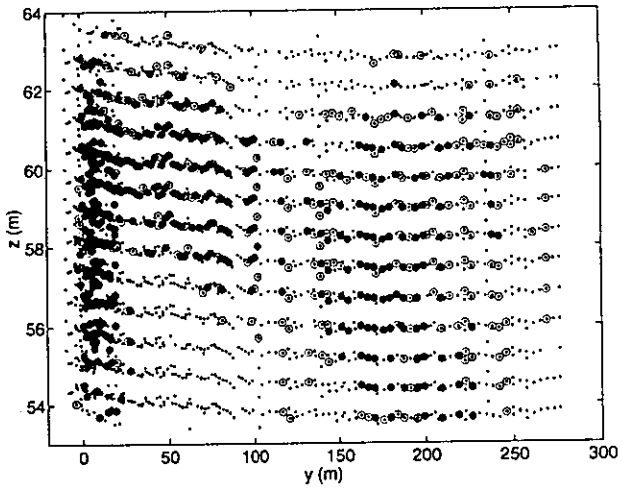

FIGURE 3

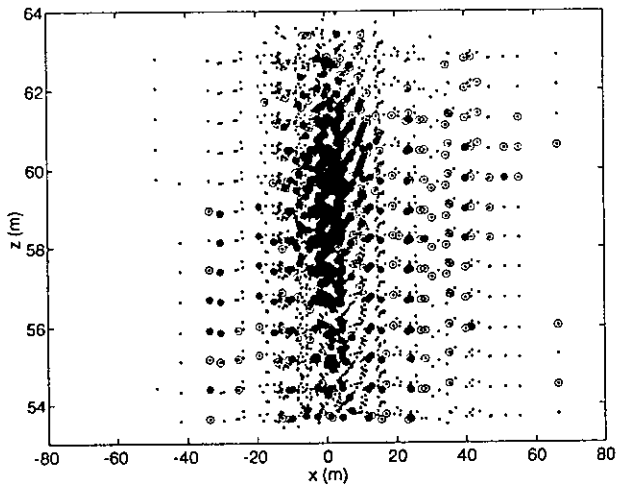

FIGURE 4

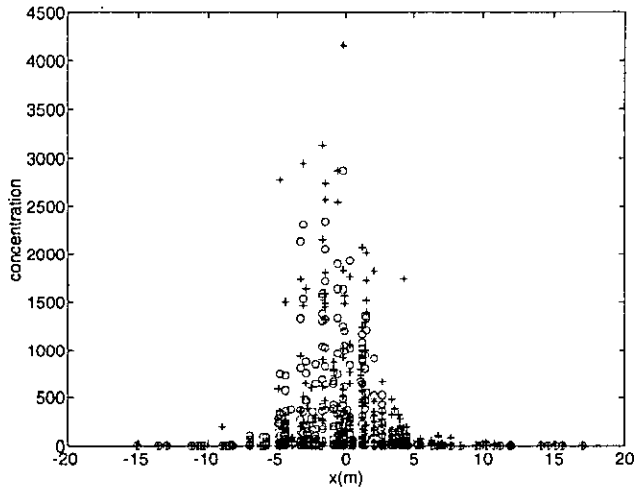

FIGURE 5

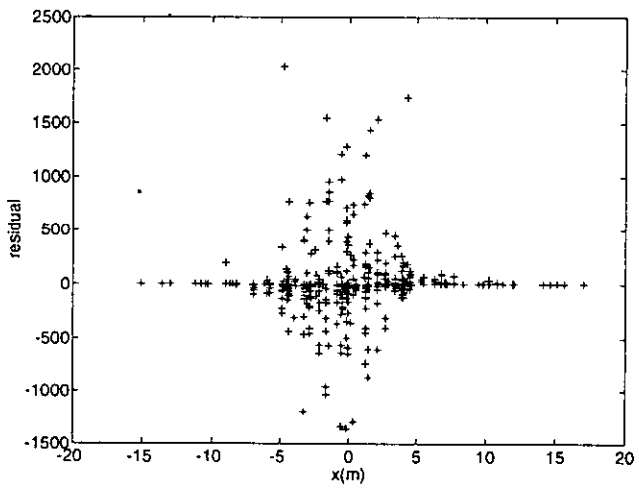

FIGURE 6

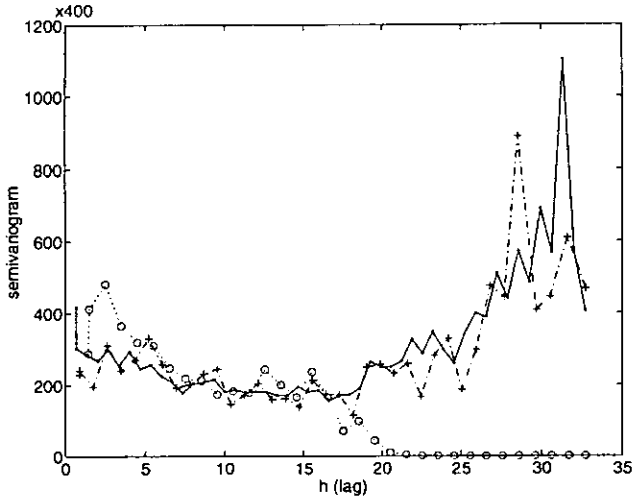

FIGURE 7

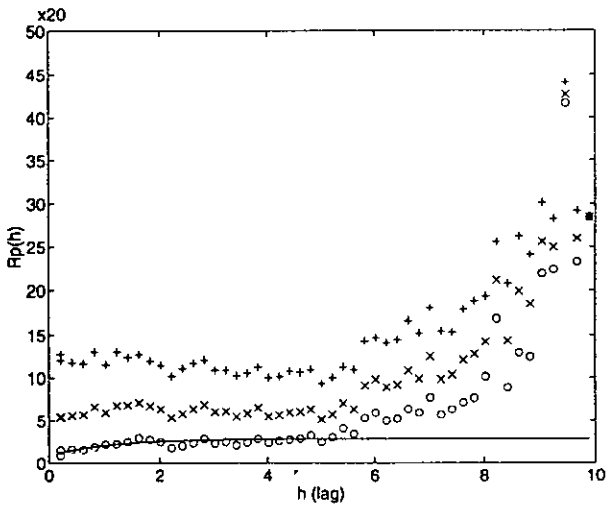

FIGURE 8

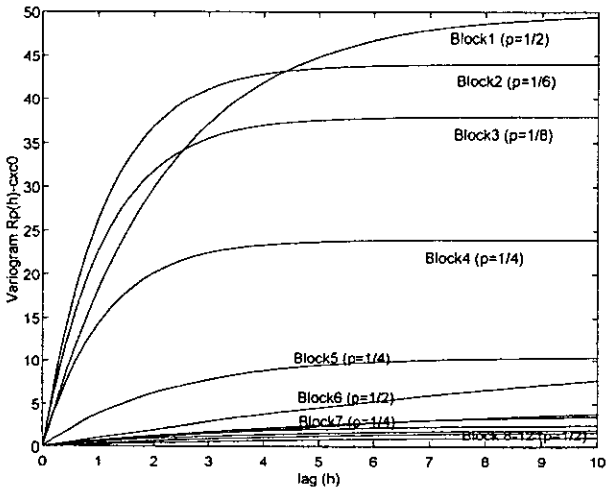

FIGURE 9

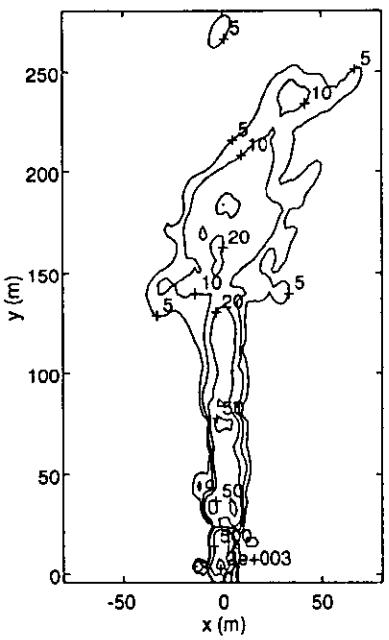

FIGURE 10

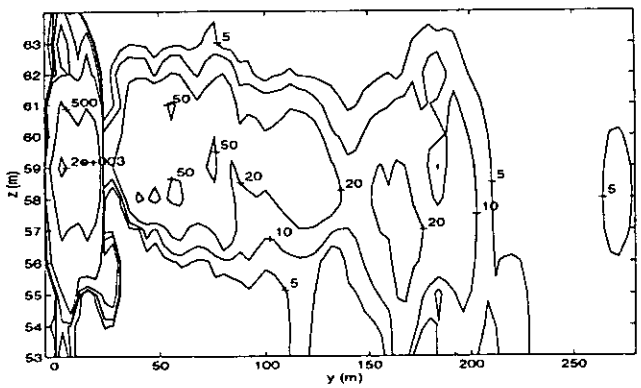

FIGURE 11

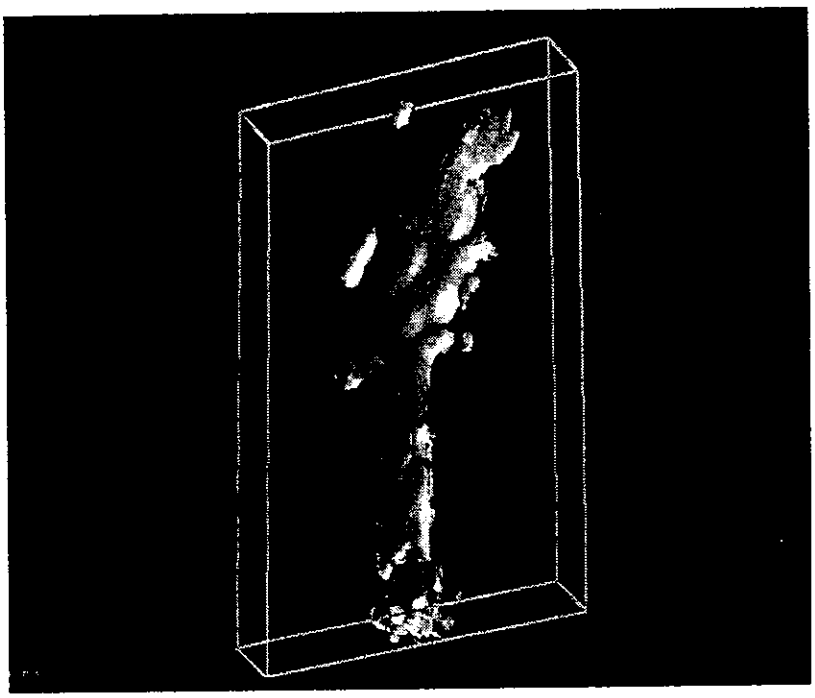

FIGURE 12

References

[1] E. E. Adams and L. W. Gelhar, *Field study of dispersion in a heterogeneous aquifer, 2. Spatial moments analysis*, Water Resour. Res. **28** (1992), no. 12, 3293–3307.

