



Large-Time Spatial Covariance of Concentration of Conservative Solute and Application to the Cape Cod Tracer Test

MARILENA PANNONE* and PETER K. KITANIDIS

Department of Civil and Environmental Engineering, Stanford University, Stanford, CA

(Received: 22 January 1999; in final form: 2 October 1999)

Abstract. Most studies on conservative transport in stationary velocity fields have focused on the description of the concentration mean. In this work, we use a Lagrangian methodology to develop an analytical expression for the spatial covariance of the concentration, based on the central limit theorem and applicable to large times after injection. We use this expression to analyze the conservative tracer test data from the natural gradient experiment conducted at Cape Cod in 1985–1988, in which bromide was quickly injected into the aquifer and the concentration was measured at many locations at certain points in time. The parameters that determine the concentration mean had been estimated in previous studies. Here, the two-particle covariance matrix, needed to describe the concentration covariance function, is derived from the measurements through a maximum likelihood method. Then, the data are interpolated on a grid using simple kriging, and contour maps of the concentration estimates are plotted. The results of cross-validation indicate that the model is consistent with the field measurements and the kriging estimates appear realistic.

Key words: concentration covariance model, two-particle moments, maximum likelihood estimation, field measurements, kriging interpolation.

1. Introduction and Background

Solute transport in natural porous formations is usually described at three distinct spatial levels of resolution: pore, laboratory, and field (Bear, 1972; Dagan, 1989). The upscaling from the pore to the laboratory and from the laboratory to the field leads to conservation equations that contain constitutive parameters globally representing the unresolved variability. Consider the advection–dispersion equation describing the transport of an ideal tracer

$$\frac{\partial c}{\partial t} + \mathbf{u} \cdot \nabla c = \nabla \cdot (\mathbf{D} \nabla c). \quad (1)$$

At the laboratory scale, \mathbf{u} is the local seepage velocity vector, the macroscopic contribution of the highly irregular velocity field within the pore network; \mathbf{D} is the local dispersion tensor, which represents the superposition of molecular diffusion and hydrodynamic dispersion due to velocity fluctuations that were averaged out; finally, c is the concentration defined over a volume that contains many pores.

* Now at Department of Environmental Physics and Engineering, University of Basilicata, Italy.

The seepage velocity fluctuates with the hydraulic conductivity, which is usually extremely variable in geologic formations. Since it cannot be measured everywhere, it is impossible to predict exactly the small-scale fluctuations of the solute concentration over a whole formation. The most common approach is to attempt to resolve the slowly varying components of the distribution, averaging over the microscale heterogeneities. For instance, the stochastic approach focuses on the prediction of the concentration ensemble mean as function of the statistics of the hydraulic conductivity. Under certain conditions, the ensemble mean concentration $\langle c \rangle$ satisfies an advection–dispersion equation

$$\frac{\partial \langle c \rangle}{\partial t} + \mathbf{U} \cdot \nabla \langle c \rangle = \nabla \cdot (\mathbf{D}_m \nabla \langle c \rangle), \quad (2)$$

where \mathbf{U} is the ensemble mean velocity and \mathbf{D}_m is the macrodispersion tensor. In statistically homogeneous velocity fields and for sufficiently large travel times, both \mathbf{U} and \mathbf{D}_m are constant (Dagan, 1982, 1984, 1989; Gelhar and Axness, 1983; Neuman *et al.*, 1987; Rubin, 1990).

However, the ensemble mean concentration $\langle c \rangle$ is just the best estimate of c , given the statistics of the small-scale variability. A useful measure of the expected deviation is represented by the concentration variance σ_c^2 , which is the mean square difference between the actual and the mean distributions (Dagan, 1982, 1984, 1989; Graham and McLaughlin, 1989; Vomvoris and Gelhar, 1990; Rubin, 1991; Kapoor and Gelhar, 1994a, 1994b; Kapoor and Kitanidis, 1996, 1998; Dagan and Fiori, 1997; Kabala, 1997; Pannone and Kitanidis, 1999).

Concentration variance and dilution kinetics have been shown to depend on the ratio between two-particle and one-particle second-order moments (Pannone and Kitanidis, 1999), defined as

$$W_{ij} = \langle X'_i Y'_j \rangle = \langle X_i Y_j \rangle - \langle X_i \rangle \langle Y_j \rangle \quad (3)$$

and

$$P_{ij} = \langle X'_i X'_j \rangle = \langle X_i X_j \rangle - \langle X_i \rangle \langle X_j \rangle, \quad (4)$$

where vectors \mathbf{X} and \mathbf{Y} indicate the coordinates of two particles that started at the same location; the angle brackets indicate ensemble averaging; and $X'_i = X_i - \langle X_i \rangle$. Note that two particles that start from the same position follow different trajectories even in a single realization, because of local dispersion that makes them sample different streamlines. In macroscopically Fickian transport processes, and under asymptotic conditions, P_{ij} increases linearly with time. However, the rate of increase of W_{ij} gradually diminishes (e.g., Pannone and Kitanidis, 1999). The smaller W_{ij} compared to P_{ij} , the less correlated the locations of the two particles. Small correlation indicates a high degree of dilution, as particles started from the same location occupy different locations in a bigger volume. More complete dilution also means smaller concentration variance (Kapoor and Kitanidis, 1996, 1998), as the concentration in any realization becomes smoother and closer to its ensemble mean.

The two-particle moments have already been used in the study of turbulence (Richardson, 1926; Batchelor, 1952; Fisher *et al.*, 1979) and have also been analyzed in the context of groundwater flow and transport in heterogeneous formations through a Eulerian approach (Rajaram and Gelhar, 1993) as well as a linearized Lagrangian approach (Dagan, 1989; Fiori and Dagan, 1999). The theoretical analysis of the two-particle covariance under general conditions is complicated, since the nonlinear interaction between local- and macrodispersive mechanisms must be considered. Nonlinear effects due to the deviation of the particle trajectory from its mean have been analyzed by Neuman and Zhang (1990); Zhang and Neuman (1990); and Zhang (1995), in terms of single-particle second moments.

Pannone and Kitanidis (1999), performed a qualitative analysis suggesting that the rate of increase of \mathbf{W} should diminish with increasing time; additionally, they presented exact results for the simple case of stratified confined formations and Fickian regime with constant macrodispersion coefficients, indicating that the two-particle covariance tends to a constant value. The analytical derivation was possible by considering the nonstationarity of the two-particle Lagrangian velocity.

The authors derived an expression for the large-time concentration variance in a Lagrangian framework, using normality as a consequence of the central value theorem to provide closure. In this work, we extend this approach to develop expressions for the large-time concentration covariance. The problem of concentration covariance has been previously studied (Graham and McLaughlin, 1989; Neuman, 1993; Zhang, 1995; Kapoor and Kitanidis, 1997). Kapoor and Kitanidis (1997), formulated the problem of concentration covariance deriving its transport equation through a leading-order approach. Their formulation included advection with the mean velocity, macrodispersion due to velocity fluctuations, and a decay term due to local dispersion. This was an extension of previous work where, however, some of the above terms were neglected.

We apply the proposed Lagrangian theory to the data from the Cape Cod tracer test (LeBlanc *et al.*, 1991; Garabedian *et al.*, 1991). The data were previously analyzed for the parameters describing the mean concentration (Garabedian *et al.*, 1991; Thierrin and Kitanidis, 1994). We develop and apply a methodology to determine the remaining parameters needed to quantify concentration variance-covariance and dilution. Once these additional parameters are determined, we perform kriging to obtain an estimate of the actual plume concentration conditional on concentration observations.

2. Large-time Concentration Covariance

The spatial covariance between the fluctuations of concentration at two different locations \mathbf{x} and \mathbf{y} is

$$R_c(\mathbf{x}, \mathbf{y}, t) = \langle c'(\mathbf{x}, t) c'(\mathbf{y}, t) \rangle = \langle c(\mathbf{x}, t) c(\mathbf{y}, t) \rangle - \langle c(\mathbf{x}, t) \rangle \langle c(\mathbf{y}, t) \rangle, \quad (5)$$

where $c(\mathbf{x}, t)$ indicates the concentration at location \mathbf{x} and time t .

Without loss of generality, the concentration function may be written as

$$c(\mathbf{x}, t) = \int_{V_0} C_0(\mathbf{a}) \delta(\mathbf{x} - \mathbf{X}(\mathbf{a}, t)) d\mathbf{a}, \quad (6)$$

where $\mathbf{X}(\mathbf{a}, t)$ is the position at time t of a particle starting at time $t = 0$ from location \mathbf{a} ; $C_0(\mathbf{a})$ is the initial concentration; V_0 is the injection volume, and δ is the Dirac function. Then,

$$c(\mathbf{x}, t)c(\mathbf{y}, t) = \int_{V_0} \int_{V_0} C_0(\mathbf{a}) C_0(\mathbf{b}) \delta(\mathbf{x} - \mathbf{X}(\mathbf{a}, t)) \delta(\mathbf{y} - \mathbf{Y}(\mathbf{b}, t)) d\mathbf{a} d\mathbf{b}, \quad (7)$$

where $\mathbf{Y}(\mathbf{b}, t)$ represents the position at time t of a different particle that started at location \mathbf{b} .

The total positions \mathbf{X} and \mathbf{Y} can be subdivided in a mean and its fluctuation;

$$\mathbf{X}(\mathbf{a}, t) = \mathbf{a} + \mathbf{U}t + \mathbf{X}'(\mathbf{a}, t) \quad (8)$$

and

$$\mathbf{Y}(\mathbf{b}, t) = \mathbf{b} + \mathbf{U}t + \mathbf{Y}'(\mathbf{b}, t). \quad (9)$$

By virtue of the stationarity of the velocity field, the marginal statistics of the random variables \mathbf{X}' and \mathbf{Y}' do not depend on the initial location. However, their joint statistics depend on the separation between the two initial positions.

