Inverse modeling of partitioning interwell tracer tests: A streamline approach

Akhil Datta-Gupta and Seongsik Yoon
Department of Petroleum Engineering, Texas A&M University, College Station, Texas, USA

D. W. Vasco
Earth Sciences Division, Berkeley National Laboratory, Berkeley, California, USA

Gary A. Pope
Petroleum and Geosystems Engineering, University of Texas at Austin, Austin, Texas, USA

Received 9 April 2001; revised 12 October 2001; accepted 12 October 2001; published 20 June 2002.

[1] Identifying the location and distribution of nonaqueous phase liquid (NAPL) in the subsurface constitutes a vital step in the design and implementation of aquifer remediation schemes. In recent years, partitioning interwell tracer tests (PITT) have gained increasing popularity as a means to characterize NAPL saturation distribution in situ. In this method a suite of conservative and partitioning tracers are injected into the contaminated site. The chromatographic separation between the conservative and the partitioning tracers can be used to infer NAPL saturation distribution. The conventional approach to the analysis of the tracer response uses a first-order method of moments to compute average NAPL saturation in the tracer swept regions and cannot provide detailed spatial distribution of the NAPL. We propose a computationally efficient streamline-based inverse method for analyzing partitioning interwell tracer tests to estimate three dimensional spatial variation of NAPL saturation in the subsurface. Our approach is based on an analogy between streamlines and seismic ray tracing and relies on an analytic sensitivity computation method that yields sensitivities of the partitioning tracer response to subsurface parameters such as porosity, hydraulic conductivity, and NAPL saturation in a single streamline simulation. The inversion of tracer response is carried out in a manner analogous to seismic waveform inversion whereby we first match the “first arrival” followed by matching the “amplitudes” of the tracer response. The power and utility of the method is illustrated using synthetic and field applications. The field example is from the Hill Airforce Base, Utah, where partitioning tracer tests were conducted in an isolated test cell with 4 injection wells, 3 extraction wells, and 12 multilevel samplers. Tracer responses from 51 sampling locations are analyzed to determine hydraulic conductivity variations and NAPL saturation distribution in the test cell. Finally, a performance comparison with simulated annealing shows that our proposed approach is faster by 3 orders of magnitude.

INDEX TERMS: 1832 Hydrology: Groundwater transport; 1829 Hydrology: Groundwater hydrology; KEYWORDS: streamline simulation, parameter estimation, partitioning tracer tests, sensitivity computations, inverse modeling

1. Introduction

[2] It is recognized that the presence of nonaqueous phase liquids (NAPLs) poses a significant impediment to aquifer restoration. In order for any remediation technique to be successful it is essential that the NAPL distribution be properly characterized. Partitioning tracer tests are a promising technique for characterizing NAPL distribution in the subsurface because of accessibility to large volumes of the contaminant. During partitioning interwell tracer tests a suite of tracers with a range of NAPL-water partitioning coefficients are injected into the subsurface and are recovered down gradient at the extraction wells. A conservative or nonpartitioning tracer is also injected during the test. Because of the presence of NAPL, partitioning tracers are retarded compared to the nonpartitioning tracer. The chromatographic separation between these tracers is utilized to estimate average NAPL saturation in the subsurface [Jin et al., 1995]. When the tracer is sampled at multiple vertical and areal locations, then the tracer tests can be used to estimate the three-dimensional (3-D) distribution of NAPL saturation in the subsurface [Annable et al., 1998]. Such information is of obvious importance in the design and implementation of appropriate remediation schemes.

[3] Characterizing the spatial distribution of NAPL saturation from partitioning tracer tests would typically require the solution of an inverse problem. Such inverse problems arise in many fields of science and engineering. Several reviews of the application of inverse methods to ground-
water problems can be found in the literature [Yeh, 1986; McLaughlin and Townley, 1996]. A major focus of the parameter estimation problems in the groundwater literature has been estimation of hydraulic conductivities based on point measurements of conductivities and steady state or transient head response [Kitanidis and Vomvoris, 1983; Carrera and Neuman, 1986; Rubin and Dagan, 1987; McLaughlin and Wood, 1988]. In recent years, inverse problems that utilize solute concentration response have received increased attention in the literature [Woodbury and Sudicky, 1992; Deng et al., 1993; Hyndman et al., 1994; Harvey and Gorelick, 1995; Ezzedine and Rubin, 1996; Medina and Carrera, 1996; Rubin and Ezzedine, 1997; Vasco and Datta-Gupta, 1999]. Computational efforts associated with such inverse modeling still remain a significant factor that deters the routine use of concentration data [Anderman and Hill, 1999]. Most of the work on inverse modeling associated with tracer data has been limited to estimation of hydraulic conductivities, porosity, and transport parameters such as dispersivities and molecular diffusion.

Inverse problems associated with estimation of the spatial distribution of NAPL saturation still remain relatively unexplored. Previous efforts toward estimating NAPL saturation using partitioning interwell tracer tests have mostly utilized the method of moments [Jin et al., 1995; Wilson and Mackay, 1995; Annable, 1998]. Such approaches are well suited to estimate the average NAPL saturation in the tracer swept region but cannot determine the spatial distribution of NAPL saturation. James et al. [1997] introduced a stochastic approach to estimate spatial distribution of temporal moments and associated covariance relations for both conservative and partitioning tracers. An optimal estimation algorithm was then used for determining NAPL saturation distribution and data conditioning. Subsequent work [James et al., 1999] utilized a nonlinear Gauss-Newton search algorithm and adjoint sensitivity calculations for estimation of NAPL distribution. More recently, Zhang and Graham [2001] proposed an extended Kalman filter approach to estimate spatially distributed residual saturation of NAPL. The stochastic approaches require prior knowledge of the spatial correlation structure for hydraulic conductivity and NAPL saturation. The computational costs of these methods often limit large-scale three-dimensional field applications.

The formulation of inverse problems associated with tracer tests typically requires the computation of sensitivities of concentration to changes in model parameters. That is, we must compute the change in concentration response resulting from a small perturbation in subsurface properties such as permeability, porosity, or fluid saturation. The computation of such sensitivities can be classified into three broad categories: perturbation approaches, direct algorithms, and the adjoint state methods. The relative merits of these methods have been discussed in the literature [Yeh, 1986]. The calculations of these sensitivities can constitute a major part of the computational efforts associated with inverse methods.

