A Bayesian geostatistical transfer function approach to tracer test analysis

Michael N. Fienen,[1] Jian Luo,[1,2] and Peter K. Kitanidis[1]

Received 12 September 2005; revised 10 January 2006; accepted 9 February 2006; published 29 July 2006.

[1] Reactive transport modeling is often used in support of bioremediation and chemical treatment planning and design. There remains a pressing need for practical and efficient models that do not require (or assume attainable) the high level of characterization needed by complex numerical models. We focus on a linear systems or transfer function approach to the problem of reactive tracer transport in a heterogeneous saprolite aquifer. Transfer functions are obtained through the Bayesian geostatistical inverse method applied to tracer injection histories and breakthrough curves. We employ nonparametric transfer functions, which require minimal assumptions about shape and structure. The resulting flexibility empowers the data to determine the nature of the transfer function with minimal prior assumptions. Nonnegativity is enforced through a reflected Brownian motion stochastic model. The inverse method enables us to quantify uncertainty and to generate conditional realizations of the transfer function. Complex information about a hydrogeologic system is distilled into a relatively simple but rigorously obtained function that describes the transport behavior of the system between two wells. The resulting transfer functions are valuable in reactive transport models based on traveltime and streamline methods. The information contained in the data, particularly in the case of strong heterogeneity, is not overextended but is fully used. This is the first application of Bayesian geostatistical inversion to transfer functions in hydrogeology but the methodology can be extended to any linear system.

Citation: Fienen, M. N., J. Luo, and P. K. Kitanidis (2006), A Bayesian geostatistical transfer function approach to tracer test analysis, *Water Resour. Res.*, *42*, W07426, doi:10.1029/2005WR004576.

1. Introduction

[2] In situ bioremediation depends on the ability to stimulate biomass in the subsurface through the delivery and mixing of nutrients and the controlled flow of contaminated groundwater. Modeling the movement of water and solutes through the aquifer provides the information vital for manipulation of subsurface conditions. The traditional approach to modeling groundwater flow and transport requires finding appropriate boundary and initial conditions and parameters, and applying them to mathematical models of groundwater flow and transport. This is the most comprehensive approach to effective management and has been the topic of many advances in hydrogeology in recent decades. However, the development and parameterization of a finely resolved 3-D model of a specific site remains expensive and time consuming.

[3] There is a pressing need for practical and efficient methods that allow for accurately simulating specific problems without the massive effort required for a highly resolved 3-D model. Such methods provide valuable management information within a reasonable budget and time horizon. At the simplest, one may assume a homogeneous domain governed by a 1-D analytical solution using a representative single set of parameters. However this simple model can conceal features important to accurate representation of the subsurface. A more general method is the transfer function (linear systems) approach [e.g., *Jury and Roth*, 1990]. Transfer functions represent the response of a system to an infinitesimally short duration input with unit magnitude (a Dirac pulse). The transfer function is equivalent to a breakthrough curve resulting from such a Dirac input. Chemical and biological reactions can be modeled based on interpreting breakthrough curves as traveltime probability density functions (pdfs) that provide contact and residence time information [e.g., *Simmons et al.*, 1995; *Kaluarachchi et al.*, 2000]. Efficient traveltime formulations [e.g., *Cvetkovic and Dagan*, 1994; *Crane and Blunt*, 1999; *Vasco and Datta-Gupta*, 1999; *Cirpka and Kitanidis*, 2000, 2001] enable the use of transfer functions as pdfs for use in modeling transport. Typically, transfer functions are found by assuming a predetermined shape [e.g., *Luo et al.*, 2006a]. This constraint can lead to oversimplification of the traveltime distribution in highly heterogeneous aquifers.

[4] Within the context of transfer function and traveltime research, our work provides a method that is less restrictive regarding assumptions of transfer function shape. Furthermore, it is stochastic, enabling the consideration of uncertainty. The development was motivated by a multiple-well recirculation tracer test conducted in a heterogeneous fractured saprolite aquifer. Data available include the injection

[1]Department of Civil and Environmental Engineering, Stanford University, Stanford, California, USA.
[2]Now at School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, Georgia, USA.

0043-1397/06/2005WR004576$09.00

history of tracers and the observed breakthrough curves at monitoring and extraction wells. The breakthrough curves can be interpreted as the convolution of the input signal and a transfer function corresponding. The transfer function can be used as a pdf for reaction modeling. From the known injection and observation data we can calculate the transfer function through deconvolution, cast as an inverse problem.

[5] A parametric approach using gamma distributions was initially applied to these data [*Luo et al.*, 2006a]. The key advantages to the parametric approach are computational efficiency and simplicity. Gamma distributions are unimodal functions of simple shape. There are cases, however, where multiple peaks in the transfer function may be encountered. The nonparametric approach allows the data to define the nature of modes and the complexity and roughness in the transfer function rather than prior assumptions made by the investigator. We will show that the method presented in this paper can recover a bimodal transfer function without any prior knowledge of multiple modes. Naturally, the advantages of generality afforded by nonparametric methods come at a computational cost.

[6] We constrain the transfer function to be nonnegative to obtain realistic confidence intervals and a best estimate that does not include physically impossible negative values. Furthermore, conditional realizations may be generated that honor the constraint; each of which can be used as the basis of a reactive model for Monte Carlo applications. This constraint is implemented using the Gibbs sampler with a reflected Brownian motion variogram characterizing the prior covariance [*Michalak and Kitanidis*, 2004].

2. Background

[7] The transfer function method has a long history in many areas, particularly signal and image processing [e.g., *Liu and Liu*, 1975; *Box et al.*, 1994]. Hydrologic applications include the unit graph (later called "unit hydrograph") runoff response of a watershed to rainfall [e.g., *Sherman*, 1932], the recharge response to infiltration [*Besbes and de Marsily*, 1984], and the discharge of Karst systems due to recharge [e.g., *Long and Derickson*, 1999; *Dreiss*, 1982].

[8] In these examples, as in the heterogeneous fractured aquifer motivating this work, it is difficult to obtain the parameters necessary to directly model a response to stimulus. Monitoring the actual response to a known stimulus, however, is relatively easy. The interpretation of such an experiment can provide significant information about the underlying system.