[2] ASCE Task Committee on Geostatistical Techniques in Geohydrology of the Ground Water Hydrology, Committee of the ASCE Hydraulics Division, *Review of geostatistics in geohydrology. I: Basic concepts; II. Applications*, J. Hydraulic Engineering **116** (1990), no. 5, 633–658.

[3] H. T. Banks and K. Kunish, *Estimation Techniques for Distributed Parameter Systems*, Birkhäuser, Boston, 1989.

[4] J. M. Boggs and E. E. Adams, *Field study of dispersion in a heterogeneous aquifer, 4. Investigation of adsorption and sampling bias*, Water Resour. Res. **28** (1992), no. 12, 3325–3336.

[5] J. M. Boggs, L. M. Beard, W. R. Waldrop, T. B. Stauffer, W. G. MacIntyre, and C. P. Antworth, *Transport of tritium and four organic compounds during a natural gradient experiment (MADE-2), EPRI Report TR-101998,*, Electric Power Research Institute, Palo Alto, CA 94304, 1993.

[6] J. M. Boggs, S. C. Young, L. M. Beard, L. W. Gelhar, K. R. Rehfeldt, and E. E. Adams, *Field study of dispersion in a heterogeneous aquifer, 1. Overview and site description*, Water Resour. Res. **28** (1992), no. 12, 3281–3291.

[7] R. M. Cooper and J. D. Istok, *Geostatistics applied to groundwater contamination. I: Methodology; II: Application; III: Global estimates*, J. Environmental Engineering **114** (1988), no. 2, 270–286; 287–299; **114** (1988), no. 4, 915–928.

[8] N. Cressie and D. M. Hawkins, *Robust estimation of the variogram: I*, Math. Geol. **12** (1980), no. 2, 115–125.

[9] J. E. Dennis, Jr. and R. B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, Printice-Hall, Englewood Cliffs, New Jersey, 1983.