In this paper we propose a streamline-based inverse approach for estimating spatial distribution of NAPL saturation using partitioning interwell tracer tests. Streamline models can be advantageous in two ways. First, the streamline simulator can serve as an efficient “forward” model for the inverse problem [Datta-Gupta and King, 1995; Crane and Blunt, 1999]. Second, and more importantly, parameter sensitivities can be formulated as one-dimensional integrals of analytic function along streamlines. The computation of sensitivity for all model parameters then requires a single simulation run. The sensitivity computations exploit the analogy between streamlines and seismic ray tracing, based on the observation that the streamline transport equations can be cast in the form of an Eikonal equation, the governing equation for travel time tomography [Vasco and Datta-Gupta, 1999, 2001]. This allows us to use efficient techniques from inverse theory to match both conservative and partitioning tracer responses. Inversion of tracer response is carried out in a manner analogous to seismic waveform inversion [Zhou et al., 1995]. This involves first matching the breakthrough or “arrival time” of the tracer response followed by matching of “amplitudes” of the tracer response, that is, the full tracer concentration history. Such a two-step procedure makes the solution less sensitive to the choice of initial model. An additional feature of the method is that it prevents the solution from prematurely matching secondary peaks [Vasco and Datta-Gupta, 1999]. This is particularly important in field applications where the tracer response is frequently characterized by multiple peaks. We demonstrate the utility of our methodology using both synthetic and field examples.

2. Methodology

In this section, we discuss our theoretical and computational framework for the analysis and inversion of partitioning interwell tracer tests using the streamline approach. The principal components are forward modeling of the tracer response, computation of analytic sensitivities of tracer response to subsurface properties, and finally, history matching or data inversion.

2.1. Forward Modeling of Tracer Transport: Streamline Approach

Forward modeling relates unknown parameters, such as hydraulic conductivity, porosity, and NAPL saturation to the tracer response at an observation well. Because tracers are often injected as a finite slug in small quantities, avoiding numerical dispersion in tracer transport modeling is a major concern. Computational burden associated with repeated forward calculations during inverse modeling is another important aspect. To ensure accuracy and efficiency of the computations, we have used a three-dimensional multiphase streamline model for flow and transport calculations [Datta-Gupta and King, 1995].

Streamline models approximate 3-D fluid flow calculations by a sum of 1-D solutions along streamlines. The choice of streamline directions for the 1-D calculations makes the approach extremely effective for modeling convection-dominated flows in the presence of strong heterogeneity. The details of streamline simulation can be found elsewhere [King and Datta-Gupta, 1998; Crane and Blunt, 1999]. Briefly, in this approach we first compute the pressure or head distribution using a finite difference solution to the conservation equations. The velocity field is obtained using Darcy’s law. A key step is streamline
simulation is the decoupling of flow and transport by a coordinate transformation from the physical space to one following flow directions. This is accomplished by defining a streamline “time of flight” as follows [Datta-Gupta and King, 1995]:

$$\tau(\psi) = \int_\psi^1 \frac{1}{|v(x)|} \, dr.$$  \hspace{1cm} (1)

Thus the time of flight is simply the travel time of a neutral tracer along a streamline $\psi$. In (1), $r$ is the distance along the streamline and $x$ refers to the spatial coordinates. In this paper, we will exploit an analogy between streamlines and seismic ray tracing to utilize efficient techniques from geophysical inverse theory. To facilitate this analogy, we will rewrite the time of flight in terms of a “slowness” commonly used in ray theory in seismology [Nolet, 1987]. The “slowness” is defined as the reciprocal of velocity as follows:

$$s(x) = \frac{1}{|v(x)|} = \frac{n(x)}{K(x)|\nabla \phi(x)|},$$  \hspace{1cm} (2)

where we have used Darcy’s law for the interstitial velocity $v$ and $n$ is the porosity, $K$ is the hydraulic conductivity, and $\phi$ is the piezometric head or pressure. The streamline time of flight can now be written as

$$\tau(\psi) = \int_\psi^1 s(x) \, dr.$$  \hspace{1cm} (3)

Consider the convective transport of a neutral tracer. The conservation equation is given by

$$\frac{\partial C(x,t)}{\partial t} + v \cdot \nabla C(x,t) = 0.$$  \hspace{1cm} (4)

where $C$ represents the tracer concentration. We can rewrite (4) in the streamline time of flight coordinates using the operator identity [Datta-Gupta and King, 1995]

$$v \cdot \nabla = \frac{\partial}{\partial \tau}.$$  \hspace{1cm} (5)

Physically, we have now moved to a coordinate system where all streamlines are straight lines and the distance is measured in units of $\tau$. The coordinate transformation reduces the multidimensional transport equation into a series of one-dimensional equations along streamlines,

$$\frac{\partial C(\tau,t)}{\partial t} + \frac{\partial C(\tau,t)}{\partial \tau} = 0.$$  \hspace{1cm} (6)

The tracer response at a producing well can be obtained by simply integrating the contributions of individual streamlines reaching the producer [Datta-Gupta and King, 1995],

$$C(t) = \int C_0\left[t - \tau(\psi)\right] d\psi = \int C_0\left[t - \int s(x) \, dr\right] d\psi.$$  \hspace{1cm} (7)

where $C_0$ is the tracer concentration at the injection well. If we include longitudinal dispersion along streamlines, then the tracer concentration at the producing well will be given by [Gelhar and Collins, 1971]

$$C(t) = \exp\left[-\frac{(t - \tau(\psi))^2}{4(\alpha \omega)}\right] \frac{d\psi}{\sqrt{4\alpha \omega}}.$$  \hspace{1cm} (8)

where $\alpha$ is longitudinal dispersivity and $\omega = \int |v(x)|^2$. During partitioning interwell tracer tests the retardation of partitioning tracers in the presence of NAPL saturation can simply be expressed as an increase in travel time along streamlines. This, in turn, results in an increased slowness as follows [Jin et al., 1995]:

$$s(x) = \frac{1}{|v(x)|} (S_w + K_N S_{Nw}) = \frac{n(x)}{K(x)|\nabla \phi(x)|} (S_w + K_N S_{Nw}).$$  \hspace{1cm} (9)

where $S_w$ and $S_{Nw}$ denote water and NAPL saturation and $K_N$ is the partitioning coefficient of tracer defined as the ratio of tracer concentration in the NAPL phase to that in the water phase. Notice that when the tracer has equal affinity toward water and NAPL ($K_N = 1$), the tracer response will be insensitive to NAPL saturation as one would expect and (9) reverts back to (2) for single-phase tracer transport. If the NAPL is mobile, the impact of NAPL saturation on the hydraulic conductivity can be accounted for through the use of appropriate relative permeability functions [Vasco and Datta-Gupta, 2001].