The statistical moments of the concentration function are obtained through averaging the actual quantities over the different possible realizations of the trajectory fluctuations. From (6),

$$\langle c(\mathbf{x}, t) \rangle = \int \int_{V_0} C_0(\mathbf{a}) \delta(\mathbf{x} - \mathbf{a} - \mathbf{U}t - \mathbf{X}') f_{\mathbf{X}'}(\mathbf{X}') d\mathbf{a} d\mathbf{X}'. \quad (10)$$

From (7),

$$\begin{aligned} \langle c(\mathbf{x}, t)c(\mathbf{y}, t) \rangle &= \iiint \int_{V_0} \int_{V_0} C_0(\mathbf{a}) C_0(\mathbf{b}) \delta(\mathbf{x} - \mathbf{a} - \mathbf{U}t - \mathbf{X}') \times \\ &\quad \times \delta(\mathbf{y} - \mathbf{b} - \mathbf{U}t - \mathbf{Y}') f_{\mathbf{X}'\mathbf{Y}'}(\mathbf{X}', \mathbf{Y}'; \mathbf{a} - \mathbf{b}) d\mathbf{a} d\mathbf{b} d\mathbf{X}' d\mathbf{Y}', \end{aligned} \quad (11)$$

where $f_{\mathbf{X}'}(\mathbf{X}')$ indicates the probability density function of \mathbf{X}' and $f_{\mathbf{X}'\mathbf{Y}'}(\mathbf{X}', \mathbf{Y}'; \mathbf{a} - \mathbf{b})$ is the joint probability density function of \mathbf{X}' and \mathbf{Y}' . Then,

$$\langle c(\mathbf{x}, t) \rangle = \int_{V_0} C_0(\mathbf{a}) f_{\mathbf{X}'}(\mathbf{x} - \mathbf{a} - \mathbf{U}t) d\mathbf{a} \quad (12)$$

and

$$\begin{aligned} \langle c(\mathbf{x}, t)c(\mathbf{y}, t) \rangle &= \int_{V_0} \int_{V_0} C_0(\mathbf{a}) C_0(\mathbf{b}) \times \\ &\quad \times f_{\mathbf{X}'\mathbf{Y}'}(\mathbf{x} - \mathbf{a} - \mathbf{U}t, \mathbf{y} - \mathbf{b} - \mathbf{U}t; \mathbf{a} - \mathbf{b}) d\mathbf{a} d\mathbf{b}. \end{aligned} \quad (13)$$

Due to the stationarity of the velocity field, and by virtue of the central limit theorem, $f_{\mathbf{X}'}$ and $f_{\mathbf{X}'\mathbf{Y}'}$ are assumed to be asymptotically normal (Pannone and Kitanidis, 1999).

This means that these distributions tend to become normal as time passes and, in practical terms, that they are approximated by normal ones. Thus, they are completely described by first and second moments of \mathbf{X}' and \mathbf{Y}' ,

$$f_{\mathbf{X}'}(\mathbf{X}') = (2\pi)^{-n/2} |\mathbf{P}|^{-1/2} \exp\left(-\frac{1}{2}\mathbf{X}'^T \mathbf{P}^{-1} \mathbf{X}'\right), \quad (14)$$

$$\begin{aligned} f_{\mathbf{X}'\mathbf{Y}'}(\mathbf{X}', \mathbf{Y}'; \mathbf{a} - \mathbf{b}) &= f_{\mathbf{Y}'|\mathbf{X}'}(\mathbf{Y}'|\mathbf{X}') f_{\mathbf{X}'}(\mathbf{X}') \\ &= (2\pi)^{-n} \left| \begin{array}{cc} \mathbf{P} & \mathbf{W} \\ \mathbf{W}^T & \mathbf{P} \end{array} \right|^{-1/2} \times \\ &\quad \times \exp\left(-\frac{1}{2} \begin{bmatrix} \mathbf{X}'^T & \mathbf{Y}'^T \end{bmatrix} \begin{pmatrix} \mathbf{P} & \mathbf{W} \\ \mathbf{W}^T & \mathbf{P} \end{pmatrix}^{-1} \begin{bmatrix} \mathbf{X}' \\ \mathbf{Y}' \end{bmatrix}\right), \end{aligned} \quad (15)$$

where n is the dimensionality of the flow domain; $|\mathbf{A}| = \det \mathbf{A}$; the vertical line indicates statistical conditioning, and the tensors of single- and two-particle covariances, \mathbf{P} and \mathbf{W} , are defined by Equations (4) and (3). \mathbf{P} and \mathbf{W} depend on t and \mathbf{W} also depends on the distance $\mathbf{a} - \mathbf{b}$.

Equations (12) through (15) apply for any initial condition. For further simplification, consider now that the solute mass is injected at a single point \mathbf{a}_0 ,

$$C_0(\mathbf{a}) = \frac{M}{\eta} \delta(\mathbf{a} - \mathbf{a}_0), \quad (16)$$

$$C_0(\mathbf{b}) = \frac{M}{\eta} \delta(\mathbf{b} - \mathbf{a}_0), \quad (17)$$

where M is the total mass and η the porosity. Then,

$$\langle c(\mathbf{x}, t) \rangle = \frac{M}{\eta} f_{\mathbf{X}'}(\mathbf{x} - \mathbf{a}_0 - \mathbf{U}t) \quad (18)$$

and

$$\langle c(\mathbf{x}, t)c(\mathbf{y}, t) \rangle = \left(\frac{M}{\eta}\right)^2 f_{\mathbf{X}'\mathbf{Y}'}(\mathbf{x} - \mathbf{a}_0 - \mathbf{U}t, \mathbf{y} - \mathbf{a}_0 - \mathbf{U}t) \quad (19)$$

where, now, the two-particle statistics are for two particles starting at the same location.

In terms of spatial covariance, we obtain

$$\begin{aligned} R_c(\mathbf{x}, \mathbf{y}, t) &= \left(\frac{M}{\eta}\right)^2 f_{\mathbf{X}'\mathbf{Y}'}(\mathbf{x} - \mathbf{a}_0 - \mathbf{U}t, \mathbf{y} - \mathbf{a}_0 - \mathbf{U}t) - \\ &\quad - \left(\frac{M}{\eta}\right)^2 f_{\mathbf{X}'}(\mathbf{x} - \mathbf{a}_0 - \mathbf{U}t) f_{\mathbf{X}'}(\mathbf{y} - \mathbf{a}_0 - \mathbf{U}t). \end{aligned} \quad (20)$$

3. Estimation of Two-particle Moments and Kriging

The concentration c at time t is modeled as a nonstationary random field with the first and second moments given in the previous section. Kriging can be used to condition the predicted concentration on the observations of c that are available at a number of locations. Before applying kriging, however, the observations can be used to determine the model parameters. The analysis from this point on is for a specific time, though we will not show the dependence on t .

Parameter estimation can be accomplished using a cross-validation approach, that is, by selecting parameters that reproduce one set of observations given another set (Kitanidis, 1991). Since the model is described through only the first and second moments, kriging predictors are reasonable. Cross-validation using kriging may be applied through a maximum likelihood approach, assuming that the concentration field is Gaussian. This procedure allows us to search for the values of the parameters that best explain the oscillations of the actual concentration field about its ensemble mean.

Consider that at a time t there are N observations z_i . If c_i is the actual concentration and e_i the measurement error,

$$z_i = c_i + e_i \quad i = 1, \dots, N. \quad (21)$$

These errors are modeled as normally distributed, with zero mean and covariance function R_e . It is reasonable to assume that they are uncorrelated and that the standard deviation σ_e linearly increases with the measured concentration

$$R_e(\mathbf{x}_i, \mathbf{x}_j) = \begin{cases} \sigma_e^2(\mathbf{x}_i) & \text{if } i = j \\ 0 & \text{if } i \neq j, \end{cases} \quad (22)$$

where

$$\sigma_e(\mathbf{x}_i) = \alpha z(\mathbf{x}_i) + \beta. \quad (23)$$

In (23), $z_i = z(\mathbf{x}_i)$ indicates the observed concentration at location \mathbf{x}_i ; for concentrations given in mg/l or ppm, $\beta = 0.1$ mg/l represents a minimum error that accounts for the uncertainty related to measurement of values near the detection limit. The coefficient α represents the relative error. We test two reasonable (G. Hopkins, M. Reinhard, personal communication) values of α : 0.05 and 0.07. Additional studies should be performed to evaluate the reliability of the data.

A priori, the observation vector \mathbf{z} is Gaussian with mean $\langle \mathbf{c} \rangle$ and covariance matrix $\mathbf{\Gamma} + \mathbf{R}_e$. Here, $\langle \mathbf{c} \rangle$ is the N by 1 vector of unconditional mean values at observation locations; $\mathbf{\Gamma}$ is the N by N symmetric matrix whose ij th element is the unconditional covariance of concentration at locations i and j , and \mathbf{R}_e is the N by N diagonal matrix whose ii th element is the error variance at location i . Thus, the negative logarithm of the pdf of \mathbf{z} given the structural parameters is, except for a constant term

$$L = \frac{1}{2} \ln |\mathbf{\Gamma} + \mathbf{R}_e| + \frac{1}{2} (\mathbf{z} - \langle \mathbf{c} \rangle)^T (\mathbf{\Gamma} + \mathbf{R}_e)^{-1} (\mathbf{z} - \langle \mathbf{c} \rangle). \quad (24)$$

The *a posteriori*, that is, given the observations, estimates of the structural parameters can then be determined by minimizing L with respect to them.

The structural parameters in this problem are the quantities M , \mathbf{U} , \mathbf{P} , and \mathbf{W} . The first three of them are needed to estimate the mean concentration and have thus been evaluated in previous studies (Garabedian *et al.*, 1991; Thierrin and Kitanidis, 1994). We will use these estimates and we will limit our attention to finding \mathbf{W} . Thus, our problem is to minimize L as function of the different W_{ij} .

