

## Hydraulic tomography: Development of a new aquifer test method

T.-C. Jim Yeh and Shuyun Liu

Department of Hydrology and Water Resources, University of Arizona, Tucson

**Abstract.** Hydraulic tomography (i.e., a sequential aquifer test) has recently been proposed as a method for characterizing aquifer heterogeneity. During a hydraulic tomography experiment, water is sequentially pumped from or injected into an aquifer at different vertical portions or intervals of the aquifer. During each pumping or injection, hydraulic head responses of the aquifer at other intervals are monitored, yielding a set of head/discharge (or recharge) data. By sequentially pumping (or injecting) water at one interval and monitoring the steady state head responses at others, many head/discharge (recharge) data sets are obtained. In this study a sequential inverse approach is developed to interpret results of hydraulic tomography. The approach uses an iterative geostatistical inverse method to yield the effective hydraulic conductivity of an aquifer, conditioned on each set of head/discharge data. To efficiently include all the head/discharge data sets, a sequential conditioning method is employed. It uses the estimated hydraulic conductivity field and covariances, conditioned on the previous head/discharge data set, as prior information for next estimations using a new set of pumping data. This inverse approach was first applied to hypothetical, two-dimensional, heterogeneous aquifers to investigate the optimal sampling scheme for the hydraulic tomography, i.e., the design of well spacing, pumping, and monitoring locations. The effects of measurement errors and uncertainties in statistical parameters required by the inverse model were also investigated. Finally, the robustness of this inverse approach was demonstrated through its application to a hypothetical, three-dimensional, heterogeneous aquifer.

### 1. Introduction

Accurate predictions of water and solute distributions and movement in geological formations require detailed knowledge of the spatial distribution of the hydraulic properties of the formations [Yeh, 1992, 1998]. Conventional aquifer tests (also known as pumping tests) assume aquifer homogeneity and yield effective hydraulic conductivity and the storage coefficient for an equivalent homogeneous aquifer. These hydraulic parameters are average properties of the aquifer over a large volume [Butler and Liu, 1993] and do not provide information of spatial distribution of the hydraulic conductivity within the volume. On the other hand, measurement of hydraulic conductivity of small-scale samples at a large number of locations is time-consuming, costly, and impractical.

To circumvent these difficulties and to efficiently gain information of the spatial distribution of hydraulic conductivity, the geophysical tomography concept has recently been employed [Gottlieb and Dietrich, 1995; Butler *et al.*, 1999]. Specifically, fully screened wells are segregated into many vertical intervals using packers. Water is pumped from or injected into an aquifer at one of the intervals to create a steady flow condition. Hydraulic head responses of the aquifer at other intervals are then monitored, yielding a set of head/discharge (or recharge) data. By sequentially pumping (or injecting) water at one interval and monitoring the steady state head response at others, many head/discharge (or recharge) data sets are obtained. Such a sequential aquifer test is referred to as hydraulic tomography. This new field method has significant advantages

over traditional pumping tests. For instance, hydraulic tomography can provide detailed information about vertical and lateral pressure head responses induced by pumping at a given location. Furthermore, by changing the position of the pump in the well, many sets of aquifer responses to pumping at different locations can be obtained. Such a large number of data sets may reduce the nonuniqueness issue of the inverse problem and may reveal the details of a heterogeneous hydraulic conductivity field.

Several researchers have recently investigated this idea of hydraulic tomography. For example, Gottlieb and Dietrich [1995] proposed a method of hydraulic tomography for identifying the permeability distribution in a hypothetical, two-dimensional saturated soil. In their study they used two boreholes to create hydraulic dipoles. The positions of source and sink are varied over both boreholes. Pore water pressure changes along the vertical were monitored in monitoring wells at other locations. They subsequently applied a least squares-based inverse approach to the pressure data to produce an image of the spatial distribution of hydraulic conductivity. Butler *et al.* [1999] applied this hydraulic tomography concept to networks of multilevel sampling wells. They developed new techniques for measuring drawdown data at a scale that had previously been unobtainable. These new techniques greatly facilitate the implementation of hydraulic tomography in the field.

Hydraulic tomography can yield many useful sets of secondary information, namely head responses, that can be used to identify heterogeneity of the aquifer. Still, a reliable and efficient inverse methodology is required to decipher the information so that a reliable image of the hydraulic conductivity field can be obtained. Classical inverse methodologies are known to have many difficulties [Yeh, 1986]. They also confront

Copyright 2000 by the American Geophysical Union.

Paper number 2000WR900114.  
0043-1397/00/2000WR900114\$09.00

an insurmountable computational burden when they are applied to estimate detailed hydraulic properties in three-dimensional geological formations [Kitanidis, 1997]. Consequently, few classical inverse models have been applied to identify small-scale heterogeneity in three-dimensional geological media. More importantly, the abundance of hydraulic head information generated by hydraulic tomography presents an even greater challenge for the classical inverse methodologies.

In the past few decades, cokriging has been used to estimate hydraulic conductivity fields from scattered measurements of pressure head in saturated flow problems [Kitanidis and Vomvoris, 1983; Hoeksema and Kitanidis, 1984]. However, cokriging is a linear estimator, and its application is limited to mildly nonlinear systems, such as groundwater flow in geological formations of mild heterogeneity (variance of natural log of conductivity  $\sigma_{\ln k}^2 = 0.1$ ). When the degree of aquifer heterogeneity is large ( $\sigma_{\ln k}^2 > 1$ ) and the linear assumption becomes inadequate, cokriging cannot provide a good estimate of the conditional mean conductivity field [Yeh et al., 1996]. In other words, it cannot take full advantage of the head information to obtain an optimal estimate of the hydraulic properties.

To overcome this shortcoming, Yeh et al. [1995, 1996] and Zhang and Yeh [1997] developed an iterative geostatistical technique in which a linear estimator was used successively to incorporate the nonlinear relationship between hydraulic properties and pressure head. This method is referred to as a successive linear estimator (SLE). They demonstrated that with the same amount of information the SLE revealed a more detailed conductivity field than cokriging. Hughson and Yeh [1998, 2000] showed that the SLE is computationally efficient compared to the classical inverse method. They extended it to the inverse problem in three-dimensional, variably saturated, heterogeneous porous media, which had not been attempted before. In their study, pressure head and moisture content measurements at 42 locations (7 wells  $\times$  6 depths) in a three-dimensional porous medium were collected at three different times during an infiltration event. This secondary information was then used to estimate saturated hydraulic conductivity  $K_s$  and  $\alpha$  parameter of the Mualem-van Genuchten unsaturated hydraulic property model [van Genuchten, 1980] at 500 locations in the porous medium.

In this paper, on the basis of the SLE we develop a sequential inverse technique for hydraulic tomography to process the large amount of data to characterize aquifer heterogeneity. While demonstrating the robustness of the inverse method, we also investigate the effect of monitoring intervals, pumping intervals, and the number of pumping locations on the final estimate of hydraulic conductivity. Guidelines for optimal design of a hydraulic tomography test are subsequently established. To further verify our results, Monte Carlo inverse simulations are performed, and the effects of measurement errors and uncertainties in statistical parameters required by the inverse model are investigated. Finally, an example is used to illustrate the effectiveness and the robustness of this sequential approach for hydraulic tomography under three-dimensional, steady flow conditions.