### 2.2. Analytic Sensitivity Computations

Sensitivity calculations constitute a critical aspect of inverse modeling. By sensitivity, we mean the partial derivative of the tracer response with respect to model parameters such as hydraulic conductivity, porosity, and NAPL saturation. Although several methods are available for computing sensitivities, for example, numerical perturbation methods, sensitivity equation methods, or adjoint state methods [Yeh, 1986; Sun and Yeh, 1990], they are somewhat limited by their computational costs and the degree of complexity required for their implementation. The streamline approach provides an extremely efficient means for computing parameter sensitivities using a single forward simulation. The sensitivities can be analytically computed and only require evaluation of one-dimensional integrals along streamlines. The sensitivity computations exploit the analogy between streamlines and seismic ray tracing since the streamline transport equations can be cast in the form of an Eikonal equation, the governing equation for travel time tomography [Vasco and Datta-Gupta, 1999].

Consider a small perturbation in a subsurface property (for example, hydraulic conductivity or saturation) from an initial model and the resulting changes in slowness and the tracer concentration in a producing well,

$$s(x) = s^0(x) + \delta s(x),$$

$$C(t) = C^0(t) + \delta C(t).$$  \hspace{1cm} (10)

In (10), $s^0$ and $C^0$ refer to the slowness and the tracer concentration, respectively, for the initial model. If we assume that the streamlines do not shift as a result of a small perturbation in the subsurface property, we can then
compute the changes in tracer time of flight along a streamline as follows:

$$
\delta \tau(\psi) = \int_\psi \delta s(\mathbf{r}) d\psi,
$$

(11a)

and the resulting changes in the tracer concentration at a producing well will now be given by a Taylor series expansion of (7)

$$
\delta C(t) = - \int \hat{C}_0(t - \tau(\psi)) \delta \tau(\psi) d\psi
= - \int \hat{C}_0 \left[ t - s^0(\mathbf{r}) d\psi \right] \left[ \delta s(\mathbf{r}) d\psi \right].
$$

(11b)

In (11b), $\hat{C}_0$ is the time derivative of the injection concentration history and acts as a weighting term for the trajectories converging at the producing well. Because the slowness in (11b) is a composite response as given in (9), its variation can be written as

$$
\delta s(\mathbf{r}) = \frac{\partial s(\mathbf{r})}{\partial K} \delta K(\mathbf{r}) + \frac{\partial s(\mathbf{r})}{\partial n} \delta n(\mathbf{r}) + \frac{\partial s(\mathbf{r})}{\partial S_w} \delta S_w
$$

(12)

We can compute the partial derivatives as follows:

$$
\frac{\partial s(\mathbf{r})}{\partial K} = \frac{1}{K^2(\mathbf{r})} \left( S_u + K_N S_N \right) \frac{\partial s(\mathbf{r})}{\partial K} + \frac{s(\mathbf{r})}{n(\mathbf{r})} \frac{n(\mathbf{r})}{S_u + K_N S_N},
$$

$$
\frac{\partial s(\mathbf{r})}{\partial n} = \frac{n(\mathbf{r})}{K(\mathbf{r})} \frac{\partial n(\mathbf{r})}{\partial n} \left( S_u + K_N S_N \right),
$$

$$
\frac{\partial s(\mathbf{r})}{\partial S_w} = \frac{1}{K(\mathbf{r})} \frac{\partial S_w}{\partial S_w} \left( 1 - K_N \right) - \delta s(\mathbf{r})(1 - K_N).
$$

[14] Note that in (13) the tracer response will be insensitive to saturation changes for $K_N = 1$. Using (11a)–(13), we can now compute the sensitivities of the streamline time of flight and the tracer concentration history at a producing well with respect to hydraulic conductivity, porosity, and water saturation. It is important to point out that the expressions in (13) only involve quantities that are readily available once we generate the velocity field and define the trajectories in a streamline simulator. Thus, in a single streamline simulation we derive all the sensitivity coefficients required to solve the inverse problem. Figure 1 shows tracer concentration sensitivity at a fixed time for hydraulic conductivity, porosity, and water saturation in a homogeneous quarter five-spot pattern computed using the analytic method. The well configuration consists of an injection well (top left corner) and a producing well (bottom right corner). For comparison purposes, we have also shown results using a numerical perturbation method whereby each parameter is perturbed at a time and the tracer response is recomputed using a forward simulation. The overall agreement between the numerical and streamline-based results supports the general validity of our approach.

### 2.3. Estimating Subsurface Properties: Inverse Modeling

[15] During inverse modeling we want to minimize the differences between the observed tracer responses and the model predictions to estimate unknown parameters. These parameters can be three-dimensional distribution of hydraulic conductivity, porosity, and in the case of partitioning tracer tests, NAPL saturation. Mathematically, the inverse problem can be expressed as the following minimization problem:

$$
\min_m \| \mathbf{d} - \mathbf{g}(\mathbf{m}) \|^2,
$$

(14)

where $\mathbf{d}$ is the data vector with $N$ observations, $\mathbf{g}$ is the forward model, $\mathbf{m}$ is the vector of $M$ parameters, and $\| \|_2$ denotes the Euclidean norm. Because of the nonlinearity between the data and model parameters, we must resort to an iterative procedure for the minimization. Using a first-order Taylor series expansion of $\mathbf{g}(\mathbf{m})$ around $\mathbf{m}^k$, an estimate of the model parameters at the $k$th iteration, we obtain the following relationship between the parameter changes and the model predictions:

$$
\mathbf{g}(\mathbf{m}) = \mathbf{g}(\mathbf{m}^k) + \mathbf{G}^k \delta \mathbf{m},
$$

(15)

where $\delta \mathbf{m}$ is the parameter perturbation at $k$th step and $\mathbf{G}$ is the sensitivity matrix with sensitivity coefficients as entries. For example, $G_{ij}$ denotes the sensitivity of the $i$th observation point with respect to the $j$th parameter and will be given by:

$$
G_{ij} = \frac{\partial d_i}{\partial m_j}.
$$

(16)

For the tracer concentration data the sensitivities with respect to permeability, porosity, and water saturation are computed using (11b). In general, we will have many observations from several wells. The differences between the observed and calculated tracer response at the $k$th iteration step will comprise a data misfit vector given by

$$
\mathbf{e} = \mathbf{d} - \mathbf{g}(\mathbf{m}^k)
$$

(17)

and the data misfit vector can be related to the changes in the model parameter estimates through the following system of equations:

$$
\mathbf{e} = \mathbf{G}^k \delta \mathbf{m}.
$$

(18)

[16] In our approach we follow an iterative minimization procedure by solving (18) at each step or equivalently minimize the linear least squares

$$
\| \mathbf{e} - \mathbf{G}^k \delta \mathbf{m} \|^2 = \sum_{i=1}^{N} \left( e_i - \sum_{j=1}^{M} G_{ij} \delta m_j \right)^2.
$$

(19)

The model parameters are updated at each step during the iterations

$$
\mathbf{m} = \mathbf{m}^k + \delta \mathbf{m}.
$$

(20)

[17] Because of the strong nonlinearity of the inverse problem, we require a large number of iterations to con-
verge to model parameter estimates. In our applications described below, many tens to over a hundred iterations were necessary to find model parameters consistent with the observations. During every iteration the streamlines are updated and the tracer concentration history is recomputed using a forward simulation. This points to the need for very efficient forward modeling and solution of the linearized inverse problem.