[9] In linear systems the response due to an arbitrary input can be modeled by superposition of multiple occurrences of the response to an input of infinitesimal duration with unit magnitude (a Dirac pulse). The functional relationship between the unit input and its response is the transfer function. The power of this approach is the ability to predict the response of a complex system characterized through a small amount of data. For the approach to be valid, the system must behave as linear and time invariant. These assumptions are discussed later.

[10] Most hydrologic linear systems applications employ a parametric transfer function in which a shape is assumed (e.g., lognormal, gamma, etc.) that can be described by a few parameters. These parameters are commonly determined by minimizing a measure of the difference between observations and simulated results calculated by convolution of the input and a transfer function. Fitting parameters to an analytical solution of the ADE is similar to a parametric transfer function deconvolution. Our method uses deconvolution to calculate nonparametric transfer functions in which fewer assumptions are made about the shape of the transfer function. For example, the number of modes and the nature of the tails may vary due to heterogeneity in an aquifer. Previous approaches have depended on accounting for multiple modes by constraining the solution to allow only one or two peaks in the transfer function [e.g., *Provencher*, 1982; *Neuman and de Marsily*, 1976] and require the specification of their locations. The ability to predict the locations of modes in the transfer function implies a level of understanding of the hydrogeology that may not be available in the early stages of an investigation. An inappropriate conceptual model may smooth out structure and inhibit the power of the data to provide the maximal amount of information.

[11] *Skaggs et al.* [1998] presented a nonparametric transfer function approach for vadose zone leaching. They used linear basis functions, a second derivative smoothness regularization term, and constrained the zeroth temporal moment to be unity. Their initial deconvolution did not impose a constrained number of modes in the transfer function, but if initial results indicated multiple modes, a statistical F test was used to determine whether constraining the number of modes resulted in a more significant result.

[12] Deconvolution to find the transfer function is an ill-conditioned inverse problem meaning small errors in measurement can result in large oscillations of the transfer function [e.g., *Blank et al.*, 1971; *Delleur and Rao*, 1971a]. These oscillations are often suppressed through regularization achieved either by enforcing smoothness or by filtering. Regularization can be applied to the transfer function, the input signal [e.g., *Neuman and de Marsily*, 1976], or both [e.g., *Daboczi*, 1998; *Long and Derickson*, 1999]. As explained in the next section, regularization is employed in this work in the form of prior covariance.

[13] The key assumptions necessary to apply the transfer function method are that the physics of the system are described as linear and time-invariant. In essence, any partial differential equation (PDE) with boundary and initial conditions that allow for a Green's functions solution meets the linearity requirement.

[14] The governing equation for breakthrough curves of a tracer in an aquifer is the general advection-dispersion equation (ADE)

$$\frac{\partial C}{\partial t} = -\mathbf{u} \cdot \nabla C + \nabla \cdot (\mathsf{D} \nabla C) \tag{1}$$

where C is concentration, t is time, $\mathbf{u}$ is the vector of velocities, and D is the dispersion tensor. Analytical solutions for various boundary conditions including 1-D step input [*Ogata and Banks*, 1961], 1-D Dirac instantaneous input [*Sauty*, 1980], and 2-D instantaneous input [*De Josselin De Jong*, 1958] are linear in the initial injected concentration or mass. *Leij et al.* [2000] presented Green's functions solutions for the ADE under various boundary and initial conditions. All of these solutions assume an initial concentration of zero. Since equation (1) is linear in the parameters, all solutions in which the solution starts at rest (i.e., concentration is zero) are themselves linear in the parameters [*Liu and Liu* 1975, pp. 59–62].

[15] The time invariance assumption requires that the system response must not change over time. For an aquifer, this means the properties such as hydraulic conductivity must be constant and also that all boundary conditions (including pumping rates of wells) must not change. These assumptions presuppose a steady state system with regard to flow. While this is restrictive, many remediation schemes are operated at steady state for long periods of time. The transfer function method is appropriate for such systems.

3. Problem Formulation

[16] The convolution integral simulates the response of a linear and time-invariant system to a particular stimulus

$$y(t) = \int_0^t x(\tau)s(t-\tau)d\tau \tag{2}$$

where $y(t)$ is the response at time t, $x(\tau)$ is the input (stimulus) at time τ, and $s(t-\tau)$ is the transfer function for lag time $t-\tau$. In the example of a tracer test, $y(t)$ is the concentration breakthrough curve (BTC) at a monitoring well and $x(\tau)$ is the concentration history at an injection well. We reiterate here that the transfer function is calculated from data for a given set of boundary and initial conditions. For example, a well field in which the pumping rates are held constant, and regional flow and recharge do not change. The transfer function calculated using the BTC at an observation well and the input function of an injection well characterizes the flow and transport between those two wells. To characterize the behavior throughout a well field, a transfer function must be calculated for each such pair.

[17] For numerical implementation the convolution integral must be expressed as a linear system of equations

$$\mathbf{y} = \mathbf{H}\mathbf{s} \tag{3}$$

where $\mathbf{y}$ is the known $n \times 1$ vector of measurements, $\mathbf{H}$ is the sensitivity matrix ($n \times m$) expressing the quadrature of the convolution integral and incorporating the known $x(t)$ input function and $\mathbf{s}$ is a vector ($m \times 1$) of unknowns (the transfer function $s(t)$ in this case). The vectors $\mathbf{y}$ and $\mathbf{s}$ and the discretized input function $\mathbf{x}$ are mapped in time to the same discretization $\Delta\tau$. To simplify the formulation, we use the commutative property of the convolution integral to restate equation (2) as