## 2. Methodology

### 2.1. Equation of Flow in Three-Dimensional Saturated Media

In this study we assume that the steady state flow field, created by the hydraulic tomography in three-dimensional, sat-

urated, heterogeneous, porous media can be described by the following equation:

$$\nabla \cdot [K(\mathbf{x})\nabla\phi] + Q(\mathbf{x}) = 0 \quad (1)$$

with boundary conditions

$$\phi|_{\Gamma_1} = \phi^*, \quad [K(\mathbf{x})\nabla\phi] \cdot \mathbf{n}|_{\Gamma_2} = q, \quad (2)$$

where  $\phi$  is total head (m),  $\mathbf{x}$  is the spatial coordinate ( $\mathbf{x} = \{x_1, x_2, x_3\}$ , m, and  $x_3$  represents the vertical coordinate and is positive upward),  $Q$  is the pumping rate ( $\text{m}^3/\text{h m}^3$ ) at the selected interval during the tomography experiment, and  $K(\mathbf{x})$  is the saturated hydraulic conductivity field in m/h. In (2), prescribed total head on the Dirichlet boundary  $\Gamma_1$  is denoted by  $\phi^*$  (m). Specified flux  $q$ , in m/h, is given on the Neumann boundary conditions  $\Gamma_2$ , and  $\mathbf{n}$  is a unit vector normal to the union of  $\Gamma_1$  and  $\Gamma_2$ .

### 2.2. Sequential Inverse Algorithm

To deal with aquifer heterogeneity, the natural log of hydraulic conductivity,  $\ln(K(\mathbf{x}))$ , of an aquifer is treated as a stationary stochastic process with an unconditional mean,  $\langle \ln K \rangle = F$  (the angle brackets denote the expected value), and the unconditional perturbation  $f$ . The corresponding steady hydraulic head distribution due to pumping in an interval in the hydraulic tomography is then presented by  $\phi(\mathbf{x}) = H(\mathbf{x}) + h(\mathbf{x})$ , where  $H = \langle \phi \rangle$  and  $h$  is the unconditional head perturbation. Suppose that we have used well log data and core samples to determine  $n_f$  conductivity values  $f_i^* = (\ln K_i^* - F)$ , where  $i = 1, 2, \dots, n_f$  (we will refer to these data sets as primary information). Additionally, we have estimated the mean and correlation structure of the conductivity field. Also assume that during a hydraulic tomography experiment we have collected  $m$  sets of  $n_h$  observed head values  $\phi_j^*$ , where  $j = n_f + 1, n_f + 2, \dots, n_f + mn_h$  during  $m$  sequential pumping tests. These head data sets are referred to as secondary information. We then seek an inverse model that can produce head and conductivity fields that preserve the observed head and conductivity values at sample locations and satisfy their underlying statistical properties (i.e., mean and covariance, etc.) and the governing flow equation. In the conditional probability concept, such a head or conductivity field is a conditional realization of  $\phi$  or  $\ln K$  field among many possible realizations of the ensemble. Consequently, a conditional conductivity field can be expressed as the sum of conditional mean conductivity and its conditional perturbation,  $K_c(\mathbf{x}) = \langle K_c(\mathbf{x}) \rangle + k_c(\mathbf{x})$ . Similarly, the conditional head field can be written as  $\phi_c = \langle \phi_c(\mathbf{x}) \rangle + h_c(\mathbf{x})$  (the subscript  $c$  denotes conditional). While many possible realizations of such conditional  $\ln K$  and  $\phi$  fields exist, the conditional mean fields, i.e.,  $\langle K_c(\mathbf{x}) \rangle$  and  $\langle \phi_c(\mathbf{x}) \rangle$ , are unique. One way to derive these conditional mean fields is to solve the inverse problem in terms of the conditional mean flow equation. The conditional mean equation can be formulated by substituting the conditional stochastic variables into the governing groundwater flow equation (1) and taking the expected value. The conditional mean flow equation then takes the form:

$$\nabla \cdot [\langle K_c(\mathbf{x}) \rangle \nabla \langle \phi_c(\mathbf{x}) \rangle] + \langle \nabla \cdot [k_c(\mathbf{x}) \nabla h_c(\mathbf{x})] \rangle + Q(\mathbf{x}) = 0. \quad (3)$$

We assume that the pumping rate  $Q(\mathbf{x})$  is deterministic. Notice that the true conditional mean  $K$  and  $\phi$  fields do not satisfy the

continuity equation (3) unless the second term in (3) is zero. The second term,  $\langle \nabla \cdot (k_c \nabla h_c) \rangle$ , becomes zero only under two conditions: (1) All the conductivity values in the aquifer are specified (i.e.,  $k_c(\mathbf{x}) = 0$ ) or (2) all the head values in the domain are known (measured) so that  $h_c(\mathbf{x})$  is zero everywhere. In practice, these two conditions will never be met, and we are currently unaware of a means by which to correctly evaluate this term. Accordingly, we will assume that this term is proportional to the conditional mean gradient such that we can rewrite the mean equation as

$$\nabla \cdot [\langle \mathbf{K}_{\text{eff}}(\mathbf{x}) \rangle \nabla \langle \phi_c(\mathbf{x}) \rangle] + Q(\mathbf{x}) = 0. \quad (4)$$

This conditional mean equation has the same form as (1), but it is expressed in terms of the conditional effective conductivity and conditional mean hydraulic head field. The conditional effective conductivity  $\langle \mathbf{K}_{\text{eff}} \rangle$  is a parameter that combines the conditional mean conductivity  $\langle K_c \rangle$  and the ratio of the second term to the conditional mean gradient.

Based on the concept of conditional mean equation, we essentially seek an inverse approach to derive the conditional effective hydraulic conductivity that will produce a conditional mean head field in (4). To do this, we used the SLE, which starts with the classical cokriging technique using observed  $f_i^*$  and  $h_j^*$  collected in one pumping test in the tomography to construct a cokriged, mean-removed log conductivity map. That is,

$$f_k(\mathbf{x}_0) = \sum_{i=1}^{n_f} \lambda_{i0} f_i^*(\mathbf{x}_i) + \sum_{j=n_f+1}^{n_f+n_h} \mu_{j0} h_j^*(\mathbf{x}_j), \quad (5)$$

where  $f_k(\mathbf{x}_0)$  is the cokriged  $f$  value at location  $\mathbf{x}_0$ . Then, conductivity  $K_k(\mathbf{x}_0)$  becomes  $\exp [F + f_k(\mathbf{x}_0)]$ . Here,  $\lambda_{i0}$  and  $\mu_{j0}$  are the cokriging weights associated with  $\mathbf{x}_0$ , which can be evaluated as follows:

$$\sum_{i=1}^{n_f} \lambda_{i0} R_{ff}(\mathbf{x}_\ell, \mathbf{x}_i) + \sum_{j=n_f+1}^{n_f+n_h} \mu_{j0} R_{fh}(\mathbf{x}_\ell, \mathbf{x}_j) = R_{ff}(\mathbf{x}_0, \mathbf{x}_\ell) \quad \ell = 1, 2, \dots, n_f, \quad (6)$$

$$\sum_{i=1}^{n_f} \lambda_{i0} R_{hf}(\mathbf{x}_\ell, \mathbf{x}_i) + \sum_{j=n_f+1}^{n_f+n_h} \mu_{j0} R_{hh}(\mathbf{x}_\ell, \mathbf{x}_j) = R_{hf}(\mathbf{x}_0, \mathbf{x}_\ell)$$

$$\ell = n_f + 1, n_f + 2, \dots, n_f + n_h,$$

where  $R_{ff}$ ,  $R_{hh}$ , and  $R_{fh}$  are covariances of  $f$  and  $h$  and the cross covariance of  $f$  and  $h$ , respectively. The covariance  $R_{hh}$  and the cross-covariance  $R_{fh}$  in (6) are derived from the first-order numerical approximation (similar to equations (9)–(11)) because of its flexibility for cases that involve bounded domains and nonstationary problems.