Calculation of the second moments for the field solute distribution (Garabedian *et al.*, 1991) produced a non-diagonal tensor of solute variance terms. However, to make the investigation possible within a reasonable effort, we have performed all the computations in a reference frame whose axes are parallel to the plume principal directions. In this case, both \mathbf{P} and \mathbf{W} are diagonal matrices

$$\mathbf{P} = \begin{bmatrix} P_{11} & 0 & 0 \\ 0 & P_{22} & 0 \\ 0 & 0 & P_{33} \end{bmatrix}, \quad (25)$$

$$\mathbf{W} = \begin{bmatrix} W_{11} & 0 & 0 \\ 0 & W_{22} & 0 \\ 0 & 0 & W_{33} \end{bmatrix}. \quad (26)$$

In order to estimate the structural parameters, W_{ii} , we have chosen the Gauss–Newton method (see Kitanidis and Lane, 1985), an iterative optimization procedure

$$\boldsymbol{\theta}_{l+1} = \boldsymbol{\theta}_l - \mathbf{F}_l^{-1} \mathbf{g}_l. \quad (27)$$

In (27), l represents the step counter, $\boldsymbol{\theta} = [W_{11} \ W_{22} \ W_{33}]^T$ is the vector of the unknowns, \mathbf{g} is the vector of the gradients of the negative loglikelihood function L with respect to $\boldsymbol{\theta}$, and \mathbf{F} indicates the Fisher information matrix. The Fisher information matrix is an approximation of the Hessian \mathbf{H} , which is the matrix of the second derivatives.

Equation (27) can be easily derived by approximating the negative loglikelihood function with a quadratic polynomial in terms of vector of unknowns. Through a Taylor series expansion about the most recent guess $\boldsymbol{\theta}_l$, where \mathbf{H} is replaced by \mathbf{F} , we get

$$L(\boldsymbol{\theta}) = L(\boldsymbol{\theta}_l) + \mathbf{g}_l^T (\boldsymbol{\theta} - \boldsymbol{\theta}_l) + \frac{1}{2} (\boldsymbol{\theta} - \boldsymbol{\theta}_l)^T \mathbf{F}_l (\boldsymbol{\theta} - \boldsymbol{\theta}_l) + O(\|\boldsymbol{\theta} - \boldsymbol{\theta}_l\|^3). \quad (28)$$

Then, neglecting the higher order terms and minimization of the right-hand side of (28) lead to (27).

The different components of \mathbf{F} are obtained by taking the expected values of the corresponding components of \mathbf{H} ; that is, by substituting $(\mathbf{z} - \langle \mathbf{c} \rangle)$ and $(\mathbf{z} - \langle \mathbf{c} \rangle)(\mathbf{z} - \langle \mathbf{c} \rangle)^T$ with their mean values $\mathbf{0}$ and $(\boldsymbol{\Gamma} + \mathbf{R}_e)$, respectively. Compared to Newton's method (that uses (28) without the substitution $\mathbf{H} \rightarrow \mathbf{F}$), the Gauss–Newton's method offers some important advantages – first, it avoids the

cumbersome numerical calculation of the second derivatives of $\mathbf{\Gamma}$; second, it often converges faster.

It is possible to show that, in this case, \mathbf{g} is given by

$$\mathbf{g} = \begin{bmatrix} \frac{\partial L}{\partial W_{11}} \\ \frac{\partial L}{\partial W_{22}} \\ \frac{\partial L}{\partial W_{33}} \end{bmatrix} \quad (29)$$

$$= \begin{bmatrix} \frac{1}{2}Tr \left[(\mathbf{\Gamma} + \mathbf{R}_e)^{-1} \frac{\partial \mathbf{\Gamma}}{\partial W_{11}} \right] - \frac{1}{2} (\mathbf{z} - \langle \mathbf{c} \rangle)^T \left[(\mathbf{\Gamma} + \mathbf{R}_e)^{-1} \frac{\partial \mathbf{\Gamma}}{\partial W_{11}} (\mathbf{\Gamma} + \mathbf{R}_e)^{-1} \right] (\mathbf{z} - \langle \mathbf{c} \rangle) \\ \frac{1}{2}Tr \left[(\mathbf{\Gamma} + \mathbf{R}_e)^{-1} \frac{\partial \mathbf{\Gamma}}{\partial W_{22}} \right] - \frac{1}{2} (\mathbf{z} - \langle \mathbf{c} \rangle)^T \left[(\mathbf{\Gamma} + \mathbf{R}_e)^{-1} \frac{\partial \mathbf{\Gamma}}{\partial W_{22}} (\mathbf{\Gamma} + \mathbf{R}_e)^{-1} \right] (\mathbf{z} - \langle \mathbf{c} \rangle) \\ \frac{1}{2}Tr \left[(\mathbf{\Gamma} + \mathbf{R}_e)^{-1} \frac{\partial \mathbf{\Gamma}}{\partial W_{33}} \right] - \frac{1}{2} (\mathbf{z} - \langle \mathbf{c} \rangle)^T \left[(\mathbf{\Gamma} + \mathbf{R}_e)^{-1} \frac{\partial \mathbf{\Gamma}}{\partial W_{33}} (\mathbf{\Gamma} + \mathbf{R}_e)^{-1} \right] (\mathbf{z} - \langle \mathbf{c} \rangle) \end{bmatrix},$$

while \mathbf{F} is defined by

$$\mathbf{F} = \begin{bmatrix} \left\langle \frac{\partial^2 L}{\partial W_{11}^2} \right\rangle & \left\langle \frac{\partial^2 L}{\partial W_{11} \partial W_{22}} \right\rangle & \left\langle \frac{\partial^2 L}{\partial W_{11} \partial W_{33}} \right\rangle \\ \left\langle \frac{\partial^2 L}{\partial W_{22} \partial W_{11}} \right\rangle & \left\langle \frac{\partial^2 L}{\partial W_{22}^2} \right\rangle & \left\langle \frac{\partial^2 L}{\partial W_{22} \partial W_{33}} \right\rangle \\ \left\langle \frac{\partial^2 L}{\partial W_{33} \partial W_{11}} \right\rangle & \left\langle \frac{\partial^2 L}{\partial W_{33} \partial W_{22}} \right\rangle & \left\langle \frac{\partial^2 L}{\partial W_{33}^2} \right\rangle \end{bmatrix}, \quad (30)$$

$$F_{ij} = \left\langle \frac{\partial^2 L}{\partial W_{ii} \partial W_{jj}} \right\rangle = \frac{1}{2} Tr \left[(\mathbf{\Gamma} + \mathbf{R}_e)^{-1} \frac{\partial \mathbf{\Gamma}}{\partial W_{ii}} (\mathbf{\Gamma} + \mathbf{R}_e)^{-1} \frac{\partial \mathbf{\Gamma}}{\partial W_{jj}} \right] \quad (31)$$

and Tr indicates matrix trace, that is, the sum of the diagonal entries.

To further improve the performance of the method, we have resorted to a type of Lavenberg–Marquardt procedure that consists of reducing the length of each step and turning the step direction slightly toward the steepest-descent direction

$$\boldsymbol{\theta}_{l+1} = \boldsymbol{\theta}_l - (\mathbf{F}_l + \nu \mathbf{I})^{-1} \mathbf{g}_l. \quad (32)$$

In (32), \mathbf{I} is the identity matrix and ν is a multiple of the smallest eigenvalue of \mathbf{F}_l . This modification is especially helpful at early iterations, by preventing divergence, if the starting point is far from the optimum. However, it tends to slow down the convergence of the algorithm.

For each sampling round of the Cape Cod tracer test that we have analyzed, the iterations have been stopped when the gradient of L with respect to the unknown

parameters was practically zero. In the neighborhood of the optimum $\boldsymbol{\theta}^*$, with $\mathbf{g}^* = \mathbf{0}$, we have

$$\begin{aligned} L(\boldsymbol{\theta}) &= L(\boldsymbol{\theta}^*) + \frac{1}{2}(\boldsymbol{\theta} - \boldsymbol{\theta}^*)^T \mathbf{H}^* (\boldsymbol{\theta} - \boldsymbol{\theta}^*) + O(\|\boldsymbol{\theta} - \boldsymbol{\theta}^*\|^3) \\ &\simeq L(\boldsymbol{\theta}^*) + \frac{1}{2}(\boldsymbol{\theta} - \boldsymbol{\theta}^*)^T \mathbf{F}^* (\boldsymbol{\theta} - \boldsymbol{\theta}^*). \end{aligned} \quad (33)$$

Thus, at a point $\boldsymbol{\theta}_l$ very near to the optimum $\boldsymbol{\theta}^*$, the gradient is

$$\mathbf{g}_l \simeq \mathbf{F}^* (\boldsymbol{\theta}_l - \boldsymbol{\theta}^*) \quad (34)$$

and

$$L(\boldsymbol{\theta}_l) - L(\boldsymbol{\theta}^*) \simeq \frac{1}{2} \mathbf{g}_l^T (\mathbf{F}^*)^{-1} \mathbf{g}_l. \quad (35)$$

The convergence criterion is then equivalent to

$$\mathbf{g}_l^T (\mathbf{F}^*)^{-1} \mathbf{g}_l \leq \epsilon, \quad (36)$$

where ϵ is a small dimensionless number. In the results presented later, we have used $\epsilon = 0.5$ and we have approximated \mathbf{F}^* by \mathbf{F} at the latest iteration.

The consistency of the model with the data can be verified after the fitting of the parameters through the analysis of the orthonormal residuals (Kitanidis, 1997). In our case, orthonormal residuals are a set of N uncorrelated random variables, expressed by

$$\varepsilon_i = \frac{\gamma_i}{d_i}, \quad i = 1, \dots, N \quad (37)$$

where

$$\gamma_i = \{\mathbf{T}(\mathbf{z} - \langle \mathbf{c} \rangle)\}_i, \quad (38)$$

$$d_i = \{[\text{diag}[\mathbf{T}(\boldsymbol{\Gamma} + \mathbf{R}_e)\mathbf{T}^T]]_i\}^{1/2} \quad (39)$$

and \mathbf{T} is a matrix such that $\mathbf{T}(\boldsymbol{\Gamma} + \mathbf{R}_e)\mathbf{T}^T$ is diagonal. In (39), $\text{diag}[\mathbf{A}]$ indicates the vector of the diagonal elements of \mathbf{A} .