$$y(t) = \int_0^t s(\tau)x(t-\tau)d\tau. \tag{4}$$

[18] We express the quadrature using the trapezoidal rule,

$$\mathbf{H}_{(n\times m)} = \begin{bmatrix} \frac{1}{2}x(\tau_1)\Delta\tau & 0 & \cdots & 0 & 0 \\ \frac{1}{2}x(\tau_2)\Delta\tau & \frac{1}{2}x(\tau_1)\Delta\tau & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{1}{2}x(\tau_{m-1})\Delta\tau & x(\tau_{m-2})\Delta\tau & \cdots & \frac{1}{2}x(\tau_1)\Delta\tau & 0 \\ \frac{1}{2}x(\tau_m)\Delta\tau & x(\tau_{m-1})\Delta\tau & \cdots & x(\tau_2)\Delta\tau & \frac{1}{2}x(\tau_1)\Delta\tau \\ 0 & \frac{1}{2}x(\tau_m)\Delta\tau & \cdots & x(\tau_3)\Delta\tau & \frac{1}{2}x(\tau_2)\Delta\tau \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \frac{1}{2}x(\tau_m)\Delta\tau & \frac{1}{2}x(\tau_{m-1})\Delta\tau \\ 0 & 0 & \cdots & 0 & \frac{1}{2}x(\tau_m)\Delta\tau \end{bmatrix} \tag{5}$$

where $\Delta\tau$ is the same discretization step used for all vectors in the problem. The measurements ($y(t)$) and the unknown transfer function ($s(t)$) are discretized as vectors

$$\mathbf{y} = [y(t_1)\cdots y(t_n)]^T \tag{6}$$

$$\mathbf{s} = [s(t_1)\cdots s(t_m)]^T \tag{7}$$

where the superscript $[\ ]^T$ indicates matrix or vector transpose.

[19] In a specific application, there is no guarantee that measured functions will have a constant discretization so all observations are linearly interpolated to the same discretization. Each value of $\mathbf{y}$ corresponding to an actual observation is mapped to the nearest interpolated point on the grid. All other interpolated values of $\mathbf{y}$ are disregarded. Thus the dimension of $\mathbf{y}$ is the same as the number of actual observations (n). The matrix $\mathbf{H}$ from equation (5) must also be trimmed by removing each row for which an observation of $\mathbf{y}$ is not available. The interpolated values of $\mathbf{x}$ remain as part of the quadrature for the convolution integral implementation.

[20] The length of the vectors $\mathbf{x}$ and $\mathbf{y}$ is dictated by the availability of measurements. The length of $\mathbf{s}$, called memory, must be specified. Commonly the lengths of all three vectors are set to be the same. In many cases, however, $\mathbf{s}$ will tend to zero much sooner than $\mathbf{x}$ or $\mathbf{y}$. Extending $\mathbf{s}$ artificially increases the computational expense and the degrees of freedom which can introduce undesirable noise. In this work, the memory was specified based on observations of the length of the nonzero part of $\mathbf{s}$ in the unconstrained case.

[21] In a perfectly linear system without any errors associated with model selection or measurement, the solution would simply require the inverse (or pseudoinverse if $m \neq n$) of the $\mathbf{H}$ matrix in equation (3). In physically realistic systems, however, there are two sources of uncertainty: intrinsic variability inherent to the physical framework; and epistemic uncertainty due to the inability of our models to perfectly simulate nature, measurement sparsity and measurement errors [*Rubin*, 2003, p. 4]. Intrinsic variability is irreducible without infinitesimal discretization and is accounted for by the prior covariance. Epistemic uncertainty can be reduced by collecting more accurate, appropriate, or numerous measurements. Epistemic errors are often called measurement errors, but that terminology implies that they are only a function of inaccurate measurements when, in fact, they also result from imperfect conceptual models. In the measurement equation (equation (3)) we therefore add an epistemic error term to account for this

$$\mathbf{y} = \mathbf{H}\mathbf{s} + \mathbf{v} \tag{8}$$

where $\mathbf{v}$ is a vector of epistemic errors, modeled as a random process. Next we discuss the solution to obtain a best estimate and confidence intervals for $\mathbf{s}$.

4. Solution Method

[22] The Bayesian geostatistical inverse method has several advantages. Primarily, the method is much better than

other methods in allowing the data to speak for themselves. The only features imposed upon the shape of the breakthrough curve are the nonnegativity constraint, some degree of continuity due to selection of the prior covariance and a limit to the duration of the transfer function. The few "structural" parameters required by the method are optimized using recursive cross-validation techniques so that they too are determined by the behavior of the data. These impositions are much less restrictive than choosing the exact shape of the function as is done in parametric approaches.

4.1. Bayesian Inversion Method

[23] Bayes' rule states

$$p''(\mathbf{s}) \propto p(\mathbf{y}|\mathbf{s})p'(\mathbf{s}) \tag{9}$$

where p'' (**s**) is the posterior probability density function (pdf), p' (**s**) is the prior pdf, and p(**y**|**s**) is the likelihood function. We start by assuming the epistemic errors **v** in the measurement equation (equation (8)) are normally distributed. Without imposing a nonnegativity constraint, p''(**s**) can be solved for in closed form using cokriging equations [*Kitanidis and Vomvoris*, 1983; *Kitanidis*, 1995]. However, in the case of transport in porous media, the transfer function must be constrained to be nonnegative in all places because its units are concentrations per unit input concentration and negative concentrations are physically impossible.

[24] In the Bayesian approach, the role of the prior pdf (p'(**s**)) is to enforce constraints on the shape of **s** including smoothness, continuity and nonnegativity. The role of the likelihood function is to characterize the misfit of the model to the data for each estimate of **s**. Therefore nonnegativity is enforced through the prior pdf. Methods to enforce nonnegativity include the power transformation [*Kitanidis*, 1997] and the method of Lagrange multipliers [*Gill et al.*, 1981]. These methods may provide reasonable best estimates, but render the equations nonlinear and may produce unreasonable confidence intervals.

[25] In the prior pdf, assume a linear generalized covariance function (**Q**)

$$Q(t_i, t_j|\theta) = -\theta|t_i - t_j| \tag{10}$$

where θ is the nonnegative slope and $|t_i - t_j|$ is the distance (in time) between the ith and jth measurements. The linear generalized covariance function is chosen because of its maxentropic properties and versatility, enforcing continuity but not smoothness; this choice enforces very little prior assumption of shape.