[18] For field-scale applications of inverse modeling, very often we have a large number of unknown parameters and limited measurements. Thus the inverse problem tends to be ill posed [Parker, 1994]. Such ill-posed problems can suffer from nonuniqueness and instability in solution. To circumvent these problems, we augment the linear system of equations (19) by incorporating additional penalty terms, a process known as regularization [Constable et al., 1987; Borchers et al., 1997; Liu and Ball, 1999]. Two common approaches are to include a model norm penalty and a model roughness penalty. The norm penalty ensures that our final model is not significantly different from our prior model. This makes physical sense because typically our initial model already incorporates sufficient geologic and other prior information. The roughness penalty accounts for the fact that tracer data is an integrated response and is best suited to resolve large-scale trends rather than small-scale fluctua-

Figure 1. Sensitivity comparison between (left) streamline-based analytic approach and (right) numerical perturbation approach for tracer response in a homogeneous quarter five-spot pattern at a fixed time. See color version of this figure at back of this issue.
tions. The penalized objective function to be minimized is given by
\[ \| \mathbf{e} - \mathbf{Gm} \|^2 + \beta_1^2 \| \mathbf{Lm} \|^2 + \beta_2^2 \| \mathbf{m} \|^2, \]  
(21)
where \( \beta \) values are the weighting factors for the model roughness and norm penalties and \( \mathbf{L} \) is a spatial difference operator, typically the second spatial derivative of parameters, measuring the model roughness [Menke, 1989].

We will solve for (21) using an efficient singular value decomposition (SVD) algorithm. Because a form of norm minimization will be implicit in our SVD solution of (21) [Parker, 1994], we shall only retain the roughness penalty term in the following formulation. The necessary equations for a minimum of (21) may then be written in the following form:
\[ [\mathbf{G}^T \mathbf{G} + \mathbf{L}^T \mathbf{L}] \mathbf{\delta m} = \mathbf{G}^T \mathbf{e}. \]  
(22)
We define an \((N + M)\) by \(M\) augmented matrix \( \mathbf{\Gamma} \)
\[ \mathbf{\Gamma} = \begin{bmatrix} \mathbf{G} \\ \mathbf{L} \end{bmatrix}, \]  
(23)
as well as the augmented data vector \( \mathbf{T} \)
\[ \mathbf{T} = \begin{bmatrix} \mathbf{e} \\ \mathbf{0} \end{bmatrix}. \]  
(24)
We can now write (22) as
\[ \mathbf{\Gamma}^T \mathbf{\Gamma} \mathbf{\delta m} = \mathbf{\Gamma}^T \mathbf{T}. \]  
(25)
The solution of equation (25) provides model parameter estimates that minimize the penalized misfit function (21). Our estimates are based upon a SVD of \( \mathbf{\Gamma} \), that is, by the representation of \( \mathbf{\Gamma} \) as the following product:
\[ \mathbf{\Gamma} = \mathbf{UAV}^T, \]  
(26)
where \( \mathbf{A} \) is a diagonal matrix and \( \mathbf{U} \) and \( \mathbf{V} \) satisfy the orthogonality relations \( \mathbf{U}^T \mathbf{U} = \mathbf{I} \) and \( \mathbf{V}^T \mathbf{V} = \mathbf{I} \). There are many texts describing the utility of the SVD [Noble and Daniel, 1977; Golub and Van Loan, 1996] and its application to inverse modeling [Menke, 1989; Parker, 1994]. We only wish to emphasize that there are efficient algorithms for constructing a partial SVD, based on a three-term recursion first proposed by Lanczos [1950]. This technique is applicable to very large inverse problems [Vasco et al., 1999b]. Once we have constructed an SVD of \( \mathbf{\Gamma} \), we may use the properties of the matrices \( \mathbf{U}, \mathbf{V}, \) and \( \mathbf{A} \) to construct an approximate or “generalized” inverse of \( \mathbf{\Gamma} \). Details may be found in the works of Menke [1989] or Parker [1994]. It is critical when forming the generalized inverse to truncate the representation (26), eliminating those singular vectors, columns of \( \mathbf{U} \) and \( \mathbf{V} \), associated with elements of \( \mathbf{A} \) close to zero. Say that there are \( p \) significant values that are greater than the cutoff. We may then write a truncated representation in place of (26)
\[ \mathbf{\Gamma} = \mathbf{U}_p \mathbf{A}_p \mathbf{V}_p^T, \]  
(27)
where the subscript denotes that only \( p \) columns of \( \mathbf{U} \) and \( \mathbf{V} \) are retained. On the basis of the truncated representation we can now construct an approximate, or generalized inverse as follows:
\[ \mathbf{\Gamma}^+ = \mathbf{V}_p \mathbf{A}_p^{-1} \mathbf{U}_p^T. \]  
(28)
Thus our model parameter estimates are given by
\[ \mathbf{\delta m} = \mathbf{V}_p \mathbf{A}_p^{-1} \mathbf{U}_p^T \mathbf{T}. \]  
(29)

2.4. SVD and Model Assessment

An important aspect of inverse modeling is the assessment of the model parameter estimates. Conventionally, this involves the construction of the model parameter covariance matrix, which contains the uncertainties associated with our estimates of permeability and saturation. The diagonal elements of the covariance matrix are the model parameter variances while the off-diagonal terms measure the parameter covariances. As shown in Appendix A, for equations normalized by the standard errors of the data, we may write the model parameter covariance matrix in terms of the singular value decomposition
\[ \mathbf{C}_m = \mathbf{V}_p \mathbf{\Omega} \mathbf{V}_p^T, \]  
(30)
where \( \mathbf{\Omega} \) is given by
\[ \mathbf{\Omega} = \mathbf{A}_p^{-2} - \mathbf{A}_p^{-1} \mathbf{U}_{2p} \mathbf{U}_{2p} \mathbf{A}_p^{-1}, \]  
(31)
and the matrix \( \mathbf{U}_{2p} \) contains the last \( M \) rows of \( \mathbf{U}_p \), corresponding to the roughness penalty, as noted in Appendix A.