In order to compare the actual performance of the model with the expected one, we have computed the following statistics:

$$Q_1 = \frac{1}{N} \sum_{i=1}^N \varepsilon_i \quad (40)$$

and

$$Q_2 = \frac{1}{N} \sum_{i=1}^N \varepsilon_i^2. \quad (41)$$

Q_1 should be a normally distributed random variable with mean 0 and variance $1/N$. Q_2 , when N is large, is approximately Gaussian with mean 1 and variance $2/N$. Thus, for model validation, we checked that the ε_i distribution was at least approximately normal, and that both Q_1 and Q_2 were reasonably close to their theoretical mean values 0 and 1, respectively. Additionally, we judged the predictive ability of the model through

$$cR = Q_2 \exp \left[\frac{1}{N} \sum_{i=1}^N \ln (d_i^2) \right], \quad (42)$$

which is a measure of the estimation error of the residuals. The smaller the value of cR , the better the predictive ability of the model.

With the correlation structure completely defined through the variance–covariance matrix $\mathbf{\Gamma}$, it is possible to perform interpolation through kriging, in order to get the best estimate of the concentration on a regular grid as a linear weighting of the data. As before, \mathbf{z} is the N by 1 vector of observations, with mean $\langle \mathbf{c} \rangle_z$ and covariance matrix $(\mathbf{\Gamma} + \mathbf{R}_e)_{zz}$; let \mathbf{s} be the M by 1 vector of unknowns with mean $\langle \mathbf{c} \rangle_s$ and covariance matrix $\mathbf{\Gamma}_{ss}$. The M by N measurement-unknown cross-covariance is $\mathbf{\Gamma}_{sz} = \mathbf{\Gamma}_{zs}^T$. Then, the best estimate in this given-mean case is

$$\hat{\mathbf{s}} = \langle \mathbf{c} \rangle_s + \mathbf{\Lambda} (\mathbf{z} - \langle \mathbf{c} \rangle_z), \quad (43)$$

where $\mathbf{\Lambda}$, the kriging coefficients matrix, is found by solving the linear system

$$\mathbf{\Lambda} (\mathbf{\Gamma} + \mathbf{R}_e)_{zz} = \mathbf{\Gamma}_{sz}. \quad (44)$$

4. Results from Data Analysis

Summarizing the key modeling assumptions made for the analysis of the data:

1. The velocity field is time-independent and statistically homogeneous.
2. The velocity field satisfies the conditions for the application of the central limit theorem to the distribution of particle displacements.
3. Sufficient time has elapsed for the normality assumption used in the determination of the concentration ensemble mean and covariance to be valid.
4. The injection volume is very small compared to the plume size at observation times.
5. The random function that describes the concentration is approximately Gaussian. (Although this assumption is not essential to apply minimum-variance linear estimation methods, these methods work best for near-Gaussian data.)
6. The statistics of the observation errors are known or can be determined.

Since the proposed model is based on a large-time approximation, it has been applied to the analysis of the data collected during two late sampling rounds of the tracer test – 273 and 426 days after injection. The optimization of (24) as

function of the two-particle covariances and the solution of (44) in terms of kriging coefficients have been carried out using all the available data (1923 at $t = 273$ days and 2002 at $t = 426$ days). The dimensions of the interpolation grid are slightly larger than the maximum size of the experimental plume – 100 m along the longitudinal direction, 20 m along the transverse horizontal direction, and 10 m along the vertical one. The grid spacing accounts for the more pronounced irregularity of the concentration in the vertical direction due to the anisotropy

$$\Delta Z = \Delta X/2 = \Delta Y/2 = 0.25 \text{ m.} \quad (45)$$

Here, X , Y and Z stand for the distances measured in meters along the plume principal directions – longitudinal, transverse horizontal, and transverse vertical, respectively. The origin of the principal reference frame is located at the sea level and in correspondence of the horizontal coordinates of the injection point.

At $t = 273$ days, the known structural parameters are (Garabedian *et al.* 1991):

$$M = 4470 \quad [\text{g}],$$

$$\mathbf{a}_0 + \mathbf{U}t = \begin{bmatrix} 115 \\ -2.1 \\ 6.6 \end{bmatrix} \quad [\text{m}],$$

$$\mathbf{P} = \begin{bmatrix} 196 & 0 & 0 \\ 0 & 5.2 & 0 \\ 0 & 0 & 1.05 \end{bmatrix} \quad [\text{m}^2].$$

The estimated effective porosity is $\eta = 0.39$.

Figures 1, 2 and 3 show three vertical longitudinal slices of the bromide plume through the centroid, and for three different models of the measurement-error covariance matrix

$$R_{e,ij}^{(A)} = R_e^{(A)}(\mathbf{x}_i, \mathbf{x}_j) = [0.07z(\mathbf{x}_i) + 0.1]^2 \delta_{ij}, \quad (46)$$

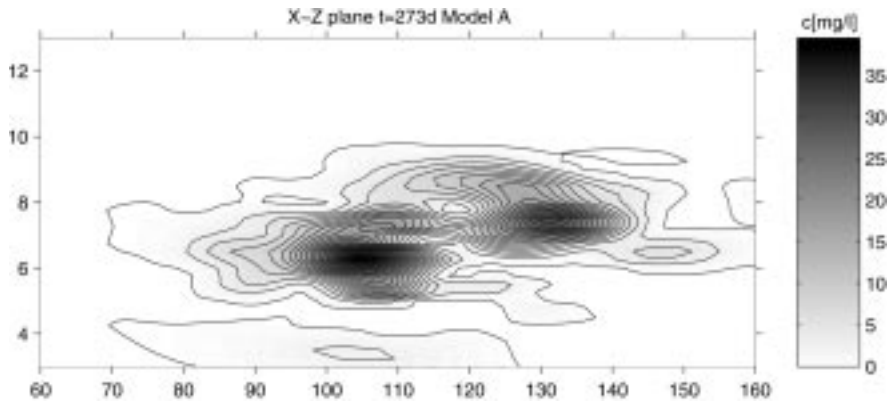


Figure 1. Vertical longitudinal slice through the centroid at $t = 273$ d (Model A).

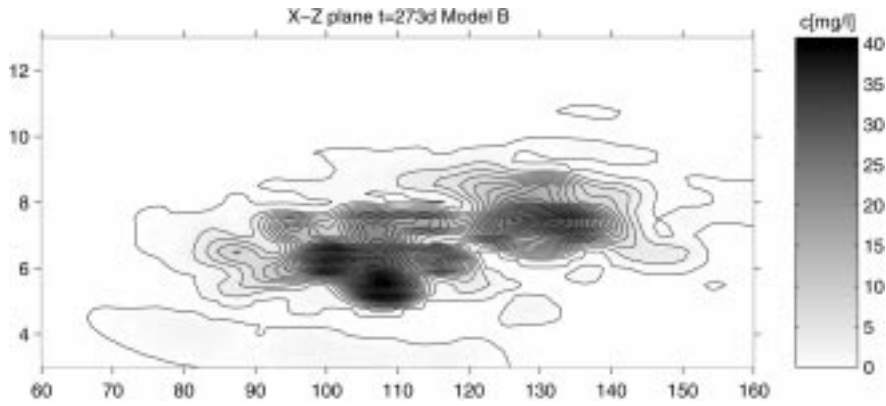


Figure 2. Vertical longitudinal slice through the centroid at $t = 273d$ (Model B).

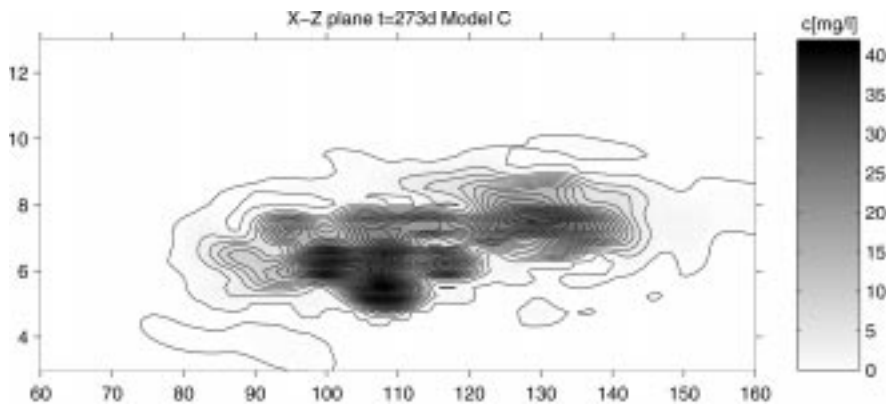


Figure 3. Vertical longitudinal slice through the centroid at $t = 273d$ (Model C).

$$R_{e,ij}^{(B)} = R_e^{(B)}(\mathbf{x}_i, \mathbf{x}_j) = [0.05z(\mathbf{x}_i) + 0.1]^2 \delta_{ij} \quad (47)$$

and

$$R_{e,ij}^{(C)} = R_e^{(C)}(\mathbf{x}_i, \mathbf{x}_j) = \delta_{ij}, \quad (48)$$

where δ_{ij} is Kronecker's delta. In these figures, and in all the figures showing concentration distributions, the exterior contour line delimits the area characterized by $c \geq 0.2$ mg/l. The concentration increment between contiguous contour lines is 2 mg/l.