[10] P. A. Dowd, *The variogram and kriging: Robust and resistant estimators, Geostatistics for Natural Resourses Characterization, Part I*, NATO ASI Ser. Ser. I Glob. Environ. Change **122** (1984), 91–107.

[11] R. A. Freeze and J. A. Cherry, *Groundwater*, Prentice-Hall, Englewood Cliffs, NJ., 1979.

[12] D. L. Freyberg, *A natural gradient experiment on solute transport in a sand aquifer, 2. Saptial moments and the advection and dispersion of nonreactive tracers*, Water Resour. Res. **22** (1986), no. 13, 2031–2046.

[13] S. P. Garabedian, D. R. Leblanc, L. W. Gelhar, and M. A. Celia, *Large-scale natural gradient tracer test in sand and gravel, Cape Cod, Massachusetts, 2. Analysis of spatial moments for a nonreactive tracer*, Water Resour. Res. **27** (1991), no. 5, 911–924.

[14] A. G. Journel and J. C. Huijbregts, *Mining Geostatistics, 5th ed.*, Academic Press, San Diego, 1991.

[15] D. R. Leblanc, S. P. Garabedian, K. M. Hess, L. W. Gelhar, R. D. Quardri, K. G. Stollenwerk, and W. W. Wood, *Large-scale natural gradient tracer test in sand and gravel, Cape Cod, Massachussetts, 1. Experimental design and observed tracer movement*, Water Resour. Res. **27** (1991), no. 5, 895–910.

[16] W. G. MacIntyre, M. Boggs, C. P. Antworth, and T. B. Stauffer, *Degradation kinetics of aromatic organic solutes introduced into a heterogeneous aquifer*, Water Resour. Res. **29** (1993), no. 12, 4045–4051.

[17] D. M. Mackay, 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** (1986), no. 13, 2017–2029.

[18] D. W. Marquardt, *An algorithm for least squares estimation of nonlinear parameters*, J. SIAM **11** (1963), 431–441.

[19] S. F. Moore and D. B. Mclaughlin, *Mapping contaminated soil plumes by kriging*, Proc. EPA National Conference on Management of Uncontrolled Hazardous Wastes, Washington, D. C. 1980, pp. 66–70.

[19] J. M. Ortega and W. C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, 1970.

[20] A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Springer-Verlag, New York, 1983.

[21] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in Fortran: The Art of Scientific Computing, 2nd ed.*, Cambridge University Press, 1992.

[22] K. R. Rehfeldt, J. M. Boggs, and L. W. Gelhar, *Field study of dispersion in a heterogeneous aquifer, 3: Geostatistical analysis of hydraulic conductivity*, Water Resour. Res. **28** (1992), no. 12, 3309–3324.

[23] M. J. L. Robin, E. A. Sudicky, R. W. Gillham, and R. G. Kachanoski, *Spatial variability of strontium distribution coefficients and their correlation with hydraulic conductivity in the Canadian Forces Base Borden aquifer*, Water Resour. Res. **27** (1991), no. 10, 2619–2632.

[24] W. Rudin, *Functional Analysis*, McGraw-Hill, New York, 1973.

[25] E. A. Sudicky, *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** (1986), no. 13, 2069–2081.

[26] A. D. Woodbury and E. A. Sudicky, *Inversion of the Borden tracer experiment data: Investigation of stochastic moment models*, Water Resour. Res. **28** (1992), no. 9, 2387–2398.

[27] J. H. Zirschky and D. J. Harris, *Geostatistical analysis of hazardous waste data*, J. Envir. Engrg., ASCE **112** (1986), no. 4, 770–785.

[28] J. H. Zirschky, G. P. Keary, R. O. Gilbert, and E. J. Middlebrooks, *Spatial estimation of hazardous waste site data*, J. Envir. Engrg., ASCE **111** (1985), no. 6, 777–789.

Richard Ewing
Institute for Scientific Computation
Texas A&M University
College Station, TX 77843-3404, USA
E-mail: ewing@isc.tamu.edu

Sungkwon Kang
Department of Mathematics
Chosun University
Kwangju 501-759, Korea
E-mail: sgkang@mail.chosun.ac.kr

Jeongook Kim
Department of Mathematics
Chonnam National University
Kwangju 500-757, Korea
E-mail: jkim@chonnam.chonnam.ac.kr

Thomas B. Stauffer
AFRL/MLQR
Tyndall AFB, FL 32403, USA
E-mail: tom.stauffer@tyndall.af.mil