In addition to model parameter uncertainty, some measure of spatial resolution can provide additional insight to the parameter estimation problem. That is, we are attempting to determine a spatially varying field of permeability, porosity, and saturation based on an integrated measure such as the tracer response at the producing wells. The best we can hope to recover is some volumetric average of the property rather than point estimates. We also have potential trade-offs between the various classes of parameters, such as between porosity and permeability estimates in a particular volume. Consideration of the spatial scale at which we can resolve variations in particular hydraulic properties leads to the idea of model parameter resolution [Menke, 1989; Vasco et al., 1997; Datta-Gupta et al., 1997]. The resolution matrix provides a measure of the spatial averaging inherent in our inverse modeling. Assuming that there is an actual or true distribution of properties, the resolution matrix relates this true model \( \mathbf{\delta m} \) to our model parameter estimates \( \mathbf{\delta m} \)
\[ \mathbf{\delta m} = \mathbf{R} \mathbf{\delta m}. \]  
(32)
The resolution matrix \( \mathbf{R} \) may be interpreted as a linear filter through which we view the actual spatial distribution of flow properties [Vasco et al., 1997]. The rows of the resolution matrix are averaging coefficients indicating the contribution of various other parameters to our estimate of a property in a given cell. In the ideal case the resolution
matrix would be an identity matrix, and we would resolve the properties of each cell perfectly, with no trade-off between adjacent estimates nor with other classes of parameters. One advantage of the resolution matrix is that it is independent of the data uncertainty and only depends on our sensitivities and the geometry of our experiment. The concept of model parameter resolution is particularly important for inverse modeling based upon hydrological data. In many situations we have a sparse distribution of data from widely spaced boreholes. As shown in Appendix A, we may construct the resolution matrix using quantities already available in the SVD of our augmented sensitivity matrix

\[ R = V_p \Psi V_p^T, \]

where

\[ \Psi = I - A_p^{-1} U_{2p} U_{2p} A_p. \]

As noted above, the matrix \( U_{2p} \) contains the last \( M \) rows, those associated with the roughness penalty, of the matrix \( U_p \).

### 2.5. Two-Step Inversion Procedure

[22] We follow a two-step procedure in determining NAPL saturation based on partitioning tracer tests. First, we invert the conservative tracer response to infer spatial distribution of hydraulic conductivity in the subsurface. Next, we invert the partitioning tracer response to find the spatial distribution of residual NAPL. The underlying assumption here is that the conservative tracer response is primarily governed by the hydraulic conductivity variations, whereas the partitioning tracer response is influenced by both hydraulic conductivity and NAPL distribution [James et al., 1997]. The streamline-based approach that we propose here facilitates the two-step procedure because we can compute sensitivities of tracer response with respect to hydraulic conductivity and NAPL saturation in a single simulation run. Each of the two steps will require several linearized iterations in order to converge to estimates of hydraulic conductivity and saturation variations that are compatible with the observations. Furthermore, it is only on the final iteration of the conservative and partitioning tracer matches that we conduct the model parameter assessment. That is, the assessment is done locally, about the final conductivity and saturation distribution models. The procedure should become clearer in our description of the applications, given next.

### 3. Applications

[23] We first consider a synthetic example and then a field application of the two-step approach. The synthetic example illustrates our procedure for analysis of partitioning tracer tests to characterize NAPL saturation in the subsurface. An application at the Hill Air Force Base, Utah, demonstrates the feasibility of the approach for analyzing partitioning tracer response in field situations.

#### 3.1. A Synthetic Example: Partitioning Tracer Test for a Single-Layer Aquifer

[24] The well configuration for this example is derived from our field application at the Hill Air Force Base discussed later. In particular, we consider a two-dimensional problem, based on a single layer within the original three-dimensional field model. There are four injection wells at the westernmost end and three extraction wells located along the eastern edge of the model (Figure 2). Between the injectors and producers there are 12 samplers. The spatial distribution of permeability, also shown in Figure 2, is generated using a stochastic moving average method. Permeabilities are generally higher toward the southern end with a region of lower permeability in the east central part. The overall permeability variation is approximately two and a half orders of magnitude, from \( \sim 1900 \) to 48,000 mD (1 mD = 9.6 \times 10^{-9} \text{ m/s}). This was the order of variation observed at the Hill test site. Tracer concentrations are monitored in the samplers as well as in the extraction wells, altogether a total of 15 tracer responses. The concentration histories for the conservative tracer are shown in Figure 3 as the solid lines. The dashed curves in Figure 3 indicate the response predicted using a homogeneous background model with a permeability of 10,000 mD. The homogeneous background served as our prior or starting model in this case. Significant discrepancies exist between the “observed” tracer histories and the calculated responses, both in arrival time and amplitude.

[25] We first match the conservative tracer response to map hydraulic conductivity assuming fixed porosity. The inversion of tracer data proceeds in two stages in a manner analogous to seismic waveform fitting [Zhou et al., 1995]. We first match the arrival times of the peak concentrations at the producing wells. Only then, after the peaks have been “lined up,” are the histories themselves matched. Our experience has shown that this two-stage procedure makes the solution more robust, that is, less sensitive to our selection of the prior model. During arrival time matching we solve a system of equations equivalent to (18) relating changes in the peak arrival times at the wells to subsurface property variations

\[ \delta T = S \delta m, \]

where \( S \) is the travel time sensitivity matrix, with elements given by (13). The arrival time inversion is followed by the
amplitude inversion whereby we match the entire tracer history. Starting with the solution produced by the travel time match, we now solve (21) and match the tracer concentration histories directly. As in the work of Vasco and Datta-Gupta [1999], we find that the improvement in parameter estimation over the arrival time inversion is rather minimal. The final permeability distribution (Figure 4) contains the major features of the original permeability field (Figure 2). However, there are discrepancies between the smaller-scale features of these two models. For example, the major low-permeability region in the inversion result is found between the second and third observation points in the second row of samplers, to the southwest of the true minimum permeability (Figure 2). Thus, as expected, we cannot recover the permeability variations on all scales.