Figures 4, 5 and 6 show the comparison between fitted (interpolated) and observed concentrations at the sampling points, as well as the distribution of the orthonormal residuals. The larger the measurement-error covariance matrix, \mathbf{R}_e , the larger the difference between observed and fitted values; however, a large difference does not mean that the model is inappropriate. If the measurement-error covariance matrix were zero, the fitted values would reproduce the observed values exactly.

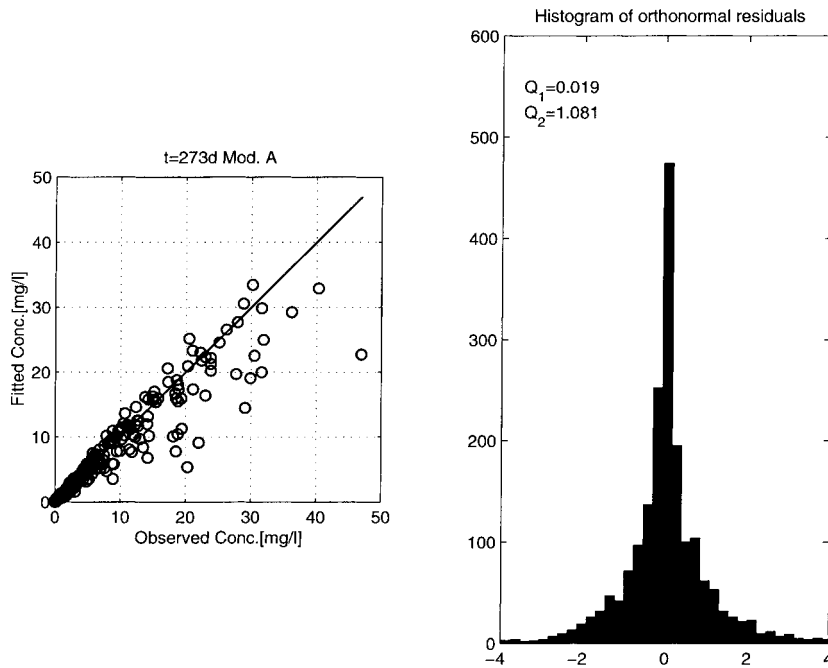


Figure 4. Comparison between fitted and observed concentrations at the sampling points, with corresponding histogram of orthonormal residuals; $t = 273d$ (Model A).

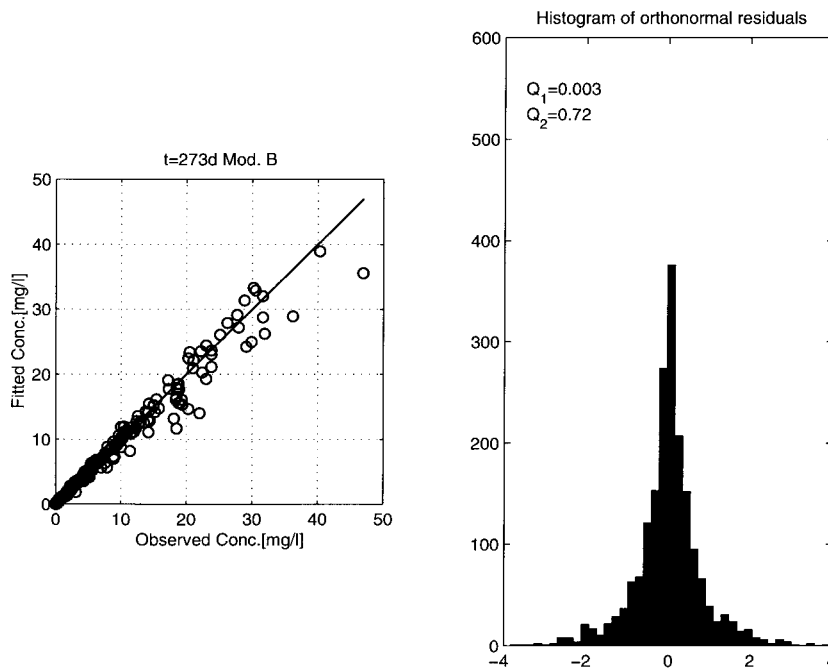


Figure 5. Comparison between fitted and observed concentrations at the sampling points, with corresponding histogram of orthonormal residuals; $t = 273d$ (Model B).

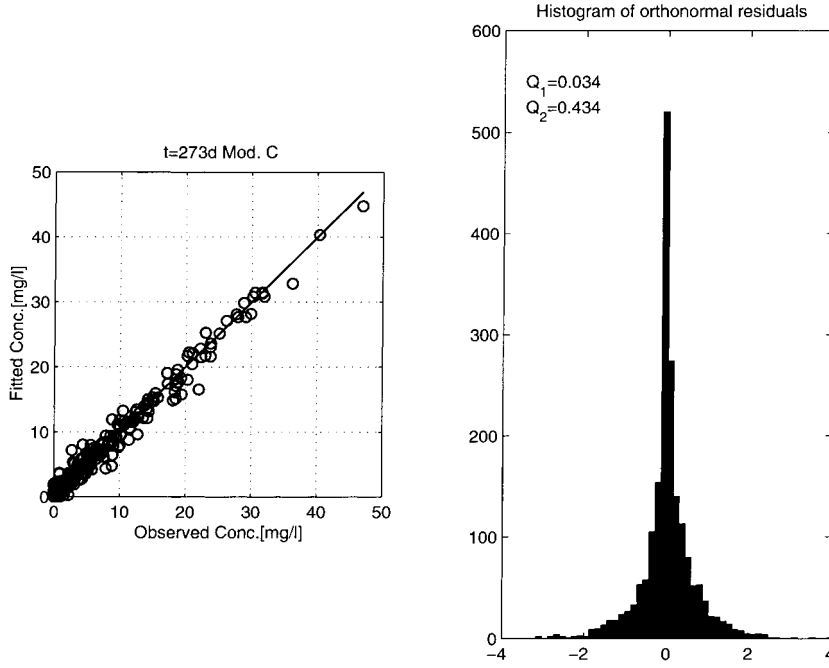


Figure 6. Comparison between fitted and observed concentrations at the sampling points, with corresponding histogram of orthonormal residuals; $t = 273d$ (Model C).

The third constant matrix (48) has been used in order to test the effect of a uniform measurement error on the predictions and on the consistency of the model. Optimization yields the following two-particle covariance tensor

$$\mathbf{W}^{(C)} = \begin{bmatrix} 182.6 & 0 & 0 \\ 0 & 2.36 & 0 \\ 0 & 0 & 0.9 \end{bmatrix} \quad [\text{m}^2]. \quad (49)$$

This error model produces the best agreement between fitted and observed values but performs less well in terms of model consistency as tested through the orthonormal residuals. Indeed,

$$Q_1^{(C)} = 0.034, \quad Q_2^{(C)} = 0.434, \quad cR^{(C)} = 1.318. \quad (50)$$

The variance of the distribution of the orthonormal residuals is clearly far from its unit theoretical mean value.

With $\alpha = 0.07$ (46), we obtain

$$\mathbf{W}^{(A)} = \begin{bmatrix} 157 & 0 & 0 \\ 0 & 3.64 & 0 \\ 0 & 0 & 0.9 \end{bmatrix} \quad [\text{m}^2] \quad (51)$$

and

$$Q_1^{(A)} = 0.019, \quad Q_2^{(A)} = 1.081, \quad cR^{(A)} = 0.185. \quad (52)$$

In this case, the values of Q_1 and Q_2 are very close to the expected ones; plus, the shape of the frequency distribution of the orthonormal residuals is approximately Gaussian.

Finally, for $\alpha = 0.05$ (47), the results are

$$\mathbf{W}^{(B)} = \begin{bmatrix} 180 & 0 & 0 \\ 0 & 3.13 & 0 \\ 0 & 0 & 0.89 \end{bmatrix} \quad [\text{m}^2], \quad (53)$$

$$Q_1^{(B)} = 0.003, \quad Q_2^{(B)} = 0.72, \quad cR^{(B)} = 0.257. \quad (54)$$

Thus, model A, Equation (46), seems to be in better agreement with the data because its Q_1 and Q_2 statistics are closer to their optimum, cR is the smallest among the three models, and the residuals appear more normally distributed than for the other two cases. The estimates of the two-particle covariance matrix \mathbf{W} depend partially on the measurement-error model. This is because the fluctuations in the observations can be attributed to either observation errors, quantified by \mathbf{R}_e , or fluctuations in the concentration, whose variance is affected by the choice of \mathbf{W} . The larger the \mathbf{R}_e , the smaller the estimate of \mathbf{W} . It is clear that the estimation of \mathbf{W} from field measurements is more difficult than the estimation of the single-particle moments. This was to be anticipated, since the two-particle covariance must be evaluated from the concentration variance, which itself is difficult to identify from the data. Nevertheless, by choosing appropriate models and by testing them through cross-validation, we arrived at reasonable estimates. The results indicate that W_{11} , the diagonal element corresponding to the flow direction, is markedly smaller than P_{11} , the one-particle variance. The ratio W_{11}/P_{11} , which is the correlation between the displacement in the flow direction of the two particles at the time of the observation, is 0.80 for model A. Specifically, the difference between the two moments ($196 \text{ m}^2 - 157 \text{ m}^2 = 39 \text{ m}^2$) cannot be attributed to the purely local dispersive contribution (the estimated field local dispersivity was about $1.5 \div 2 \text{ mm}$) and may indicate nonlinear synergies between advection and local dispersion.

Figures 1, 2 and 3 should be compared with Figure 7 that depicts the unconditional mean and compares its predictions at observation points with the actual measurements. The unconditional mean is the prediction in the absence of data or if the data are noninformative; it can be obtained formally when \mathbf{W} is very small in relation to \mathbf{P} . The unconditional mean is a Gaussian function that underestimates the highest concentration values. Note that the contour plots of Figures 1, 2, 3, and 7 show concentration distributions characterized by the same zeroth, first, and second moments. Their pronounced differences are due to different values of \mathbf{W} , which affects the concentration variance and the dilution index, a global measure of concentration nonuniformity (Kitanidis, 1994). Larger \mathbf{W} signifies larger concentration variance and lower dilution index (i.e., less complete dilution).