As discussed in section 1, the information of hydraulic head may not be fully utilized because of the nonlinear relationship between  $f$  and  $h$  and the linear assumption embedded in cokriging. To circumvent this problem, a successive linear estimator is used. That is,

$$\hat{Y}_c^{(r+1)}(\mathbf{x}_0) = \hat{Y}_c^{(r)}(\mathbf{x}_0) + \sum_{j=n_f+1}^{n_f+n_h} \omega_{j0}^{(r)} [\phi_j^*(\mathbf{x}_j) - \phi_j^{(r)}(\mathbf{x}_j)], \quad (7)$$

where  $\omega_{j0}$  is the weighting coefficient for the estimate at location  $\mathbf{x}_0$  with respect to the head measurement at location  $\mathbf{x}_j$  and  $r$  is the iteration index.  $\hat{Y}_c^{(0)}$  is an estimate of the conditional

mean of  $\ln K$ , which is equal to the cokriged log conductivity field  $f_k + F$  at  $r = 0$ . The residual about the mean estimate at an iteration  $r$  is  $y^r$  (i.e.,  $y^r = \ln K - \hat{Y}_c^{(r)}$ ). In (7),  $\phi_j^{(r)}$  is the head at the  $j$ th location of the solution to (4) at iteration  $r$ , and  $\phi_j^*$  is the observed head at location  $j$  (i.e.,  $\phi_j^* = H_j + h_j^*$ ). The values of  $\omega$  are determined by solving the following system of equations:

$$\sum_{j=n_f+1}^{n_f+n_h} \omega_{j0}^{(r)} \varepsilon_{hh}^{(r)}(\mathbf{x}_\ell, \mathbf{x}_j) + \theta \delta_{\ell\ell} = \varepsilon_{hy}^{(r)}(\mathbf{x}_0, \mathbf{x}_\ell) \quad (8)$$

$$\ell = n_f + 1, n_f + 2, \dots, n_f + n_h,$$

where  $\varepsilon_{hh}$  and  $\varepsilon_{hy}$  are the error covariance (or conditional covariance function) and error cross covariance (or conditional cross covariance), respectively, at each iteration,  $\theta$  is a stabilizing term, and  $\delta_{\ell\ell}$  is an identity matrix. During the iteration the stabilizing term is added to the diagonal terms of the left-hand-side matrix of (8) to numerically condition the matrix and thus to assure a stable solution. A larger term can result in a slower convergence rate, and a smaller  $\theta$  value may lead to numerical instability. In our approach, this stabilizing term is determined dynamically as the product of a constant weighting factor and the maximum value of the diagonal terms of  $\varepsilon_{hh}$  at each iteration.

The solution to (8) requires knowledge of  $\varepsilon_{hy}$  and  $\varepsilon_{hh}$ , which is approximated at each iteration. On the basis of the first-order analysis for a finite element groundwater flow model [Dettinger and Wilson, 1981], hydraulic head at the  $r$ th iteration can be written as a first-order Taylor series:

$$\phi = \hat{\phi}_c^{(r)} + h^{(r)} = G(\hat{Y}_c^{(r)} + y^{(r)}) \approx G(\hat{Y}_c^{(r)}) + \left. \frac{\partial G(\hat{Y}_c^{(r)})}{\partial \ln K} \right|_{\hat{Y}_c^{(r)}} y^{(r)}, \quad (9)$$

where  $G(\hat{Y}_c^{(r)})$  represents the resulting head of the conditional mean equation (4) evaluated with parameters  $\hat{Y}_c^{(r)}$ . The first-order approximation of the residual  $h^{(r)}$  can then be written as

$$h^{(r)} \approx \left. \frac{\partial G(\hat{Y}_c^{(r)})}{\partial \ln K} \right|_{\hat{Y}_c^{(r)}} y^{(r)} = J^{(r)} y^{(r)}, \quad (10)$$

where  $J$  can be evaluated using an adjoint state sensitivity method [Sykes et al., 1985; Sun and Yeh, 1992; Li and Yeh, 1998] subject to boundary conditions. Using (11), we then derive the approximate covariance of  $h^{(r)}$  and cross covariances between  $y^{(r)}$  and  $h^{(r)}$ .

$$\varepsilon_{hh}^{(r)} = J^{(r)} \varepsilon_{yy}^{(r)} J^{(r)T}, \quad (11)$$

$$\varepsilon_{hy}^{(r)} = J^{(r)} \varepsilon_{yy}^{(r)}$$

where  $J$  is the sensitivity matrix of  $n_h \times N$ , superscript  $T$  stands for the transpose, and  $\varepsilon_{yy}$  is the covariance of  $y$ , which is given by

$$\varepsilon_{yy}^{(1)}(\mathbf{x}_0, \mathbf{x}_k) = R_{ff}(\mathbf{x}_0, \mathbf{x}_k) - \sum_{i=1}^{n_f} \lambda_{i0} R_{ff}(\mathbf{x}_i, \mathbf{x}_k) - \sum_{j=n_f+1}^{n_f+n_h} \mu_{j0} R_{fh}(\mathbf{x}_j, \mathbf{x}_k) \quad (12)$$

at iteration  $r = 0$ , where  $k = 1, 2, \dots, N$ , and  $\lambda$  and  $\mu$  are cokriging coefficients. Equation (12) is the cokriging variance

if  $\mathbf{x}_0 = \mathbf{x}_k$ . For  $r \geq 1$  the covariances are evaluated according to

$$\varepsilon_{yy}^{(r+1)}(\mathbf{x}_0, \mathbf{x}_k) = \varepsilon_{yy}^{(r)}(\mathbf{x}_0, \mathbf{x}_k) - \sum_{i=n_r+1}^{n_f+n_h} \omega_{i0}^{(r)} \varepsilon_{yh}^{(r)}(\mathbf{x}_i, \mathbf{x}_k). \quad (13)$$

These covariances are approximate conditional covariances. The accuracy of this approximation was investigated by *Hanna and Yeh* [1998] and will be discussed in section 5.

After updating  $Y_c(\mathbf{x})$  the mean flow equation (4) is solved again with the newly updated  $Y_c(\mathbf{x})$  for a new head field,  $\phi$ . Then, the change of  $\sigma_f^2$  (the variance of the estimated conductivity field) and the change of the biggest head misfit among all the monitoring locations between two successive iterations are evaluated. If both changes are smaller than prescribed tolerances, the iteration stops. If not, new  $\varepsilon_{hy}$  and  $\varepsilon_{hh}$  are evaluated using (11). Equation (8) is then solved to obtain a new set of weights, which are used in (7) with  $(\phi_j^* - \phi_j^{(r)})$  to obtain a new estimate of  $Y_c(\mathbf{x})$ .

The above discussion describes the SLE for only one set of primary and secondary information during a hydraulic tomography experiment. This algorithm can also simultaneously include all of the head data collected during all the pumping operations in the sequence. Nevertheless, the system of equations in (6) and (8) can become extremely large and ill conditioned, and stable solutions to the equations can become difficult to obtain [*Hughson and Yeh*, 2000].

To avoid this problem, the head data sets are used sequentially. Specifically, our method starts the iterative process with the available conductivity measurements and the head data set collected from one of the pumping operations. Once the estimated field converges to the given criteria, the newly estimated conductivity field  $\hat{Y}_c$  is the effective conductivity conditioned on head data due to pumping at the first location, and the residual conductivity covariance is the corresponding conditional conductivity covariance. Subsequently, the conditional effective conductivity is used to evaluate the conditional mean head and sensitivity matrix, associated with pumping at the next location. Based on (11), the sensitivity matrix in conjunction with the conditional conductivity covariance then yields the head covariance and cross covariance of head and conductivity that reflect pumping at the next location, which are subsequently employed in (8) to derive the new weights [*Li and Yeh*, 1999]. With the conditional mean heads, the new weights, and the observed heads, (7) yields the conductivity estimate, representing the first estimate based on the information from the pumping at the new location. The iterative process is then employed to include the nonlinear relationship between head and conductivity. The same procedure is used for the next pumping location. In essence, our sequential approach uses the estimated hydraulic conductivity field and covariances, conditioned on previous sets of head measurements, as prior information for the next estimation based on a new set of pumping data. It continues until all the data sets are fully utilized. Such a sequential approach allows accumulation of high-density secondary information obtained from hydraulic tomography, while maintaining the covariance matrix at a manageable size that can be solved with the least numerical difficulties. *Vargas-Guzman and Yeh* [1999] provided a theoretical proof to show that such a sequential approach is identical to the simultaneous approach for linear systems.