[26] With a linear semivariogram we constrain the pdf of **s** to be nonnegative by reflecting about zero using the method of images. This method is summarized here; for the full derivation, see [*Michalak*, 2003; *Michalak and Kitanidis*, 2003, 2004]. The result is a reflected Gaussian distribution analogous to Brownian motion which is a Markov process, continuous in time and space with a stationary and memoryless transition probability. This means the pdf of **s** at time t is based only on the value of **s** at the precedent time. Employing the reflected Brownian motion model for the prior pdf renders the posterior pdf non-Gaussian. Markov chain Monte Carlo (MCMC) methods allow for the sampling of any multidimensional pdf. The MCMC method selected for this case is the Gibbs sampling algorithm. Prior applications of the MCMC method in Earth science inverse modeling using both the Gibbs sampler [*Geman and Geman*, 1984] and the Metropolis-Hastings algorithm [*Metropolis et al.*, 1953; *Hastings*, 1970] include those of *Mosegaard and Tarantola* [1995], *Oliver et al.* [1997], *Eidsvik et al.* [2002], and *Omre and Lodoen* [2004].

[27] The prior pdf can be expressed as the product of a Markov chain

$$p'(\mathbf{s}) = \prod_{i=2}^{m} p(s_i|s_{i-1}) \tag{11}$$

where m is the total number of unknown parameters (i.e., the discretized values of **s**) and

$$p(s_i|s_{i-1}) = \frac{1}{2\sqrt{\pi\theta(t_i - t_{i-1})}} \left[\begin{array}{c} \exp\left(-\frac{1}{2}\frac{(s_i - s_{i-1})^2}{2\theta(t_i - t_{i-1})}\right) \\ + \exp\left(-\frac{1}{2}\frac{(s_i + s_{i-1})^2}{2\theta(t_i - t_{i-1})}\right) \end{array} \right] \text{ for } s_i \geq 0 \tag{12a}$$

$$p(s_i|s_{i-1}) = 0 \text{ for } s_i < 0. \tag{12b}$$

[28] Assuming the epistemic errors are independent and uncorrelated, their covariance is

$$\mathbf{R} = \sigma_R^2 \mathbf{I} \tag{13}$$

where σ_R^2 is the epistemic error variance and **I** is an $n \times n$ identity matrix. Further assuming the errors are normally distributed with zero mean, the likelihood function is

$$p(\mathbf{y}|\mathbf{s}) \propto \frac{1}{\sqrt{\det(\mathbf{R})}} \exp\left[-\frac{1}{2}(\mathbf{y} - \mathbf{Hs})^T \mathbf{R}^{-1} (\mathbf{y} - \mathbf{Hs})\right]. \tag{14}$$

[29] The posterior pdf is then constructed by multiplying the new prior pdf (equation (11)) with the likelihood function (equation (14))

$$p''(\mathbf{s}|\mathbf{y}) \propto \prod_{i=2}^{m} p(s_i|s_{i-1}) \frac{1}{\sqrt{\det(\mathbf{R})}} \cdot \exp\left[-\frac{1}{2}(\mathbf{y} - \mathbf{Hs})^T \mathbf{R}^{-1} (\mathbf{y} - \mathbf{Hs})\right]. \tag{15}$$

4.2. Summary of Gibbs Sampler Algorithm for Conditional Realizations and Confidence Intervals

[30] The Gibbs sampling algorithm is a form of substitution sampling based on the definition of conditional probability using marginal distributions. In the two dimensional case, the Gibbs sampling algorithm is strictly substitution sampling, but in higher dimensions the two methods diverge somewhat. An excellent discussion of the Gibbs sampler algorithm was presented by *Casella and George* [1992]. Generally, a marginal posterior pdf is derived such that the probability of one element of **s**, referred to as s_i, can be

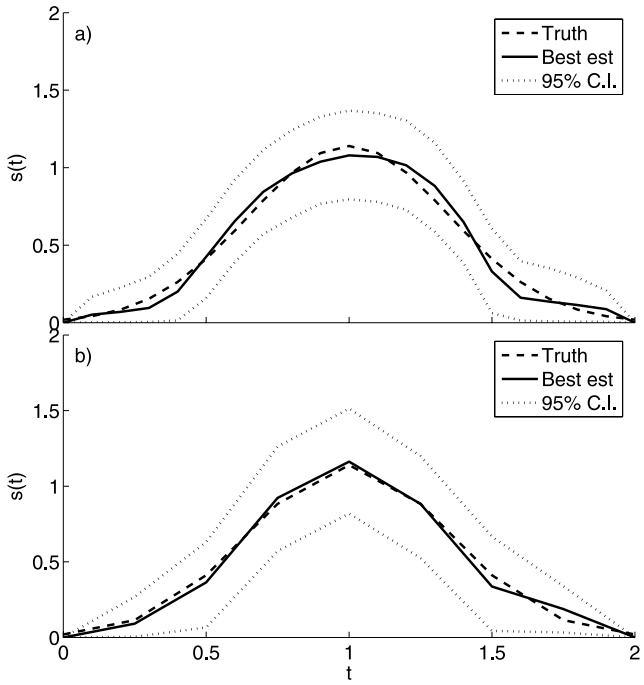

Figure 1. Deconvolution for the Gaussian example. (a) $\Delta t = 0.1$ and (b) $\Delta t = 0.25$. The dashed line is the actual transfer function ($s(t)$), the solid line is the best estimate (**Hs**) using 50,000 realizations, and the dotted lines are the 95% confidence interval.

calculated conditional on the other elements. This is done as a Markov Chain, removing each element in turn to calculate its marginal conditional on the other elements. Optimal structural parameters for the epistemic uncertainty (σ_R^2) and the variogram parameter (θ) are found using the recursive cross validation technique in the expectation-maximization (EM) algorithm [*Dempster et al.*, 1977]. More detail of the Gibbs sampler for this application is presented in the auxiliary material.[1]

[31] In several parts of the algorithm described thus far, numerical instability can arise from the disparity of scales in the independent and dependent parameter units. Particularly when m becomes large, the product of m values of the order of the dependent variable units may result in very large or very small numbers. As a result, it is desirable to transform the entire problem to a normalized space where all values are in the region of unity.

5. Verification Examples

[32] Several synthetic examples were chosen to verify the forward convolution process and to test the deconvolution method under controlled conditions with known outcomes.