Figures 8, 9 and 10 show horizontal slices through the centroid for the three different measurement-error models. The horizontal slices are generally less irregular

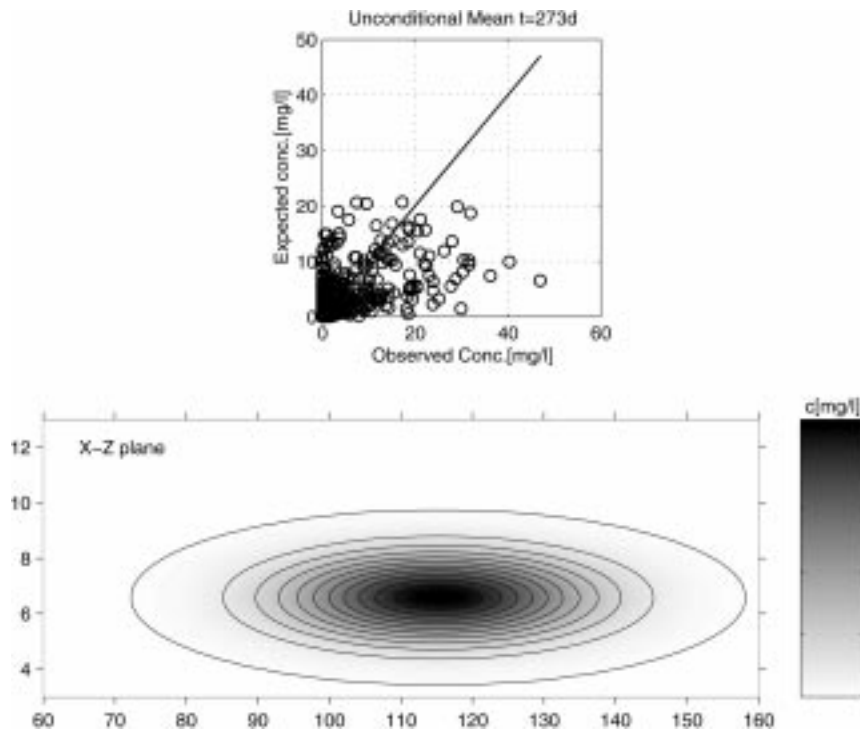


Figure 7. Unconditional mean distribution at $t = 273$ d: Comparison between expected and observed concentrations, with vertical longitudinal slice through the centroid.

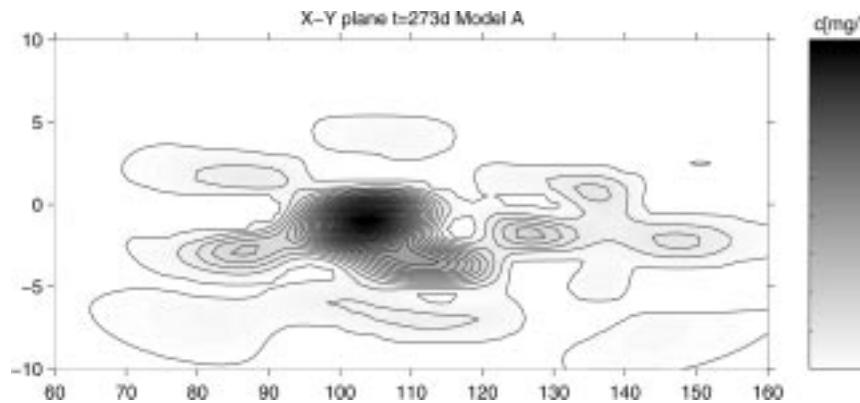


Figure 8. Horizontal slice through the centroid at $t = 273$ d (Model A).

than the vertical distributions. Model $\mathbf{R}_e^{(A)}$ predicts a peak of concentration very close to the measured one (46.9 mg/l), and practically at the same location. Figures 11 and 12 compare two parallel vertical longitudinal slices taken at the same horizontal distance from the centroid, for $\alpha = 0.05$. These plots, which stress that our model describes variability of three-dimensional concentration, show a clearly asymmetric distribution and increased nonuniformity away from the centroid.

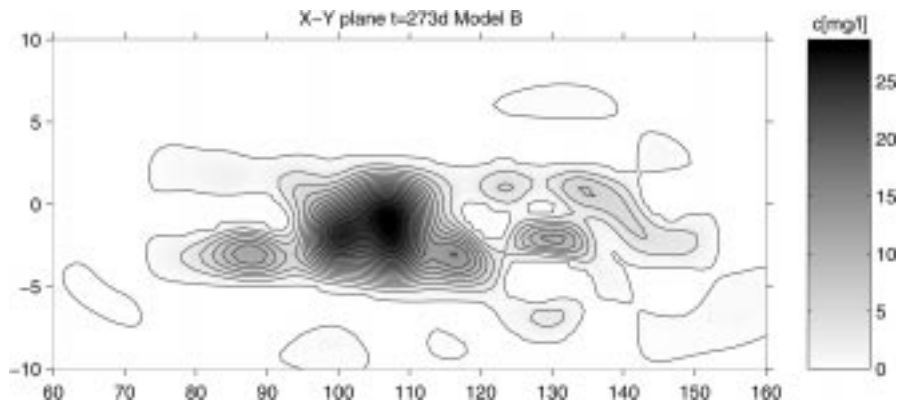


Figure 9. Horizontal slice through the centroid at $t = 273d$ (Model B).

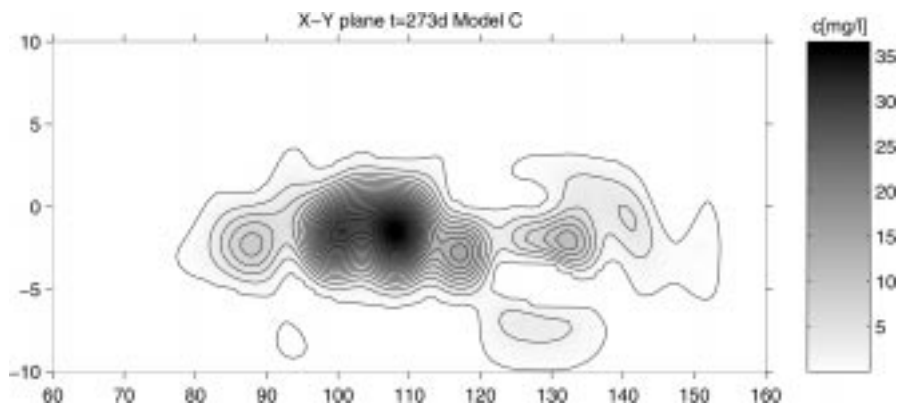


Figure 10. Horizontal slice through the centroid at $t = 273d$ (Model C).

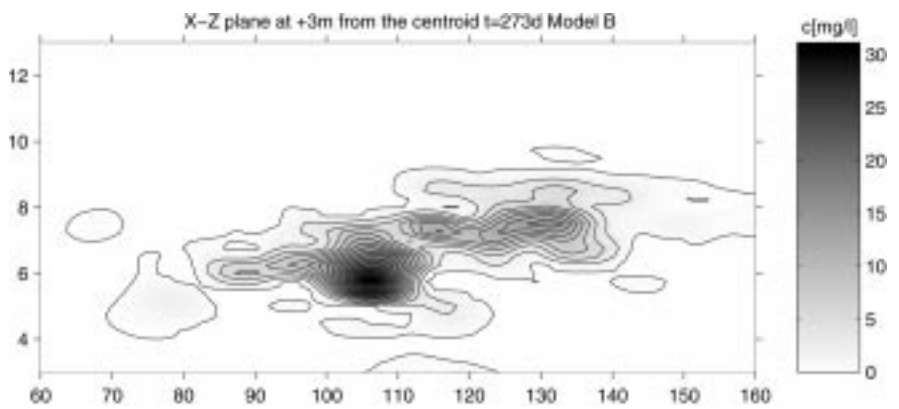


Figure 11. Vertical longitudinal slice at +3 m from the centroid; $t = 273d$ (Model B).

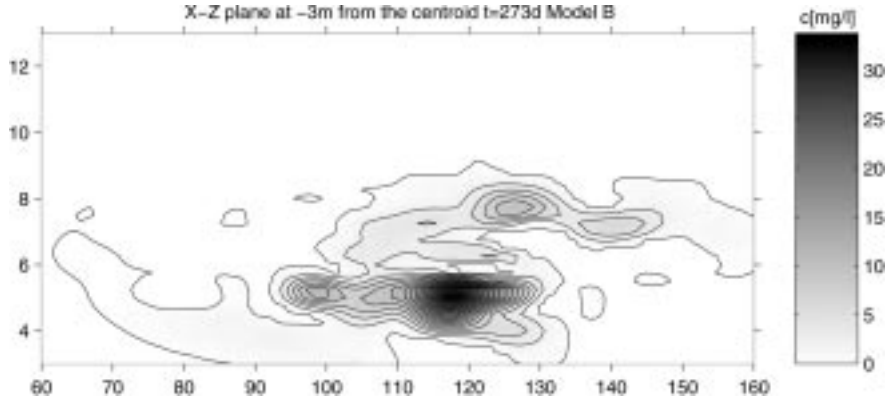


Figure 12. Vertical longitudinal slice at -3 m from the centroid; $t = 273$ d (Model B).