### 3. Design Criteria for Hydraulic Tomography

The design of the monitoring network, the pumping location, the number of pumping tests, and the pumping rate can influence the effectiveness of hydraulic tomography. In this section, the "optimum" design of hydraulic tomography is investigated by applying our sequential geostatistical inverse model to a two-dimensional, vertical, hypothetical aquifer. The hydraulic tomography experiment considered consists of two fully screened wells, separated into many vertical intervals by packers, in a confined aquifer. Water is pumped from the aquifer at one of the vertical intervals, and after steady state flow is established, head responses of the aquifer are monitored at the other intervals. The same procedure is then repeated at different pumping locations.

#### 3.1. Aquifer Description and Evaluation Criteria

The hypothetical confined aquifer was assumed to be 20 m  $\times$  20 m and was discretized into 400 elements of 1 m<sup>2</sup>. Each element was assigned a conductivity value using a random field generator [*Gutjahr*, 1989]. This generated conductivity field had a geometric mean of 0.44 m/h and an exponential correlation structure with a variance of 0.63 for  $\ln K$ . The correlation structure was anisotropic: A horizontal correlation scale of 12 m and a vertical correlation scale of 4 m are used. The left and right sides of the aquifer were constant head boundaries (with a prescribed hydraulic head of 80 m), while the top and the bottom sides were set to be no-flux boundaries.

The performance of each network design was evaluated using the average absolute error norm L1 and the mean-square error norm L2, which are defined as follows:

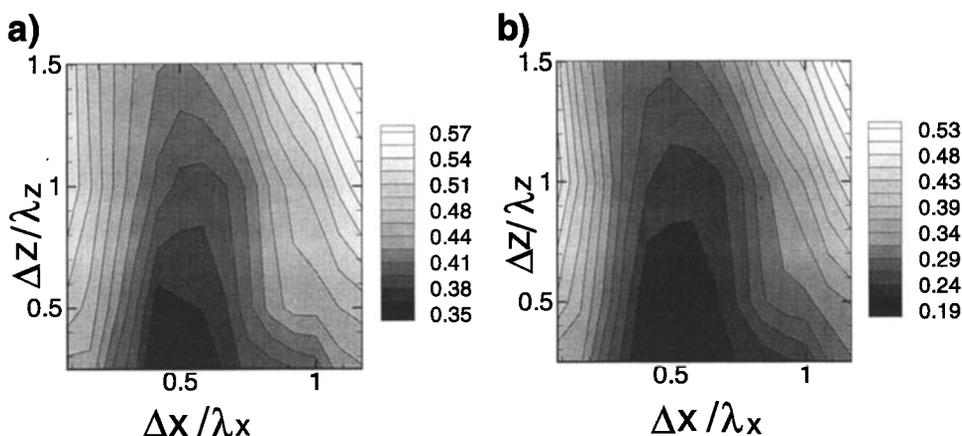
$$L1 = \frac{1}{n} \sum_{i=1}^n |\hat{f}_i - f_i|, \quad L2 = \frac{1}{n} \sum_{i=1}^n (\hat{f}_i - f_i)^2, \quad (14)$$

where  $f_i$  and  $\hat{f}_i$  represent the true and estimated perturbation of the log-transformed conductivity, respectively,  $i$  indicates the element number, and  $n$  is the total number of elements. The smaller the L1 and L2 values are, the better the estimate is.

Conditional variance of estimated conductivity  $\varepsilon_{ff}(\mathbf{x}_0, \mathbf{x}_0)$  (see equations (12) and (13)) is also used to evaluate the performance of the network design. The smaller the variance is, the more accurate the estimate. If the value of conductivity at a location is known exactly, the conditional variance at that location is zero.

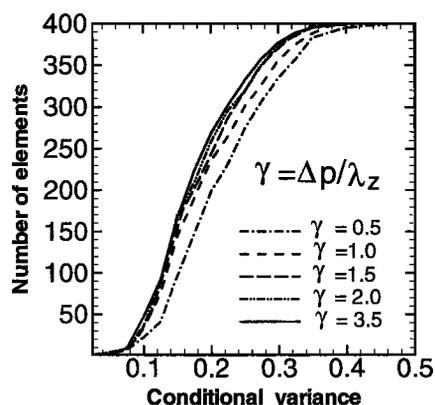
#### 3.2. Optimal Monitoring Network

One factor that must be considered during the design of a hydraulic tomography experiment is the separation distance of the two wells and the interval of packer placements within the well. To address this issue, the first well in the numerical experiments was fixed at one location in the aquifer ( $x = 13.5$  m, where  $x$  is the horizontal coordinate), and the second well was located at various distances to create different configurations. Subsequently, many monitoring network designs using different combinations of well separation distances and packer intervals were examined. For each monitoring network design a steady state flow was established by pumping at the fixed point (13.5 m, 13.5 m) and at a constant rate of 20 m<sup>3</sup>/h. The aquifer head values collected at each monitoring network were then used with our inverse model to estimate the hydraulic conductivity field.



**Figure 1.** (a) Contour map of norm L1 for different designs of hydraulic tomography. (b) Contour map of norm L2 for different designs of hydraulic tomography. Here,  $\Delta x/\lambda_x$  represents the ratio of the well spacing to the horizontal correlation scale and  $\Delta z/\lambda_z$  represents the ratio of the distance between packers to the vertical correlation scale.

Effects of horizontal well spacing and vertical packer intervals are shown in Figures 1a and 1b, where the contour maps of L1 and L2 for different values of  $\Delta x/\lambda_x$  and  $\Delta z/\lambda_z$  are plotted. The correlation scales in the  $x$  direction and the  $z$  direction are denoted by  $\lambda_x$  and  $\lambda_z$ ,  $\Delta x$  is the separation distance between the two wells, and  $\Delta z$  is the vertical distance between neighboring packers. The “optimal” horizontal and vertical intervals are defined as those that yield the minimum of L1 and L2 over the entire domain. According to Figures 1a and 1b the optimal distance between the two wells (horizontal interval) is approximately half of the horizontal correlation scale. This distance cannot be too large or too small because the best estimate of conductivity values is near the vicinity of the wells where pressure changes are collected (see discussion in section 5). The optimal vertical distance between packers along the well (vertical interval) should be as small as possible (at least smaller than 0.5 times the vertical correlation scale). Also shown is that the separation distance between the two wells has more influence on the conductivity estimates than that between vertical monitoring points along the well.



**Figure 2.** Comparison of the goodness of  $f$  estimate for various pumping intervals, represented by the number of elements whose conditional variance is less than a given threshold value. Here,  $\gamma$  is the ratio of the pumping interval  $\Delta p$  to the vertical correlation scale  $\lambda_z$ .