5.1. Deconvolution Without Errors

[33] A Gaussian and a chapeau (hat) example were chosen to verify deconvolution using the inverse method. Both examples are generally smooth with continuous first (and higher for the Gaussian) derivatives. The Brownian motion variogram is not able to take advantage of such information and is quite rough at the small scale. Additionally, the chapeau function contains a sharp discontinuity which should be challenging to reproduce. In both cases, the convolution of the prespecified $x(t)$ and $s(t)$ functions was calculated using the MATLAB function conv. The resulting convolution $y(t)$ was used in conjunction with $x(t)$ as data simulating the input and response functions of a physical process (in a tracer test, $x(t)$ is the injection history, and $y(t)$ is the breakthrough curve observed in a monitoring or pumping well).

[34] The results for the Gaussian example are shown for $\Delta t = 0.1$ and $\Delta t = 0.25$ in Figures 1a and 1b, respectively. For these simulations, 50,000 realizations were used. Excellent agreement is seen between the true transfer function and the best estimate. The posterior variability indicated by the confidence intervals is reasonable considering that the linear variogram enforces continuity but not local smoothness. Therefore any individual realization is expected to have high-frequency variability about the mean.

[35] The results for two chapeau examples are shown for $\Delta t = 0.1$ and $\Delta t = 0.2$ in Figures 2a and 2b, respectively. As above, 50,000 realizations were employed. As in the Gaussian case, the best estimate makes a reasonable reproduction of the actual function. The corners are smoothed somewhat and the confidence intervals are wider than in the Gaussian case. The wider confidence intervals are expected due to the behavior of the linear variogram which has a difficult time reproducing a smooth curve with only a single break in slope. We reiterate that the linear variogram was chosen for its flexibility and tractability and this example was chosen to push it to its practical limit.

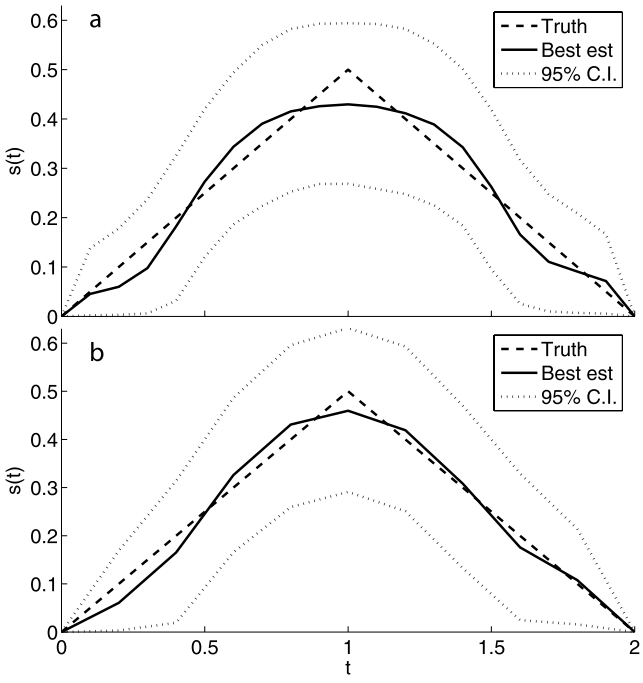

Figure 2. Deconvolution for the chapeau example. (a) $\Delta t = 0.1$ and (b) $\Delta t = 0.2$. The dashed line is the actual transfer function ($s(t)$), the solid line is the best estimate (**Hs**) using 50,000 realizations, and the dotted lines are the 95% confidence interval.

[1] Auxiliary material is available at ftp://ftp.agu.org/apend/wr/2005wr004576.

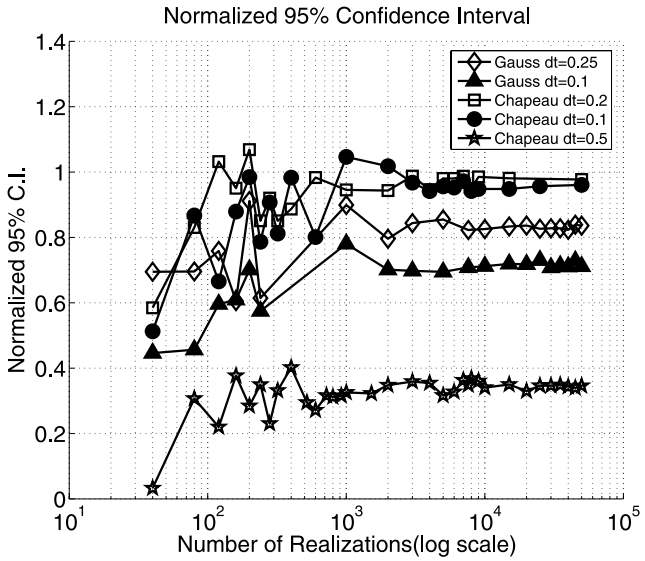

Figure 3. Plot of the behavior of the normalized 95% confidence interval (κ) as a function of the number of conditional realizations for several synthetic cases.

5.2. Analysis of Confidence Intervals

[36] Finding reasonable posterior confidence intervals is, in part, a function of the number of conditional realizations used in the deconvolution. Too few realizations yield artificially narrow confidence intervals indicating that the algorithm has not adequately sampled parameter space. There should be a critical number of realizations after which parameter space is adequately traversed and the statistics of the posterior estimates should be stable. For both the Gaussian and chapeau cases, the behavior of the confidence intervals was evaluated as a function of number of conditional realizations. The metric (κ) to evaluate the characteristics of the confidence intervals was based on the *L*2 norm of the 95% confidence interval relative to the *L*2 norm of the best estimate

$$\kappa = \frac{\sqrt{(\hat{\mathbf{s}}_{UCL} - \hat{\mathbf{s}}_{LCL})^T (\hat{\mathbf{s}}_{UCL} - \hat{\mathbf{s}}_{LCL})}}{\sqrt{\hat{\mathbf{s}}^T \hat{\mathbf{s}}}} \tag{16}$$

where $\hat{\mathbf{s}}$ is the best estimate, $\hat{\mathbf{s}}_{UCL}$ is the 95% upper confidence limit on the best estimate, and $\hat{\mathbf{s}}_{LCL}$ is the lower 95% confidence limit on the best estimate. Figure 3 shows the evolution of κ over the number of realizations used in the inverse problem. On the basis of these results, for the synthetic examples, κ becomes stable around 5000 iterations. To be conservative, 10,000 iterations is considered a reasonable value for most problems. The magnitude of κ stabilizes to a lower level when given more degrees of freedom, albeit at a slower rate. Furthermore, κ is lower in the case where the model is more appropriate for a linear variogram in the prior (i.e., κ is lower for the Gaussian case than for the chapeau case).