[26] In order to discern those features of our model that are resolvable and to quantify model parameter uncertainties, we now conduct an assessment of our estimates. Specifically, we constructed model parameter resolution and covariance matrices corresponding to the solution in Figure 4. The sensitivities are taken from the final amplitude iteration. Using the Lanczos recursion, we completed a singular value decomposition of the augmented sensitivity matrix. The singular values decay rapidly, as indicated in Figure 5. We used a cutoff $\rho = 100$ in our truncated representation and in the generalized inverse used to conduct the model parameter assessment (equations (27) and (28), respectively). The diagonal elements of the resolution matrix $R$ (see equation (32)) are shown in Figure 6a. As stated earlier, a resolution matrix close to the identity matrix signifies well-resolved parameters. Therefore a diagonal element near 1 means that the parameter is well resolved. Conversely, diagonal values near zero imply that the parameter is poorly determined and signifies trade-off with other parameters. In Figure 6a we plot the diagonal elements of the resolution matrix at the position of the block to which they correspond. We see that we obtain the best resolution, near unity in some cases, along the rows of samplers. This makes physical sense because the flow is predominantly along the rows of samplers. Note that there is some offset in this pattern, as well as variations in amplitude, most likely due to the lateral variations in permeability and the corre-
responding deflections in the streamlines. The model parameter standard errors are plotted in Figure 6b. They are calculated using (30) and (31), where the cutoff $p$ is identical to that used to compute the resolution matrix. There is no clear pattern in the spatial distribution of the standard errors. The overall level of standard error is below 1000 mD, which appears high, but it is much lower than the variation in permeability, from $\sim$1900 to 48,000 mD. Thus it seems that we can resolve the variations between the samplers with adequate precision. However, between the rows of samplers the resolution is quite poor. This is evident in our solution (Figure 4) where we are unable to recover the region of lowest permeability in the east central portion of the test cell, between the second and third row of samplers.

The second stage of the inversion entails estimating the distribution of saturation within the aquifer based on the partitioning tracer response. For our synthetic test the distribution of NAPL in the subsurface is correlated with the pattern of heterogeneity in permeability. The spatial variation of NAPL saturation is shown in Figure 7. The NAPL saturation is generally lower to the south of the test region with a peak saturation in the east central portion of the area between the wells. The background permeability distribution used in matching the partitioning tracer was that of Figure 4, the result of inverting the conservative tracer. This permeability distribution was not varied during the ensuing inversion for NAPL saturation. The reference tracer responses, based on the saturation distribution of Figure 7 and the permeability variation of Figure 2, are shown in Figure 8 as the solid lines. Also shown in Figure 8 are the partitioning tracer responses calculated using the inversion result of Figure 4 and a homogeneous NAPL saturation of 7.5 percent. There are notable differences between the reference values and the predictions based upon a uniform NAPL saturation. Figure 8 compares the reference values with those predicted by our inversion result. The final model of saturation variation is shown in Figure 9. Our inversion result reproduces the general features of the synthetic model (Figure 7). In particular, there are lower NAPL saturations to the south and a concentration of higher NAPL saturations in the east central portion of the region. There are also differences between the two models. For example, the peak NAPL concentration in our inversion result is shifted somewhat to the west of the actual maximum.

As in our inversion of the conservative tracer, we conducted a model assessment about the final model of our iterative inversion for NAPL saturation (Figure 9). Again, our method is based on the SVD of the augmented matrix (23). The resulting singular values, diagonal elements of the matrix $A$, are shown in Figure 10. They decay rapidly in amplitude, even in the presence of an amplitude penalty term. On the basis of the singular value spectrum (Figure 10), we used a value of $p = 50$ for the singular value cutoff in the expressions (30), (31), (33), and (34). The diagonal elements of the resolution and covariance matrices are displayed in Figures 11a and 11b, respectively. As for the inversion of the conservative tracer, the best resolved saturation estimates lie along or nearly along the rows of the samplers. The computed error associated with our model parameter estimates is quite low overall, of the order of 1% or less. For the most part the errors are highest to the west.

Figure 6. Model assessment using resolution and covariance analysis: (a) resolution of permeability estimates and (b) variance of permeability estimates.

Figure 7. Reference saturation distribution for the synthetic example. See color version of this figure at back of this issue.
Similarly, the resolution appears to be higher to the west. In general, it appears that we can resolve saturation distribution with acceptable uncertainty in the zones constrained by the rows of samplers. We gain some insight concerning the nature of our model parameter resolution if we examine individual rows of the resolution matrix. Two rows of the resolution matrix $R$, those associated with blocks 25 and 74, are portrayed graphically in Figure 12. The gray scale depicts the degree of averaging between the parameter and surrounding model parameters. For example, our estimate of the saturation in cell 25, which is moderately to poorly resolved, is a lateral average of values between two adjacent samplers. The saturation estimate for block 74 is better resolved. There appears to be little trade-off between our saturation estimate for cell 74 and the saturation estimates of adjacent cells.

3.2. Inversion of Hill Air Force Base Partitioning Interwell Tracer Tests

The details of the partitioning tracer tests conducted at the Hill Air Force Base can be found in the work of Annable et al. [1994, 1998]. The aquifer consists of sand, gravel (with some large cobbles), and clays with a mean permeability of 20 Darcies. The base of the aquifer is defined by an impermeable clay layer. The NAPL, which is lighter than water, is in the form of a plume covering several acres. An isolation test cell was installed for the purpose of evaluating the use of cosolvents as a remediation tool. The cell consisting of a sealed sheet pile barrier system measures $3.5 \times 4.3$ m and extends to a depth of 9.1 m below the ground surface, some 3 m below the confining unit of the aquifer. Multiple tracers were injected using four injection wells at one end of the cell. Tracer responses were measured at three extraction wells at the opposite end and 12 multilevel samplers between the injection and extraction wells as depicted in Figure 13. For our analysis we have used bromide as the conservative tracer and 2,2-dimethyl-3-pentanol (DMP) as the partitioning tracer from the suites of field tracer response. We chose DMP as the partitioning tracer because of its higher parti-
We model the lower portion of the test cell using a 14 × 11 × 10 grid blocks with dimensions of ~0.3 m horizontally and 0.15 m vertically. The choice of the grid was largely dictated by the spacing of multilevel samplers to capture spatial variations between samplers both laterally and vertically. To start with, we assume a uniform distribution of hydraulic conductivity within the test cell equal to the mean permeability of 20 Darcies. In general, the initial model should incorporate all available prior knowledge about the hydraulic conductivity distribution and may be constructed using geostatistical methods [Datta-Gupta et al., 2001]. Tracer responses from the initial model and observed tracer responses at five selected sampling locations are shown in Figure 14. First, conservative tracer responses measured at each sampling location are matched to infer hydraulic conductivity distribution within the test cell. These include the tracer data at 48 multilevel sampling locations and the three extraction wells. We assumed an effective porosity of 0.20 based on the hydraulic tests reported by Annable et al. [1998], and this was kept fixed during the inversion. Figure 15 shows the result after matching the peak arrival times at the same five locations as in Figure 14. In general, matching peak arrival times also result in a substantial improvement in the overall tracer history match as can be seen here. The observed and

Figure 11. Model assessment using resolution and covariance analysis: (a) resolution of saturation estimates and (b) variance of saturation estimates. See color version of this figure at back of this issue.