### 3.3. Optimal Pumping Interval

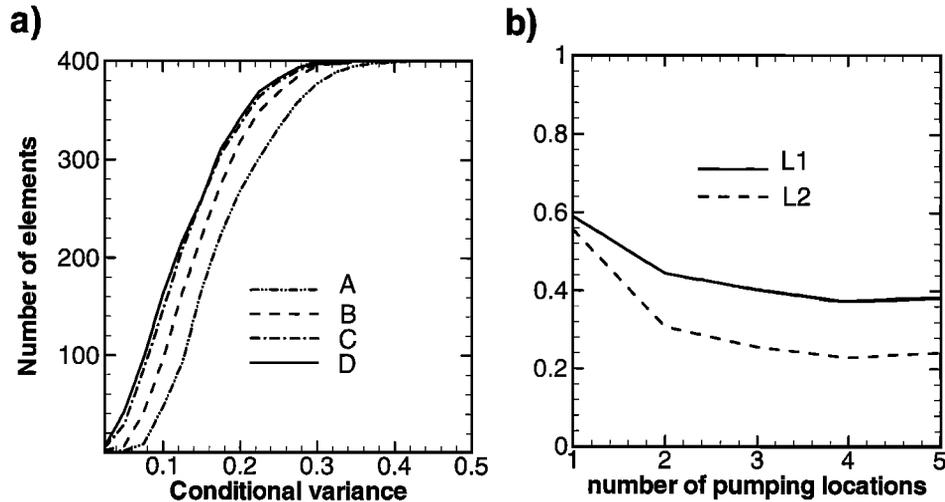
The main idea of hydraulic tomography is to collect a large number of aquifer responses using the same monitoring network by changing the pumping locations along the wells. It is important to know how the choice of pumping interval (the distance between two adjacent pumping locations) influences the effectiveness of hydraulic tomography.

To address this issue, numerical experiments were conducted. On the basis of the results of the previous analysis the horizontal separation distance of the two wells in the experiments was chosen to be 6 m, and the vertical monitoring interval was chosen to be 2 m. Consequently, two wells were set up at  $x = 7.5$  m and  $x = 13.5$  m, and 10 monitoring intervals along each well were employed. The well at  $x = 7.5$  m was chosen as the well where various pumping intervals would be considered. By changing the distance between two adjacent pumping locations from 2 m to 4 m, 6 m, 8 m, etc., the effect of the pumping interval was then evaluated. Figure 2 shows the number of elements of the aquifer with the conditional variance less than some given values for different designs of the pumping interval. On the basis of Figure 2 the size of the pumping interval has little effect on the conditional variance. However, a slightly better estimate is obtained if the pumping interval is greater than 2 m, which is half of the vertical correlation scale. The same result is also obtained by evaluating L1 and L2 for different pumping intervals. Consequently, we conclude that the pumping interval should be greater than the half of the vertical correlation scale.

### 3.4. Optimal Number of Pumping Locations

From the analysis of the optimal pumping interval we found that once the pumping interval is greater than the half of the vertical correlation scale, a further increase of the interval does not significantly improve the  $f$  estimate. Nonetheless, for a given aquifer thickness the larger the pumping interval we select, the fewer the pumping locations we have, and the less information we can obtain from the tomography. Therefore it is imperative to determine the optimal number of pumping locations so that hydraulic tomography can provide sufficient secondary information.

The influences of increasing numbers of pumping locations



**Figure 3.** (a) Comparison of the goodness of  $f$  estimate for various numbers of pumping locations, represented by the number of elements whose conditional variance is less than a given threshold value. A, two pumping locations; B, three pumping locations; C, four pumping locations; and D, five pumping locations. (b) Norm L1 and L2 versus number of pumping locations.

on the effectiveness of the tomography are shown in Figures 3a and 3b. Figure 3a plots the number of  $f$  estimates with conditional variance lower than the specified threshold value for different numbers of pumping locations. For a given threshold value of the conditional variance (for instance, 0.1), as the number of pumping locations increases from 2 to 4, the number of  $f$  estimates with conditional variance smaller than 0.1 increases from 47 to 147. As the number of pumping locations increases to 5, the number of good  $f$  estimates increases from 147 to 164, showing that the rate of improvement decreases. The same trend is also shown in Figure 3b, where the values of L1 and L2 decrease significantly when the number of pumping locations increases from 2 to 4. Then, the decrease becomes moderate, and L1 and L2 gradually approach a constant value when five pumping locations are used. The results show that an increase in the number of pumping locations improves the final  $f$  estimate, but the improvement diminishes as more pumping locations are used, indicating that certain data sets generated from hydraulic tomography may provide redundant information. On the basis of this example the optimal number of pumping location is five (20 m/4 m; here 20 m is the aquifer depth, and 4 m is the vertical correlation scale). For a generic aquifer we may conclude that the optimal number of pumping location is the ratio of the aquifer depth to the vertical correlation scale.

### 3.5. Effect of Pumping Rate

Our numerical experiments show that the pumping rate does not affect the final estimate of conductivity. Under steady state flow conditions an increase in the pumping rate leads to an increase in the hydraulic gradient, which subsequently affects the sensitivity of head with respect to saturated conductivity. Such an increase in hydraulic gradients also results in an increase in head variance, but the cross correlation between the head and the conductivity remains the same. Consequently, the increase in the pumping rate does not affect the cokriging weights and does not influence the estimate [Li and Yeh, 1998]. In other words, different pumping rates will yield identical results. One must recognize, though, that in practice, pressure

head data may be corrupted by noises. Thus an increase in pumping rate may increase the signal-to-noise ratio such that the inversion of hydraulic tomography data can yield better results.

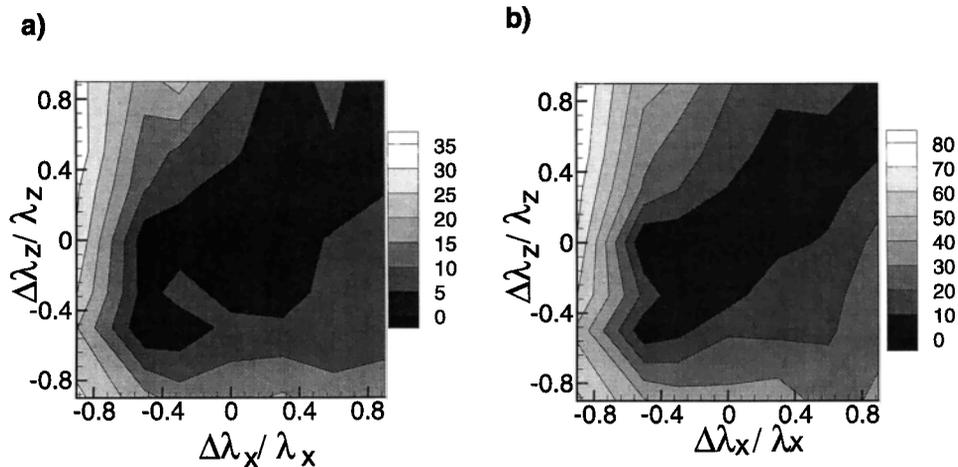
## 4. Uncertainty Analysis

Our inverse method for the hydraulic tomography requires the knowledge of mean, variance, and correlation structure of the conductivity field, head data sets and the associated pumping rates, and some conductivity values if they are available. While head data and pumping rates can be collected during the tomography, several means can be employed to obtain the mean, variance, and correlation structure of the conductivity. For example, one can estimate them on the basis of core samples and well logs if they are available, or one can employ the structure identification approach developed by *Kitanidis and Vomvoris* [1983]. Geophysical survey is an alternative for determining correlation scales [Rea and Knight, 1998], and the traditional aquifer test analysis assuming aquifer homogeneity is a good way to estimate the mean conductivity.

Nevertheless, these statistical parameters are estimates and not known precisely, and measurement errors in pressure heads are inevitable. Therefore the influence of the uncertainty in the statistical parameters and the effects of measurement errors on the estimate by our sequential inverse method are discussed next.