At $t = 426$ days, the known structural parameters are (Garabedian *et al.*, 1991)

$$M = 4940 \quad [\text{g}],$$

$$\mathbf{a}_0 + \mathbf{U}t = \begin{bmatrix} 181 \\ -21.2 \\ 5.4 \end{bmatrix} \quad [\text{m}],$$

$$\mathbf{P} = \begin{bmatrix} 334 & 0 & 0 \\ 0 & 7.7 & 0 \\ 0 & 0 & 0.81 \end{bmatrix} \quad [\text{m}^2].$$

The optimization of the negative loglikelihood function using Model A produces the following two-particle covariance tensor

$$\mathbf{W}^{(A)} = \begin{bmatrix} 226.9 & 0 & 0 \\ 0 & 7.13 & 0 \\ 0 & 0 & 0.5 \end{bmatrix} \quad [\text{m}^2]. \quad (55)$$

Figures 13 and 14 show the vertical and the horizontal slice through the centroid, respectively. Figure 15 illustrates the comparison between fitted and measured concentrations at the observation points, as well as the distribution of the orthonormal residuals. The latter is characterized by a mean and a variance given by

$$Q_1^{(A)} = 0.014, \quad (56)$$

$$Q_2^{(A)} = 0.79. \quad (57)$$

The concentration vertical field shown in Figure 13 is much more diluted than the corresponding distribution at $t = 273$ days (Figure 1). The maximum detected value decreases from ~ 40 mg/l to ~ 25 mg/l; furthermore, the shape of

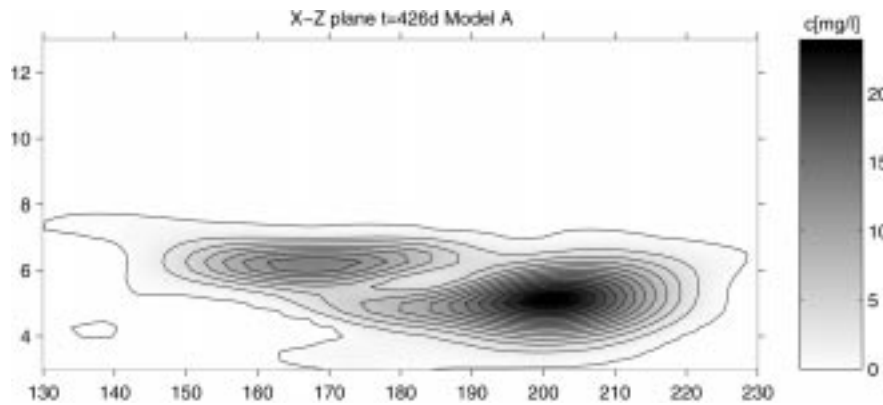


Figure 13. Vertical longitudinal slice through the centroid at $t = 426d$ (Model A).

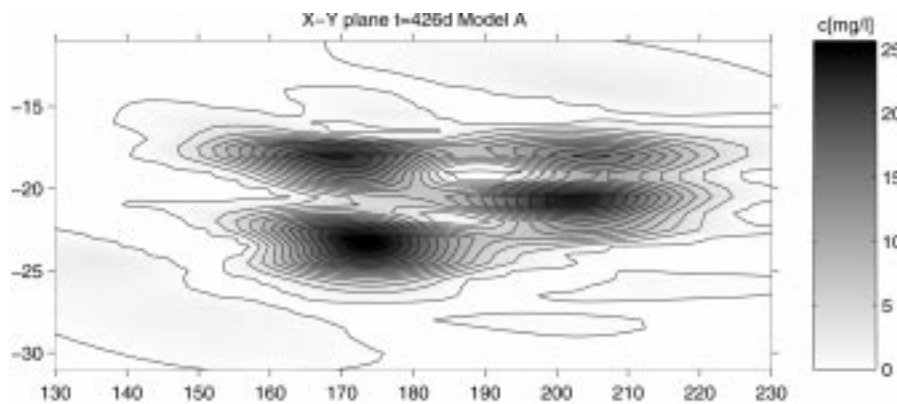


Figure 14. Horizontal slice through the centroid at $t = 426d$ (Model A).

the iso-concentration lines is clearly more regular. The correlation coefficients $\rho_i = W_{ii}/P_{ii}$ are (0.68, 0.93, 0.62) compared to (0.80, 0.70, 0.86) at time 273. We anticipated that the correlation coefficient would decrease over time, indicating that \mathbf{W} increases more slowly than \mathbf{P} as a consequence of local dispersion and perhaps other dispersive processes enhancing dilution. The correlation coefficient corresponding to the longitudinal and the vertical transverse direction indeed dropped from day 273 to day 426. However, the coefficient corresponding to the horizontal transverse direction increased significantly. Compared to the horizontal slice at $t = 273$ days (Figure 8), the concentration distribution of Figure 14 shows that the plume has spread out in the horizontal transverse direction, Y , and the peak concentration has been reduced; however, the concentration in the Y direction seems to be more nonuniform than it was at the earlier time.

Such a result is hard to explain from a purely theoretical standpoint and may be indicative of experimental errors or difficulties associated with trying to estimate the two-particle covariance from a tracer test.

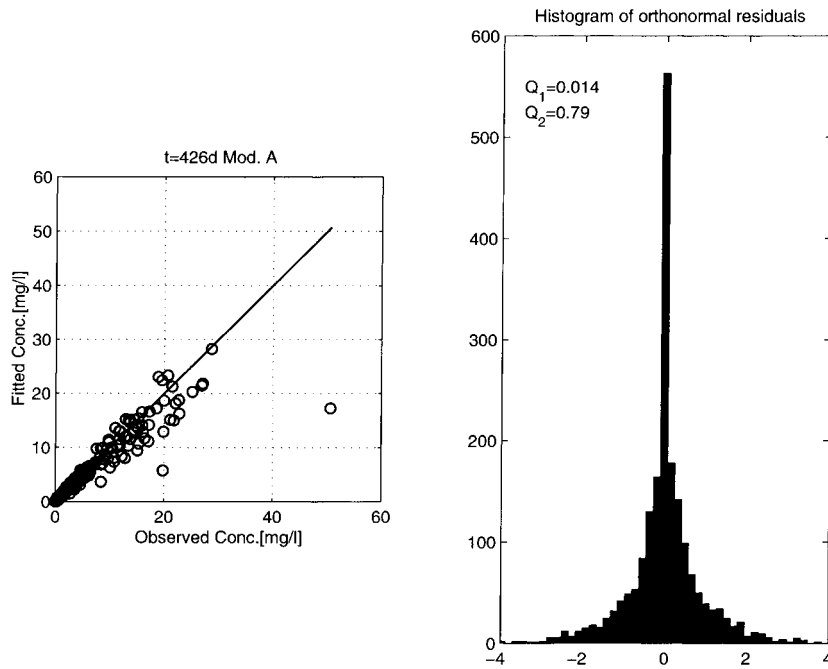


Figure 15. Comparison between fitted and observed concentrations at the sampling points, with corresponding histogram of orthonormal residuals; $t = 426d$ (Model A).

The best estimates at the sampling points are in good agreement with the observations. There is only one point totally outside the range, for which $c \simeq 50$ mg/l. However, this value is two times larger than the one immediately lower; thus, it may have been affected by a particularly large measurement error. The total recovered mass, 4940 g, is quite close to the injected one (4900 g).

5. Discussion and Conclusions

This work includes the development of an analytic expression for the large-time covariance of the concentration of a conservative solute and the analysis of data from the Cape Cod field experiment.

In the first part, we extended the analysis of Pannone and Kitanidis (1999) to derive an expression for the concentration covariance at a given time t , which was shown to depend on the pdf of the single-particle trajectory as well as on the joint pdf of the trajectories of two distinct particles. Then, we assumed that, at large times, these pdf are both Gaussian; as a consequence, the concentration covariance depends on one-particle and two-particle first and second moments only. The one-particle moments had already been analyzed in previous studies. The two-particle covariances are constitutive variables introduced by Pannone and Kitanidis (1999) as global measures of solute nonuniformity.

Since the normality assumption is key to the derivation of the analytical expression, it behooves us to discuss its validity. The normality of the distribution of particle displacement is predicated on the validity of the central limit theorem, considering the total displacement as the sum of many successive steps. In its simplest form, the central value theorem states that the sum of many independent and identically distributed random variables tends to become Gaussian as the sum increases (Parzen, 1960). The theorem can be extended so that, according to Benjamin and Cornell (1970, p. 251), "Under very general conditions, as the number of variables in the sum becomes large, the distribution of the sum of random variables will approach the normal distribution". In our case, the theorem applies if each step of the particle displacement is correlated with only a limited number of other steps, as when the velocity is a stationary random field. The relation of large-time transport theory and the central limit theorem has been studied (e.g., Bhattacharya, 1985; Bhattacharya *et al.*, 1989), for the statistics of a single particle.

For the joint probability density function of the displacements of two particles to be normal, it is sufficient that the marginal pdf and also the pdf of their difference are normal. The marginal pdf, that is, the pdf of a single-particle displacement, is accepted to generally tend to a Gaussian (Dagan, 1989). This conclusion has been directly verified in Dagan and Fiori (1997), on the basis of a 2D numerical experiment performed by A. Bellin. Using the same computational techniques as in Bellin *et al.* (1992), they generated and analyzed data from over 1500 Monte Carlo realizations, for stationary log-conductivity distribution with exponential covariance function, and showed that the single-particle displacement is normal except at very short times. The normality of the single-particle displacement has been explicitly used by Neuman and Zhang (1990) in order to get a quasi-linear solution for the second moments of the plume. Under the same conditions, the pdf of the distance between the two particles tends to become normal because the distance is the sum of many successive increments and each increment is correlated only with a limited number of the others, due to the finite correlation of the velocity field.

Thus, the normality assumption applies to cases in which the central limit theorem is valid, which is considered as the normative case. However, in some circumstances, the validity of the central limit theorem cannot be taken for granted and alternative models have been proposed. Berkowitz and Scher (1998) have used a continuous time random walk approach to describe the anomalous transport taking place in fractured systems and in highly heterogeneous porous formations. In such cases, the flow domain turns out to be made of high velocity channels surrounded by large areas of very slow flow. The concentration distribution becomes increasingly asymmetric and the moment of inertia of the plume increases at the same rate as the square of the displacement of the centroid. Our normality assumption would be inconsistent with such a model.