Figure 12. Spatial averaging associated with saturation estimates for two model cells (grid blocks).

Figure 13. Hill Air Force Base test cell diagram.
calculated peak arrival times at all 51 sampling locations are shown in Figure 16. The computation time for the arrival time match was just 37 s on a Pentium III. The estimated hydraulic conductivity field is shown in Figure 17. Note that we have mapped the hydraulic conductivity distribution under the assumption that the porosity variation is much less compared to that of hydraulic conductivity. In general, the effective porosity may be correlated to hydraulic conductivity, and additional information such as pressure data will be necessary to account for the porosity variation.

Next we invert the partitioning tracer response starting with the hydraulic conductivity field derived from the conservative tracer response and a uniform initial NAPL saturation distribution. If there is some prior knowledge of the spatial variability of NAPL saturation, it can be incorporated at this stage. As discussed before, the norm penalty in (21) attempts to preserve the initial saturation distribution while matching the tracer data. Figure 18 compares the observed and calculated partitioning tracer response at five selected locations based on the initial model. Figure 19 shows the results after inversion. The observed and calculated tracer concentrations are shown in Figure 15.

Table 1. Tracer Partitioning Coefficients for Hill Air Force Base Case

<table>
<thead>
<tr>
<th>Tracer</th>
<th>$K_N$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bromide</td>
<td>0.0</td>
</tr>
<tr>
<td>Ethanol</td>
<td>0.1</td>
</tr>
<tr>
<td>n-Pentanol</td>
<td>1.4</td>
</tr>
<tr>
<td>n-Hexanol</td>
<td>4.6</td>
</tr>
<tr>
<td>2,2-Dimethyl-3-pentanol</td>
<td>12.9</td>
</tr>
</tbody>
</table>

Annable et al. [1994].

Figure 14. Conservative tracer responses at five selected sampling locations for a uniform initial hydraulic conductivity model.

Figure 15. Conservative tracer responses at five selected sampling locations for the final hydraulic conductivity model.

Figure 16. Observed and calculated concentration peak arrival times at all sampling locations before and after inversion for hydraulic conductivity.
lated peak arrival times at all 51 sampling locations after 14 iterations are shown in Figure 20. The final NAPL saturation distribution is shown in Figure 21. The total computation time for this example was 1 min and 34 s on a Pentium III. Our results indicate an average NAPL saturation of \( \approx 6\% \) with higher saturation toward the lower part of the test cell. The average NAPL saturation is slightly higher than the previous estimates of 4.6 to 5.4\% based on the moment analysis [Annable et al., 1998]. The spatial distribution of NAPL appears to be consistent with soil core analysis, indicating higher NAPL saturation clustered toward the lower portion of the cell [Annable et al., 1994]. For comparison purposes, we also calculated tracer

![Figure 17](image1.png)

**Figure 17.** Hydraulic conductivity field estimated from the conservative tracer data for the Hill case (log permeability scale). See color version of this figure at back of this issue.

![Figure 18](image2.png)

**Figure 18.** Partitioning tracer responses at five selected sampling locations for a uniform initial nonaqueous phase liquid (NAPL) saturation model.

![Figure 19](image3.png)

**Figure 19.** Partitioning tracer responses at five selected sampling locations for the final NAPL saturation model.

![Figure 20](image4.png)

**Figure 20.** Observed and calculated concentration peak arrival times at all sampling locations before and after inversion for NAPL saturation.
Mretardation factors defined as the ratio of average travel times for the partitioning tracer and the conservative tracer at sampling locations [Jin et al., 1995; Annable et al., 1998]. A qualitative comparison of our inversion result with the spatial distribution of retardation factors (Figure 22) seems to further validate our results. Streamline distribution for the final model is shown in Figure 23.

[32] The streamline approach allows us to readily calculate the volumetric sweep from the distribution of tracer travel times. Figure 24 shows the time-of-flight contours that reflect the conservative tracer front locations in the test cell at various times. The tracer swept volume at any given time can be calculated by computing the pore volume within the corresponding time of flight contour,

\[ V_s(t) = \sum_i \int d\tau(\psi_i)\theta(t - \tau)q(\psi_i), \tag{36} \]

where \( \theta \) is the Heaviside function and \( q(\psi_i) \) is volumetric flow rate assigned to the streamline \( \psi_i \). Figure 25 shows the volumetric sweep efficiency as a function of injection time where the volumetric sweep efficiency is defined by the ratio of tracer swept volume to the total pore volume. For all practical purposes, the test cell appears to be completely swept after an injection period of 5 days.

[33] A final step in inverse modeling or parameter estimation is an assessment of the solution. As discussed before, we resort to a resolution analysis of our estimates for this purpose. Figure 26 shows the resolution of saturation estimates using the partitioning tracer data. Higher resolution is observed toward the bottom portion of the cell with high NAPL saturations. Also, high resolution is observed along flow paths between injectors and producers where sensitivities are high. Similar observations were also made by James et al. [1997] using a stochastic inversion approach.

[34] One of the major advantages of the streamline approach presented here is its computational efficiency because of the analytic computation of parameter sensitivities. For benchmarking purposes we compared our approach with a gradient-free global optimization technique, simulated annealing [Mauldon et al., 1993; Datta-Gupta et al., 1995]. In simulated annealing, parameter values are perturbed at random by drawing from a predefined probability distribution for each parameter. The method provides a mechanism of “probabilistic hill climbing” that allows the solution to escape from local minimum. Figure 27 shows the NAPL saturation distribution using the
simulated annealing approach. Overall, the NAPL distribution follows similar trends as in Figure 21. However, the computation time required was 15 hours for conservative tracer inversion followed by 16 hours for partitioning tracer inversion. This is an increase of more than 3 orders of magnitude compared to our proposed approach. Table 2 summarizes the computation times.

4. Summary and Conclusions

[35] We have described a streamline-based inversion technique to estimate spatial variation of NAPL saturation using partitioning interwell tracer tests. The streamline approach is fast and also results in improved accuracy because of reduced numerical dispersion and grid orientation effects. The streamline time-of-flight formalism provides us with a straightforward approach for estimating swept volume during tracer tests. Finally, we may apply similar methodology to the inversion of transient pressure data [Vasco et al., 2000] and multiphase flow data [Vasco and Datta-Gupta, 2001]. Thus the methods described here are of general use in solving the inverse problem for flow and transport data.