### 4.1. Uncertainty in the Mean and Variance of Hydraulic Conductivity

Without collecting a large number of hydraulic conductivity data sets the mean and variance estimates involve uncertainty. How the uncertainty affects the estimate of hydraulic conductivity by our inverse method needs to be addressed. Several numerical experiments were conducted, and the results show that the uncertainty in the mean conductivity can cause the shift of the mean of our estimated conductivity field. The pattern of heterogeneity remains almost the same. On the other hand, the uncertainty associated with the variance of



**Figure 4.** (a) Contour map of the change of L1 in percentage for different values of  $\Delta\lambda_x/\lambda_x$  (ratio of the change of horizontal correlation scale to the true correlation scale) and  $\Delta\lambda_z/\lambda_z$  (ratio of the change of vertical correlation scale to the true correlation scale). (b) Contour map of the change of L2 in percentage for different values of  $\Delta\lambda_x/\lambda_x$  and  $\Delta\lambda_z/\lambda_z$ .

conductivity has no influence on the final estimate. This is attributed to the fact that our inverse approach relies on the correlation and cross correlation, which do not involve the variance. Specifically, as the variance term appeared on both sides of the system of equations (6) and (8), it is factored out and canceled when solving the equations for weights.

**4.2. Uncertainty in Correlation Scales**

In order to study the effect of the uncertainty in correlation scales we used the previous hypothetical aquifer with the optimal network design as our base case. Then, we conducted many test cases in which the correlation scales in horizontal and vertical directions were either overestimated or underestimated up to 90%. For each test case, the percent changes in the values of the norm L1 and L2 from the base case were computed, and the changes for all the cases were then contoured. Figures 4a and 4b show that our inverse solution is not very sensitive to the uncertainty in correlation scales unless the uncertainty is so large that it completely alters the direction of anisotropy (e.g., the upper left corners of Figures 4a and 4b). This can be attributed to the fact that the correlation structure

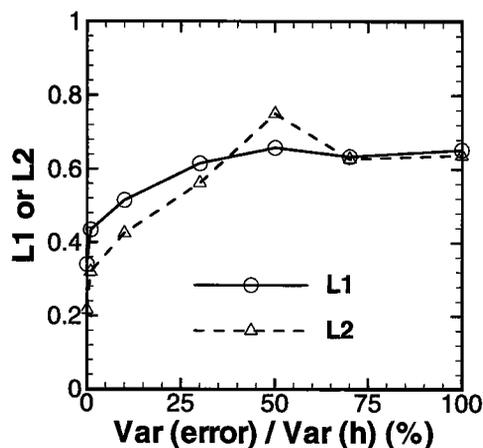
only provides a description of the average size of heterogeneity. Once more point measurements (such as head information from tomography) become available, the impact of the information about the average size of heterogeneity fades out rapidly. The same argument applies to the effect of uncertainty in the shape of the correlation structure (correlation functions).

**4.3. Measurement Errors in Pressure Heads**

To investigate impacts of these errors, we generated random head measurement errors with a zero mean and a variance that is equal to a specified fraction of the head variance for each monitoring location. The head variance was calculated based on the first-order analysis for the given pumping rate. The pressure head measurements at the monitoring locations were then perturbed with these errors and then were used in the inversion. Results of the inversion show that the estimate by our sequential inverse model is very sensitive to the errors: Small measurement errors can lead to erroneous estimates of the conductivity field.

To extract useful information from the head data corrupted with errors, the error variances were added to the diagonal terms of the head covariance matrix, corresponding to the head measurement locations, when the weights for the SLE were sought. Because of the addition of the variance the estimated conductivity field becomes smooth, and head values at the sampling locations do not agree with the observed values.

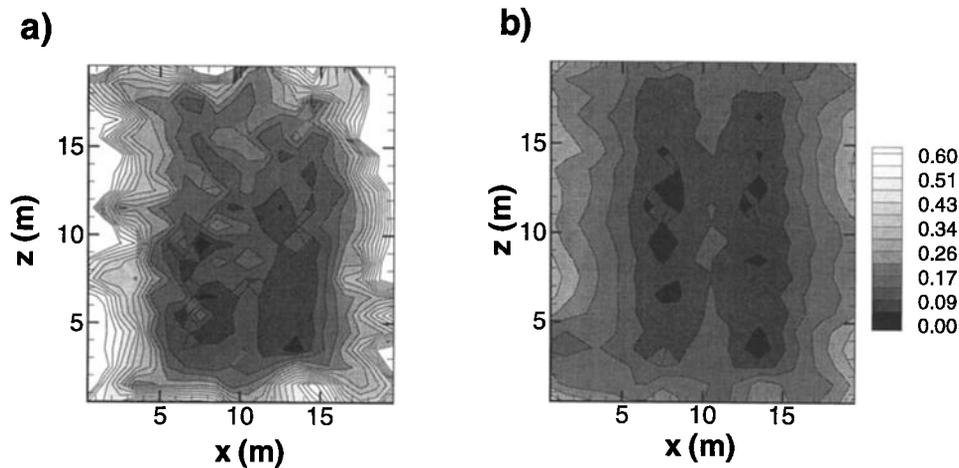
The effects of errors on our estimates using the above approach are shown in Figure 5, where the L1 and L2 norms are plotted as a function of the error. As demonstrated in Figure 5, the norm L1 and L2 grow rapidly when the error increases, which indicates that the estimate of the conductivity field is sensitive to the errors. However, when the variance of the error approaches more than 50% of the head variance, the rate of increase in L1 and L2 declines. This implies that the larger error the data set has, the less useful information the data set contributes. Therefore our inverse method results in a smooth conductivity field that is close to the unconditional mean value.



**Figure 5.** Norm L1 and L2 versus the variance of measurement errors as a percentage of head variance.

**4.4. Effect of Gravel Pack**

Wells are usually gravel-packed over the screen interval. Several numerical experiments were conducted to address the



**Figure 6.** (a) Contour map of the conditional variance of  $\ln K$  (Monte Carlo simulations). (b) Contour map of the conditional variance of  $\ln K$  (our sequential inverse approach).

effects of omitting the gravel pack in the inverse modeling. In the experiments, gravel packs of different uniform conductivity value around the two wells were considered in the forward simulation to produce head measurements. The measured heads were then used in the inversion. Results of the inversion show that the influence of the gravel pack depends upon the contrast between the hydraulic conductivity of the backfilled gravel and the mean conductivity of the aquifer system. If the conductivity of the gravel pack is close to the geometric mean of the conductivity of the aquifer, then its effect is negligible. However, if the gravel pack has a conductivity value that is several orders of magnitude greater than the geometric mean of the aquifer, then its impact is significant. Since the gravel pack changes the head distribution, the effect can be minimized by treating it as the head measurement error.

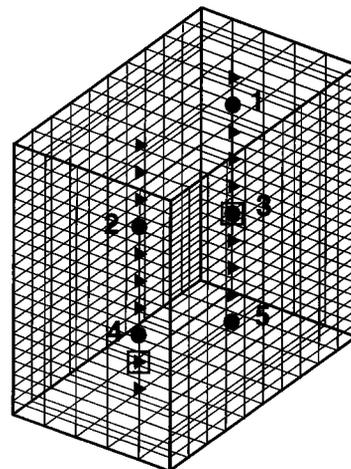
## 5. Monte Carlo Simulations

As mentioned in section 2.2, the conditional covariance functions  $\varepsilon_{hh}$  and cross-covariance functions  $\varepsilon_{hy}$  are approximations. One way to evaluate the accuracy of the approximations is to compare them to those provided by the Monte Carlo simulations. Accordingly, we conducted Monte Carlo simulations using 30 realizations of  $f$  fields. For each realization, five steady state pressure head data sets were produced on the basis of the optimum design of hydraulic tomography as discussed in section 3. In addition, two  $f$  measurements, one from each well, were included as primary information. Then, our proposed sequential inverse approach was employed. The difference between our estimate and the true conductivity field for each realization at each element was then accumulated to determine the conditional variance at each element.