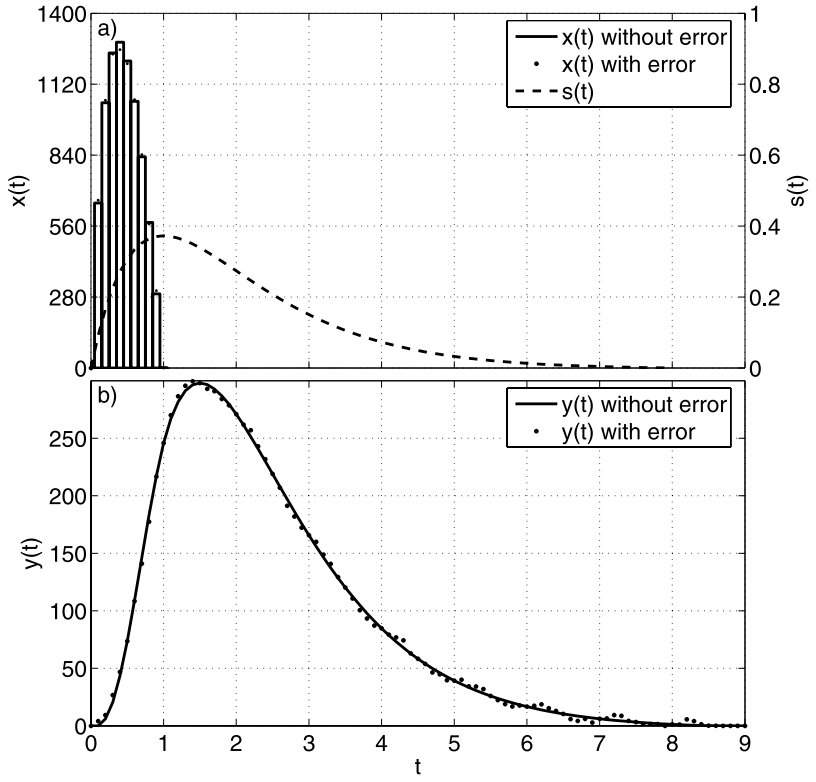

Figure 4. Synthetic functions with imposed errors for the *Neuman and de Marsily* [1976] synthetic example. Δt in all cases is 0.1. (a) Input function ($x(t)$) without error (solid line) and with errors (dots). The dashed line is the true impulse response function ($s(t)$). Note that ($s(t)$) is at a much smaller scale, read on the right y axis. (b) Output function (solid line): $y(t) = x(t) * s(t)$, where $*$ indicates convolution and the true $x(t)$ was used. The dots indicate $y(t)$ with errors introduced. Details of the nature of the errors are in the text.

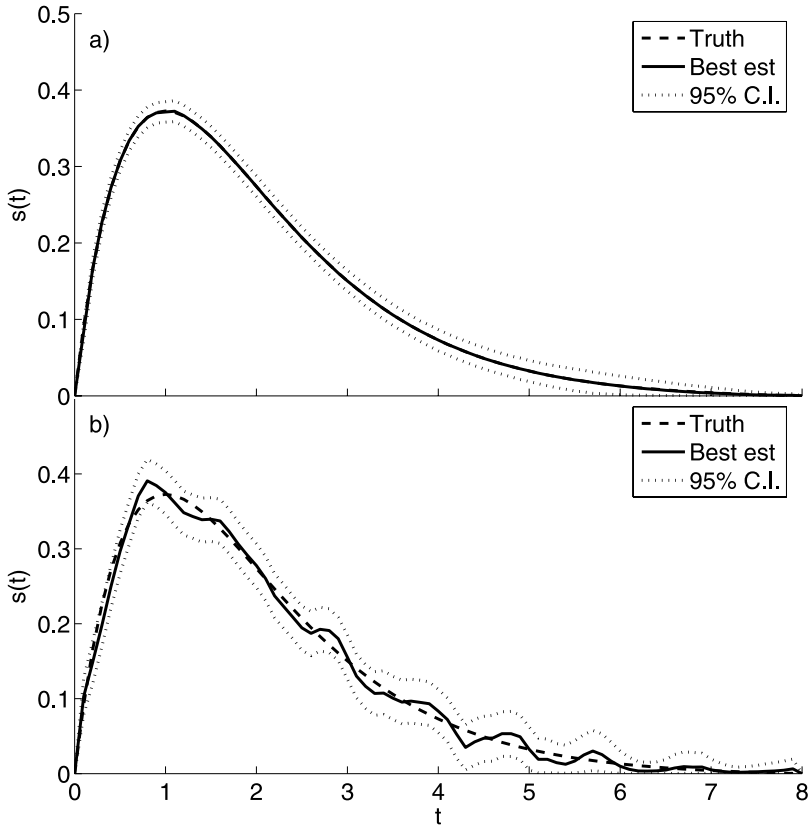

Figure 5. Transfer function calculated for the *Neuman and de Marsily* [1976] example. (a) Error free case and (b) case with introduced errors.

5.3. *Neuman and de Marsily* [1976] Example

[37] *Neuman and de Marsily* [1976] presented a parametric programming example in which good results were obtained for deconvolution by constraining the location and number of modes and by enforcing nonnegativity. The example they focused on was adapted from *Delleur and Rao* [1971b]. Oscillatory noise was introduced through errors particularly in the output function. They smoothed the transfer function to suppress noise. We reexamine this function both because analytical solutions are available for $x(t)$, $y(t)$ and $s(t)$, and to evaluate the Bayesian geostatistical inverse method's ability to suppress the noise without assuming a known number or location of modes.