[36] Primary advantage of our proposed inverse method is its computational efficiency, more than 3 orders of magnitude faster than simulated annealing for the example presented here. This makes our method particularly attractive for analysis of large-scale field tests. The speed of computation can be attributed to our analytical sensitivity calculations using the streamline approach. The streamline approach also allows us to exploit the analogy with seismic ray tracing and use efficient techniques from geophysical inversion. We have adopted an inversion scheme that is analogous to seismic waveform inversion whereby we first match the arrival times followed by amplitudes of the tracer response. This makes the data inversion robust, relatively insensitive to our initial model and also prevents the solution from being trapped in secondary tracer peaks. The latter is especially important for field application because field tracer tests are very often characterized by multipeaked response. We have demonstrated the power and utility of our method through synthetic and field examples. The field example is from the Hill Airforce Base, Utah, where partitioning tracer tests were conducted in an isolated test cell with 4 injection wells, 3 extraction wells, and 12 multilevel samplers. Tracer responses from 51 sampling locations are analyzed to determine hydraulic conductivity variations and NAPL saturation distribution in the test cell. The estimated spatial distribution of NAPL saturation was found to be consistent with soil core analysis.

Appendix A

A1. Resolution Matrix

[37] Our presentation of the resolution matrix begins with the assumption that there is a "true" model perturbation, \( \delta m \), which generates the residuals

\[
\varepsilon = G \delta m. \tag{A1}
\]

If we substitute (A1) for the data component of \( T \) in equation (29), we arrive at
\[ \hat{m} = V_p A_p^{-1} U_p^T G \begin{bmatrix} I_m \end{bmatrix} \hat{m}, \quad (A2) \]

which gives us a relationship between the ‘true’ model and our estimated model. From our SVD we have the representation

\[ \Gamma = \begin{bmatrix} G \\ L \end{bmatrix} = U A V^T. \quad (A3) \]

We can write the columns of the matrix \( U \) in a partitioned form as follows:

\[ U = \begin{bmatrix} U_1 \\ U_2 \end{bmatrix}, \quad (A4) \]

where the first \( N \) rows, those associated with \( U_1 \), correspond to the data constraints while the next \( M \) rows are related to the regularization, the roughness penalty. Hence we have the following representation of \( G \):

\[ G = U_1 A V^T. \quad (A5) \]

Therefore truncating the above representation to only include the \( p \) vectors that are above the singular value cutoff, we obtain the following form for the resolution matrix:

\[ R = V_p A_p^{-1} U_p^T U_1 A_2 V_2^T. \quad (A6) \]

Because the columns of \( U_1 \) only contain \( N \) of the \( N + M \) elements of the singular vectors they are not necessarily orthogonal. However, we may use the orthonormality of the singular vectors to write the product as

\[ U_1^T U_1 = I - U_2^T U_2. \quad (A7) \]

Thus the resolution matrix may be written as

\[ R = V_p \Psi V_p^T, \quad (A8) \]

where

\[ \Psi = I - A_p^{-1} U_2^T U_2 A_p. \quad (A9) \]

The form (A8) should be compared to the situation in which no roughness regularization is included in the inverse modeling [Menke, 1989]

\[ R = V_p V_p^T. \quad (A10) \]

Though the form (A8) is more involved than the case in which there is no roughness penalty we should mention a few points concerning computational efficiency. First, there is an efficient iterative scheme for calculating the singular values, the elements of \( A \), the Lanczos algorithm [Lanczos, 1950]. This approach is practical for very large inverse problems involving of the order of a million equations and hundreds of thousands of unknowns [Vasco et al., 1999b]. The Lanczos algorithm also provides the columns of the matrix \( V_p \) with little additional computational cost. We may obtain the columns of \( U_{2p} \) based on the relationship between the columns of \( V \) and the columns of \( U_2 \)

\[ L v_i = \lambda_i u_{2i}. \quad (A11) \]

Note that \( L \) is the roughness matrix that is extremely sparse, typically containing just a few nonzero elements per row.

### A2. Covariance Matrix

[38] A procedure similar to that followed in our derivation of the resolution matrix can be used to find the form of the covariance matrix in the presence of a roughness penalty. We begin with the observation that our model parameter estimates are linearly related to the residuals as is evident from equation (29). Utilizing the partitioned form for \( U \) given in equation (A4), we have

\[ \hat{m} = V_p A_p^{-1} U_p^T e = \Gamma^T e, \quad (A12) \]

where \( \Gamma^T \) is the generalized inverse. Given this linear relationship between the residuals and the model parameter estimates, we may relate the data covariances \( C_d \) to the model parameter covariances \( C_m \):

\[ C_m = \Gamma^T C_d \Gamma^T. \quad (A13) \]

Substituting our expression for the generalized inverse and assuming that we have normalized the equations by their associated standard error, such that the covariance matrix associated with the data becomes the identity matrix, (A13) becomes

\[ C_m = V_p A_p\psi^{-1} V_p^T. \quad (A14) \]

Alternatively, we may write the model parameter covariance matrix as

\[ C_m = V_p \Omega V_p^T, \quad (A15) \]

where \( \Omega \) is given by

\[ \Omega = A_p^{-2} - A_p^{-1} U_{2p}^T U_{2p} A_p^{-1}. \quad (A16) \]

Similar computational considerations apply to both \( \Omega \) and \( \Psi \) (see above).

### Acknowledgments

This work is based on work supported by the National Science Foundation under grant 9873275. We are grateful to the University of Florida research group, especially Mike Annable, Wendy Graham, and Suresh Rao, for sharing the Hill Air Force Base data.

### References


Figure 1. Sensitivity comparison between (left) streamline-based analytic approach and (right) numerical perturbation approach for tracer response in a homogeneous quarter five-spot pattern at a fixed time.
Figure 2. Reference permeability distribution generated using a stochastic moving-average method.

Figure 4. Estimated permeability distribution derived using the conservative tracer response.

Figure 7. Reference saturation distribution for the synthetic example.

Figure 9. Estimated saturation distribution derived using the partitioning tracer response.
Figure 11. Model assessment using resolution and covariance analysis: (a) resolution of saturation estimates and (b) variance of saturation estimates.

Figure 17. Hydraulic conductivity field estimated from the conservative tracer data for the Hill case (log permeability scale).
Figure 21. NAPL saturation distribution estimated from partitioning tracer data for the Hill case (cut-away view).

Figure 22. Spatial distribution of retardation factors computed at sampling locations.

Figure 23. Three-dimensional streamline pattern of the final model.

Figure 24. Tracer swept volume at various times (days).
Figure 26. Resolution of estimated NAPL saturation.

Figure 27. Region of high NAPL saturation distribution estimated using simulated annealing technique for the Hill case.