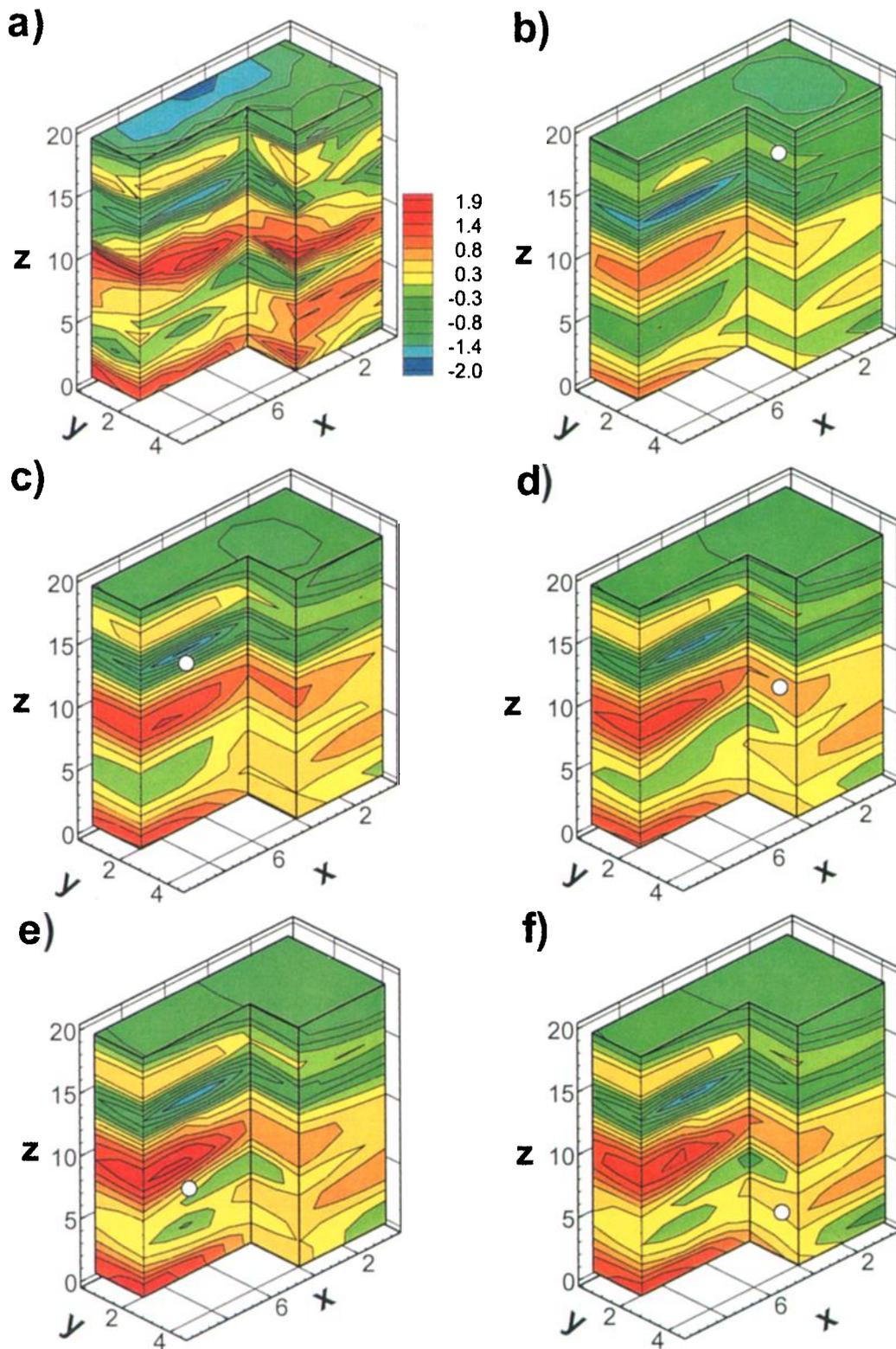
The spatial distribution of the resulting conditional variance is shown in Figure 6a. Compared to the conditional variance of  $f$  calculated using the linear approximation (Figure 6b), the conditional variances obtained from Monte Carlo simulations are larger. Nonetheless, these two conditional variance maps show a similar pattern. That is, lower variances of  $f$  occur at locations along the wells where either pressure head or conductivity was measured. Such a result indicates that significant improvements, due to primary or secondary information, are limited to the vicinity of the measurement locations.

## 6. Three-Dimensional Case Example

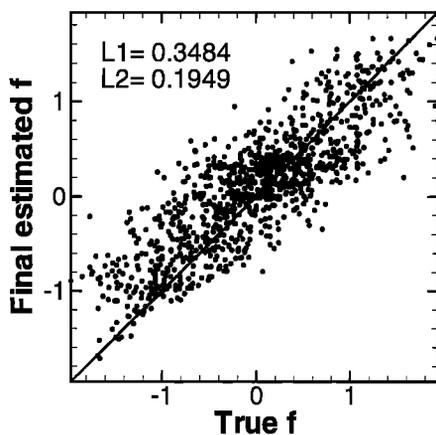
To demonstrate the robustness of the sequential iterative approach for real-world problems, it was applied to hydraulic tomography in a three-dimensional hypothetical aquifer that had dimensions of  $10 \text{ m} \times 5 \text{ m} \times 20 \text{ m}$ . The aquifer was discretized into 1000 elements with dimensions of  $1 \text{ m} \times 1 \text{ m} \times 1 \text{ m}$ . Four sides of the aquifer were constant hydraulic head boundaries with a prescribed value of 80 m while the top and bottom of the aquifer were no-flow boundaries (Figure 7). We assumed that the heterogeneous hydraulic conductivity field had a horizontal correlation scale of 12 m and a vertical correlation scale of 4 m. We further assumed that the geometric mean of the hydraulic conductivity was 0.3452 m/h with a variance of  $\ln K$  equal to 0.625. With these assumed parameters the conductivity value for each element was generated using the spectral method [Gutjahr, 1989] (Plate 1a). On the basis of the optimum-sampling scheme discussed in section 3, a total of 20 pressure measurement locations were used. Two



**Figure 7.** Schematic diagram of hydraulic tomography for a three-dimensional aquifer. Right triangles indicate monitoring locations, circles represent pumping locations (the numbers next to them indicate the order of pumping), and squares represent  $f$  measurement locations.



**Plate 1.** (a) The synthetic true  $f$  field, (b) estimated  $f$  field using data generated by pumping at location 1, (c) estimated  $f$  field using data generated by pumping at locations 1 and 2, sequentially, (d) estimated  $f$  field using data generated by pumping at locations 1, 2, and 3, sequentially, (e) estimated  $f$  field using data generated by pumping at locations 1, 2, 3, and 4, sequentially, and (f) estimated  $f$  field using data generated by pumping at locations 1, 2, 3, 4, and 5, sequentially.



**Figure 8.** Scatterplot of estimated  $f$  field using five data sets sequentially versus true  $f$  field.

conductivity measurements were taken on each of the wells located in the synthetic flow domain (Figure 7). A three-dimensional steady state flow field created by pumping at a selected interval, with a discharge of  $20 \text{ m}^3/\text{h}$ , was then simulated, and the head responses at other intervals were monitored. By sequentially pumping at five different vertical locations (Figure 7), five pressure/discharge data sets were obtained.