5.3.1. Error-Free Neuman and de Marsily Case

[38] The input function was given as

$$x(t) = t(1-t)e^{(1-t)}\left(e^8 - 41\right) \text{ for } 0 < t < 1 \tag{17}$$

and the transfer function was given as

$$s(t) = t\left(e^{(8-t)} - 1\right)/\left(e^8 - 41\right) \text{ for } 0 < t < 8. \tag{18}$$

[39] The output function $y(t) = x(t) * s(t)$ was given explicitly as

$$y(t) = \frac{e^{(9-t)}t^3(2-t)}{12} + e^{(1-t)}\left(t^2 + 3t + 4\right) - e^{(4-t)} \text{ for } 0 \le t < 1 \tag{19a}$$

$$y(t) = \frac{e^{(9-t)}(2t-1)}{12} + (11 - 3t) - e^{(4-t)} \text{ for } 1 \le t < 8 \tag{19b}$$

$$y(t) = \frac{e^{(9-t)}}{12}\left[\begin{array}{c} 12t^2 - 130t + 335 - \\ (t-8)^2(152 - 14t - t^2) \end{array}\right] + (11 - 3t) \text{ for } 8 \le t \le 9. \tag{19c}$$

The functions are shown as the solid and dashed lines in Figures 4a and 4b.

[40] The deconvolution results with 10,000 realizations are presented in Figure 5a in which the solid line is the best estimate of $s(t)$, the dotted lines show the upper and lower 95% confidence limits, and the dashed line is the true value for $s(t)$. The function was estimated in the range $0 \le t \le 8$ with $\Delta t = 0.1$. The estimate of $s(t)$ exactly overlaps the true value with tight confidence intervals showing that, in the absence of epistemic error, the method is able to reproduce $s(t)$ exactly in this example.

5.3.2. Neuman and de Marsily Case With Errors

[41] Following *Neuman and de Marsily* [1976], random noise was added to $x(t)$ and a systematic low-frequency noise was added to $y(t)$. The equations for the signal corruption were not reported in their paper, so an estimate was made based on visual inspection of their Figure 3 and the description in the paper.

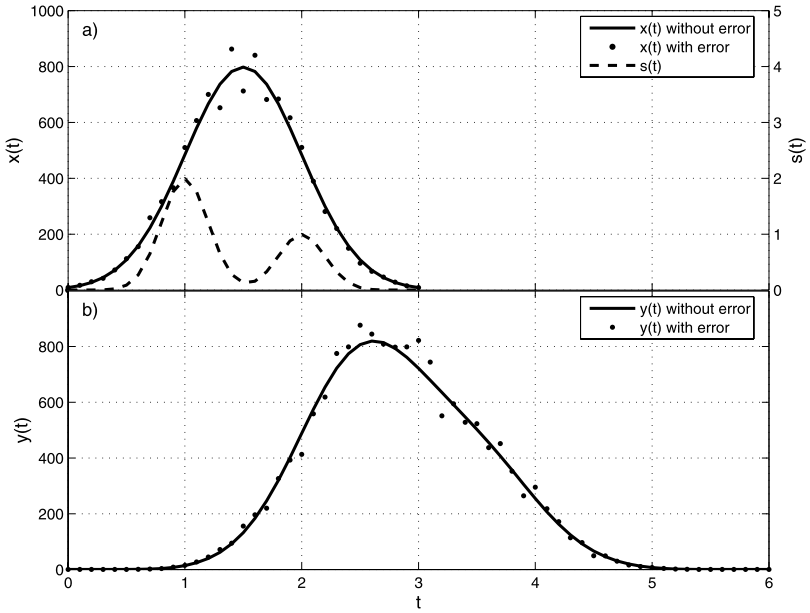

Figure 6. Synthetic functions with imposed errors for the bimodal synthetic example. Δt in all cases is 0.1. (a) Input function ($x(t)$) without error (solid line) and with errors (dots). The dashed line is the true impulse response function ($s(t)$). Note that ($s(t)$) is at a much smaller scale, read on the right y axis. (b) shows the output function (solid line): $y(t) = x(t) * h(t)$ where * indicates convolution, and the true $x(t)$ was used. The dots indicate ($y(t)$) with errors introduced. Details of the nature of the errors are in the text.

[42] The equations for the corrupted signals follow:

$$x_e(t) = x(t) + v_x(t) \quad \forall t \tag{20}$$

$$y_e(t) = y(t) + u_y(t)\sin(2\pi t)\forall t \tag{21}$$

where $x_e(t)$ and $y_e(t)$ are the values of $x(t)$ and $y(t)$ (calculated from equations (17) and (19) with errors introduced, $v_x(t)$ ~$N(0,123)$, where $N(\mu,\sigma^2)$ is a Gaussian distribution with mean μ and variance σ^2, and $u_y(t)$ is a uniformly distributed random value between 0 and 7. The values for the period of the oscillatory errors in $y_e(t)$ and the amplitude in both functions were determined such that the errors would be close (by visual inspection) to those measured on Figure 3 of *Neuman and de Marsily* [1976]. The original paper specified that the error in $y(t)$ was only added to the descending limb, but we also applied errors to the ascending limb. Negative values in $x_e(t)$ and $y_e(t)$ were set to zero.

[43] Figure 5b shows the results of the deconvolution with 10,000 realizations; the solid line is the best estimate of $s(t)$, the dotted lines show the upper and lower 95% confidence limits, and the dashed line is the true value for $s(t)$. The best estimate shows some oscillations of similar character to the best results obtained by *Neuman and de Marsily* [1976]. However, we did not specify the number or location of modes (whereas the previous work specified the location at $t = 1$). In the optimization for the structural parameters, the epistemic error value (σ_R^2) was two orders of magnitude higher than in the error-free case.

[44] This example shows the robustness of the present method to handle errors in the input and output functions, and its ability to correctly identify the location and magnitude of the mode. Furthermore, the optimal values of the structural parameters reproduced the smoothness and general shape of $s(t)$ with a minimum of prior assumptions.

5.4. Bimodal Example

[45] The final synthetic example employed a bimodal transfer function to stretch the method and verify that both modes could be accurately calculated. Using a Gaussian input function, it is difficult to conclude from visual inspection of the output function that a bimodal transfer function is even appropriate. The deconvolution was performed in the error-free case and in a case with extreme error imposed on both the input and output functions.