Next, we analyzed the evolution of the spatial structure of the bromide distribution from an experiment at Cape Cod (Garabedian *et al.*, 1991; LeBlanc *et al.*, 1991). The concentration mean was already modeled in previous studies as normal.

Adding the spatial covariance function, developed in this study, we could perform simple kriging in order to interpolate between observation points. The parameters needed were the total mass, the ensemble mean trajectory, the tensor of one-particle covariances, and the tensor of two-particle covariances. The first three parameters were known from previous studies (Garabedian *et al.*, 1991) but the two-particle covariance needed to be estimated from the data. The estimates obtained using a maximum likelihood method (Kitanidis and Lane, 1985) depended on the model of the accuracy of the observations. We presented the results of three such models for data from sampling period 273 days. In two of the models the standard observation error increased linearly with the measured value and in the third model the standard error was constant. To evaluate which of the three models was the most reliable, we used cross-validation with orthonormal residuals (Kitanidis, 1991). On the basis of cross-validation, we found that the model with the larger observation error appeared to be more consistent with the data.

Using the estimated parameters, we resorted to simple kriging to interpolate the data on a regular grid which was then used for plotting contour maps of the concentration. A particularly interesting conclusion is that the contour maps of the best estimates depend on the assumed accuracy of the observations and the two-particle covariance tensor, for constant solute mass, centroid location, and moments of inertia. The results of kriging give reasonable representations of plumes, with values vanishing away from the centroid, larger nonuniformity at areas of steeper slope, etc. Thus, the model seems to have worked very well as interpolation tool. However, it is difficult to make the case that the estimated two-particle covariance tensors represent faithfully conditions at the Cape Cod site, due to model and data limitations. The data may need to be revisited after more analysis and simulations of transport in domains with known parameters.

Acknowledgements

Funding for this study was provided through a post-doctoral scholarship by University of Basilicata (Potenza) - Italy, and by NSF under projects EAR-9523922, 'Constitutive Relations of Solute Transport and Transformations', and EAR-9522651, 'Factors Affecting Solute Dilution in Heterogeneous Formations'. We thank K. M. Hess, S. P. Garabedian and D. R. Leblanc for providing the data and useful discussions; O. Cirpka and three anonymous reviewers for their comments, that have greatly improved the manuscript.

References

- Batchelor, K.: 1952, Diffusion in a field of homogeneous turbulence, 2. The relative motion of particles, *Proc. Cambridge Philos. Soc.* **48**, 345–362.
- Bear, J.: 1972, *Dynamics of Fluids in Porous Media*, Elsevier, New York, pp. 764.

- Bellin, A., Salandin, P. and Rinaldo, A.: 1992, Simulation of dispersion in heterogeneous formations: Statistics, first-order theories, convergence of computations, *Water Resour. Res.* **28**(9), 2211–2227.
- Benjamin, J. R., and Cornell, C. A.: 1970, *Probability, Statistics, and Decision for Civil Engineers*, MacGraw-Hill, NY, pp. 684.
- Berkowitz, B. and Scher, H.: 1998, Theory of anomalous chemical transport in random fracture networks, *Physical Rev. E* **57**(5), 5858–5869.
- Bhattacharya, R.: 1985, A central limit theorem for diffusions with periodic coefficients, *The Annals of Probab.* **13**(2), 385–396.
- Bhattacharya, R. N., Gupta, V. K. and Walkers, H. F.: 1989, Asymptotics of solute dispersion in periodic porous media, *Siam J. Appl. Math.* **49**(1), 86–98.
- Dagan, G.: 1982, Stochastic modeling of groundwater flow by unconditional and conditional probabilities, 2. The solute transport, *Water Resour. Res.* **18**(4), 835–848.
- Dagan, G.: 1984, Solute transport in heterogeneous porous formations, *J. Fluid Mech.* **145**, 151–177.
- Dagan, G.: 1989, *Flow and Transport in Porous Formations*, Springer-Verlag, New York, pp. 465.
- Dagan, G. and Fiori, A.: 1997, The influence of pore-scale dispersion on concentration statistical moments in transport through heterogeneous aquifers, *Water Resour. Res.* **33**(7), 1595–1605.
- Fiori, A. and Dagan, G.: 1999, Concentration fluctuations in transport by groundwater: Comparison between theory and field experiments, *Water Resour. Res.* **35**(1), 105–112.
- Fisher, H. B., List, E. J., Koh, R. C. Y., Imberger, J. and Brooks, N. H.: 1979, *Mixing in Inland and Coastal Waters*, Academic Press, NY, pp. 483.
- Garabedian, S. P., LeBlanc, D. R., Gelhar, L. W. and Celia, M. A.: 1991, Large-scale natural gradient test in sand and gravel, Cape Cod, Massachusetts, 2: Analysis of spatial moments for nonreactive tracer, *Water Resour. Res.* **27**(5), 911–924.
- Gelhar, L. W. and Axness, C. L.: 1983, Three-dimensional stochastic analysis of macrodispersion in aquifers, *Water Resour. Res.* **19**(1), 161–180.
- Graham, W. D. and McLaughlin, D.: 1989, Stochastic analysis of nonstationary subsurface solute transport: 1. Unconditional moments, *Water Resour. Res.* **25**(2), 215–232.
- Kabala, Z. J.: 1997, Analytical solutions for the coefficient of variation of the volume-averaged solute concentration in heterogeneous aquifers, *Stoch. Hydrol. Hydraul.* **11**(4), 331–348.
- Kapoor, V. and Gelhar, L. W.: 1994a, Transport in three-dimensionally heterogeneous aquifers: 1. Dynamics of concentration fluctuations, *Water Resour. Res.* **30**(6), 1775–1788.
- Kapoor, V. and Gelhar, L. W.: 1994b, Transport in three-dimensionally heterogeneous aquifers: 2. Predictions and observations of concentration fluctuations, *Water Resour. Res.* **30**(6), 1789–1801.
- Kapoor, V. and Kitanidis, P. K.: 1996, Concentration fluctuations and dilution in two-dimensionally periodic heterogeneous porous media, *Transport in Porous Media* **21**(1), 91–119.
- Kapoor, V. and Kitanidis, P. K.: 1997, Advection–diffusion in spatially random flows: formulation for the concentration covariance, *Stoch. Hydrol. Hydraul.* **11**(5), 397–422.
- Kapoor, V. and Kitanidis, P. K.: 1998, Concentration fluctuations and dilution in aquifer, *Water Resour. Res.* **34**(5), 1181–1193.
- Kitanidis, P. K. and Lane, R. W.: 1985, Maximum likelihood parameter estimation of hydrologic spatial processes by Gauss–Newton method, *J. Hydrol.* **79**, 53–71.
- Kitanidis, P. K.: 1991, Orthonormal residuals in geostatistics: model criticism and parameter estimation, *Math. Geol.* **23**(5), 741–758.
- Kitanidis, P. K.: 1994, The concept of the dilution index, *Water Resour. Res.* **30**(7), 2011–2026.
- Kitanidis, P. K.: 1997, *Introduction to Geostatistics*, Cambridge University Press, pp. 249.
- LeBlanc, D. R., Garabedian, S. P., Hess, K. M., Gelhar, L. W., Quadri, R. D., Stollenwerk, K. G. and Wood, W. W.: 1991, Large-scale natural gradient test in sand and gravel, Cape Cod, Massachusetts, 1: Experimental design and observed movement, *Water Resour. Res.* **27**(5), 895–910.

- Neuman, S. P., Winter, C. L. and Newman, C. M.: 1987, Stochastic theory of field-scale Fickian dispersion in anisotropic porous media, *Water Resour. Res.* **23**(3), 453–466.
- Neuman, S. P. and Zhang, Y. K.: 1990, 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.
- Neuman, S. P.: 1993, Eulerian–Lagrangian theory of transport in space-time nonstationary velocity fields: Exact nonlocal formalism by conditional moments and weak approximations, *Water Resour. Res.* **29**(3), 633–645.
- Pannone, M. and Kitanidis, P. K.: 1999, Large time behavior of concentration variance and dilution in heterogeneous formations, *Water Resour. Res.* **35**(3), 623–634.
- Parzen, E.: 1960, *Modern Probability Theory and Its Applications*, Wiley, NY, pp. 464.
- Richardson, L. F.: 1926, Atmospheric diffusion shown on a distance-neighbour graph, *Proc. R. Soc. London, Ser. A* **110**, 709.
- Rajaram, H. and Gelhar, L. W.: 1993, Plume scale-dependent dispersion in heterogeneous aquifers 2., Eulerian analysis and three-dimensional aquifers, *Water Resour. Res.* **29**(9), 3249–3260.
- Rubin, Y.: 1990, Stochastic modeling of macrodispersion in heterogeneous porous media, *Water Resour. Res.* **26**(1), 133–141.
- Rubin, Y.: 1991, Transport in heterogeneous porous media: prediction and uncertainty, *Water Resour. Res.* **27**(7), 1723–1738.
- Thierrin, J. and Kitanidis, P. K.: 1994, Solute dilution at the Borden and Cape Cod groundwater tracer tests, *Water Resour. Res.* **30**(11), 2883–2890.
- Vomvoris, E. G. and Gelhar, L. W.: 1990, Stochastic analysis of the concentration variability in a three-dimensional heterogeneous aquifer, *Water Resour. Res.* **26**(10), 2591–2602.
- Zhang, Y. K. and Neuman, S. P.: 1990, A quasi-linear theory of non-Fickian and Fickian subsurface dispersion, 2. Application to anisotropic media at the Borden site, *Water Resour. Res.* **26**(5), 903–913.
- Zhang, Q.: 1995, Transient behavior of mixing induced by a random velocity field, *Water Resour. Res.* **31**(3), 577–591.
- Zhang, D.: 1995, Impacts of dispersion and first-order decay on solute transport in randomly heterogeneous porous media, *Transport in Porous Media* **21**, 123–144.