For each set of data the SLE was used to determine the conditional effective hydraulic conductivity field. The field obtained from this set of head measurements was used as prior information for the next estimation of the conductivity field, using the next set of pressure/discharge data. This procedure was performed sequentially.

Plate 1b shows the  $f$  estimates based on only the head data set created by pumping at location 1. Plate 1c shows the  $f$  estimates when the head data set created by pumping at location 2 was included. Accordingly, Plate 1f shows the final  $f$  estimates when the fifth data set was included. As illustrated in Plates 1a–1f, the major features of the heterogeneity are captured in the first sequence of the inversion. By incorporating the secondary information sequentially in the inversion more details of heterogeneity are revealed, and the estimate field increasingly resembles the true one. A scatterplot of the final estimated  $f$  versus the true  $f$  values along with the two statistical norms L1 and L2 is displayed in Figure 8.

## 7. Conclusion

The sequential inverse approach using data yielded from hydraulic tomography is a promising tool for characterizing aquifer heterogeneity. By using the secondary information sequentially, the size of the covariance matrix in our inverse approach remains small, so that the matrix equations can be solved with ease. Thus inversion of the large amount of secondary information collected during hydraulic tomography becomes feasible. Compared to the results yielded from the Monte Carlo simulations, the residual variance produced from our sequential approach reflects the pattern of the conditional variance.

Results of our numerical experiments show that the hydraulic tomography can be most effective if the horizontal separation distance between wells is set to be half of the horizontal correlation scale. The vertical interval between two pressure

monitoring locations should be no more than half of the vertical correlation scale. The optimal number of pumping locations is equal to the ratio of the aquifer depth to the vertical correlation scale. The pumping rate has no effect on the estimate.

Our analysis also leads to the conclusions that the uncertainty in the input variance for our inverse model has no influence on the estimates. Similarly, the uncertainty in correlation scales has no significant effect on the estimate unless the correlation scales are extremely underestimated or overestimated. Abundant secondary information (such as pressure head) in space can greatly reduce the effect caused by inaccurate knowledge of the correlation structure. If large measurement errors associated with pressure head exist, our inverse approach yields a smoother estimate than that obtained from the error-free data, reflecting the fact that less information is extracted from the measurements. Consequently, accurate measurements of the secondary information are needed to make hydraulic tomography successful.

Finally, hydraulic tomography appears to be a promising field technique for providing abundant secondary information for characterizing aquifer heterogeneity. Using our sequential inverse model, hydraulic tomography can reveal more detailed aquifer heterogeneity than classical aquifer tests.

**Acknowledgments.** This research is funded in part by a DOE EMSP96 grant through Sandia National Laboratories (contract AV-0655#1) and in part by an EPA grant R-827114-01-0.

## References

- Butler, J. J., Jr., and W. Z. Liu, Pumping tests in nonuniform aquifers: The radially asymmetric case, *Water Resour. Res.*, 29(2), 259–269, 1993.
- Butler, J. J., Jr., C. D. McElwee, and G. C. Bohling, Pumping tests in networks of multilevel sampling wells: Motivation and Methodology, *Water Resour. Res.*, 35(11), 3553–3560, 1999.
- Dettinger, M. D., and J. L. Wilson, First order analysis of uncertainty in numerical models of groundwater flow, 1, Mathematical development, *Water Resour. Res.*, 17(1), 149–161, 1981.
- Gottlieb, J., and P. Dietrich, Identification of the permeability distribution in soil by hydraulic tomography, *Inverse Probl.*, 11, 353–360, 1995.
- Gutjahr, A., Fast Fourier transforms for random field generation, New Mexico Tech project report, 106 pp., Socorro, 1989.
- Hanna, S., and T.-C. J. Yeh, Estimation of co-conditional moments of transmissivity, hydraulic head, and velocity fields, *Adv. Water Resour.*, 22, 87–93, 1998.
- Hoeksema, R. J., and P. K. Kitanidis, An application of the geostatistical approach to the inverse problem in two-dimensional groundwater modeling, *Water Resour. Res.*, 20(7), 1003–1020, 1984.
- Hughson, D. L., and T.-C. J. Yeh, A geostatistically based inverse model for three-dimensional variably saturated flow, *Stochastic Hydrol. Hydraul.*, 12(5), 285–298, 1998.
- Hughson, D. L., and T.-C. J. Yeh, An inverse model for three-dimensional flow in variably saturated porous media, *Water Resour. Res.*, 36(4), 829–839, 2000.
- Kitanidis, P. K., Comment on “A reassessment of the groundwater inverse problem” by D. McLaughlin and L. R. Townley, *Water Resour. Res.*, 33(9), 2199–2202, 1997.
- Kitanidis, P. K., and E. G. Vomvoris, A geostatistical approach to the inverse problem in groundwater modeling and one-dimensional simulations, *Water Resour. Res.*, 19(3), 677–690, 1983.
- Li, B., and T.-C. J. Yeh, Sensitivity and moment analysis of head in variably saturated regimes, *Adv. Water Resour.*, 21, 477–485, 1998.
- Li, B., and T.-C. J. Yeh, Cokriging estimation of the conductivity field under variably saturated flow conditions, *Water Resour. Res.*, 35(12), 3663–3674, 1999.
- Rea, J., and R. Knight, Geostatistical analysis of ground-penetrating

- radar data: A means of describing spatial variation in the subsurface, *Water Resour. Res.*, *34*(3), 329–339, 1998.
- Sun, N.-Z., and W. W.-G. Yeh, A stochastic inverse solution for transient groundwater flow: Parameter identification and reliability analysis, *Water Resour. Res.*, *28*(12), 3269–3280, 1992.
- Sykes, J.-F., J. L. Wilson, and R. W. Andrews, Sensitivity analysis of steady state groundwater flow using adjoint operators, *Water Resour. Res.*, *21*(3), 359–371, 1985.
- van Genuchten, M. T., A closed-form equation for predicting the hydraulic conductivity of unsaturated soils, *Soil Sci. Soc. Am. J.*, *44*, 892–898, 1980.
- Vargas-Guzman, A. J., and T.-C. J. Yeh, Sequential kriging and cokriging: Two powerful geostatistical approaches, *Stochastic Environ. Res. Risk Assess.*, *13*, 416–435, 1999.
- Yeh, T.-C. J., Stochastic modeling of groundwater flow and solute transport in aquifers, *J. Hydrol. Processes*, *6*, 369–395, 1992.
- Yeh, T.-C. J., Scale issues of heterogeneity in vadose-zone hydrology, in *Scale Dependence and Scale Invariance in Hydrology*, edited by G. Sposito, Cambridge Univ. Press, New York, 1998.
- Yeh, T.-C. J., A. L. Gutjahr, and M. Jin, An iterative cokriging-like technique for groundwater flow modeling, *Ground Water*, *33*(1), 33–41, 1995.
- Yeh, T.-C., M. Jin, and S. Hanna, An iterative stochastic inverse method: Conditional effective transmissivity and hydraulic head fields, *Water Resour. Res.*, *32*(1), 85–92, 1996.
- Yeh, W. W.-G., Review of parameter identification procedures in groundwater hydrology: The inverse problem, *Water Resour. Res.*, *22*(1), 95–108, 1986.
- Zhang, J., and T.-C. J. Yeh, An iterative geostatistical inverse method for steady flow in the vadose zone, *Water Resour. Res.*, *33*(1), 63–71, 1997.
- 
- S. Liu and T.-C. J. Yeh, Department of Hydrology and Water Resources, University of Arizona, John W. Harshbarger Building, P. O. Box 210011, Tucson, AZ 85721-0011. (ybiem@mac.hwr.arizona.edu)

(Received October 12, 1999; revised April 18, 2000; accepted April 18, 2000.)