5.4.1. Error-Free Bimodal Case

[46] The input function ($x(t)$) is a Gaussian function

$$x(t) = N(1.5, 0.5) \times 1,000 \text{ for } 0 < t < 3. \tag{22}$$

[47] The transfer function ($s(t)$) is the superposition of two Gaussian functions

$$s(t) = N(1, 0.2) + \frac{N(2, 0.2)}{2} \text{ for } 0 < t < 3. \tag{23}$$

[48] The output function ($y(t)$) is calculated by the numerical convolution of $x(t)$ and $s(t)$

$$y(t) = x(t) * s(t) \text{ for } 0 < t < 6. \tag{24}$$

The functions are shown as the solid and dashed lines in Figures 6a and 6b.

[49] Figure 7a shows the results of the deconvolution with 10,000 realizations in which the solid line is the best estimate of $s(t)$, the dotted lines show the upper and lower

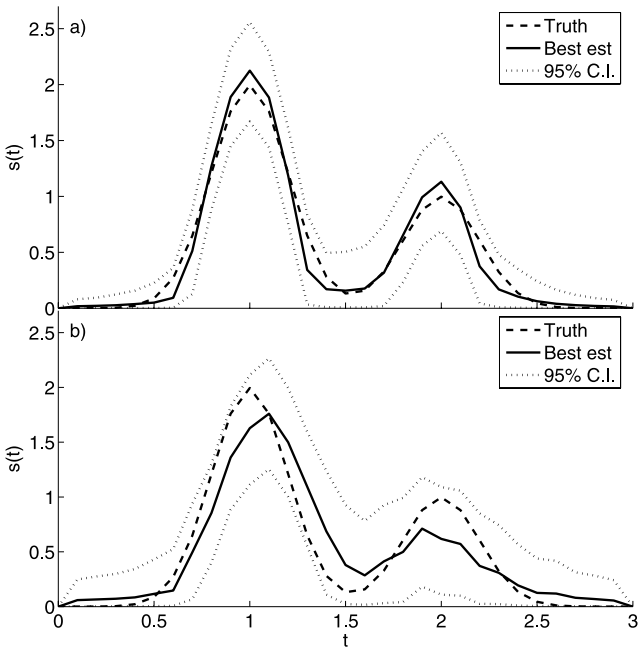

Figure 7. Transfer function calculated for bimodal example. (a) Error-free case and (b) case with introduced errors.

95% confidence limits, and the dashed line is the true value for $s(t)$. The fit is remarkable, with both modes estimated at the correct locations and magnitudes. The confidence intervals are also reasonable allowing for some variability in the conditional realizations, but all conditional realizations would retain the bimodal character with the second mode smaller than the first.

5.4.2. Bimodal Case With Errors

[50] Building on the success of the error-free results, we stretched the method by imposing error on both the input and output functions. In many cases, the input function is considered known without error while epistemic error is assumed in the acquisition of the output function. In practice, however, the same measurement technique is likely used to measure $x(t)$ and $y(t)$, so errors can be introduced in either case. The main consequence of this assumption is that the forward model (expressed as the **H** matrix) is formulated using $x(t)$, so errors are introduced even in the problem formulation.

[51] The error-free version of $x(t)$ was used to make an error-free version of $y(t)$ using equation (24), and then error was introduced on $x(t)$ and $y(t)$.

[52] The errors were normally distributed random values with mean zero and standard deviation scaled to the value of the function and added point by point

$$x_e(t) = x(t) + v_x(t) \quad \forall t \tag{25a}$$

$$y_e(t) = y(t) + v_y(t) \quad \forall t \tag{25b}$$

where $x_e(t)$ and $y_e(t)$ are the values of $x(t)$ and $y(t)$ with errors introduced, $v_x(t) \sim N(0, (x(t)/15)^2)$, and $v_y(t) \sim N(0, (y(t)/10)^2)$. Negative values in $x_e(t)$ and $y_e(t)$ were set to zero.

[53] The results of the deconvolution with 10,000 realizations are presented in Figure 7b. As in the error-free case, remarkable agreement with the actual transfer function is seen. The errors at the peaks of $x_e(t)$ and $y_e(t)$ shifted the peaks of the transfer function. Nonetheless, the first peak is consistently higher than the second and their locations are quite close to the actual values. As expected, the confidence intervals are wider due to a higher value for the epistemic error as calculated using the EM method. The optimal value for the variogram slope (θ) is similar to the error-free case, but the epistemic error (σ_R^2) is nearly two orders of magnitude higher.

[54] The bimodal case presented was arbitrary, but such a transfer function could arise in geologically realistic environments, particularly a case with two high-conductivity channels meandering through lower-conductivity floodplain deposits.

5.4.3. Bimodal Case Compared With Gamma Approach

[55] To illustrate the flexibility of the nonparametric method relative to a popular parametric method we perform the deconvolution of the bimodal transfer function without errors constraining the estimated transfer function to be a three-parameter gamma distribution of the form

$$s(t) = \frac{\beta}{\theta^{\alpha}\Gamma(\alpha)} t^{\alpha-1} \exp\left(-\frac{t}{\theta}\right) \tag{26}$$

where α is the shape parameter, β is the mass recovery parameter, θ is the scale parameter, and $\Gamma(\alpha)$ is the single-parameter gamma function defined as

$$\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} \exp(-x) dx. \tag{27}$$

[56] The actual input, transfer, and output functions were created using Equations (22)–(24)). Deconvolution was performed by minimizing the objective function

$$J = (\mathbf{y} - \mathbf{Hs})^T (\mathbf{y} - \mathbf{Hs}) \tag{28}$$

which is the least squares criterion on misfit of the model (**Hs**) relative to the measurements (**y**). The three parameters α, β and θ were fit in the least squares minimization. Figure 8a shows the optimal transfer function when constrained to be a gamma distribution in the bimodal case. The area under the curve of the gamma transfer function is similar to the bimodal, and Figure 8b shows that the output function $y(t)$ is reproduced very well. This highlights the danger in making conclusions based simply on least squares fitting of an arbitrary function. It would be tempting to conclude that, based on reproduction of $y(t)$, the gamma transfer function is a good fit to the actual transfer function. This, however, could mask an important finding, such as a channel or a mobile-immobile flow condition that the multimodal transfer function indicates.

6. Application to Field Tracer Test Data

[57] This work was motivated by an ongoing bioremediation experiment at the Field Research Center of the U.S. Department of Energy Natural and Accelerated Bioremedi-

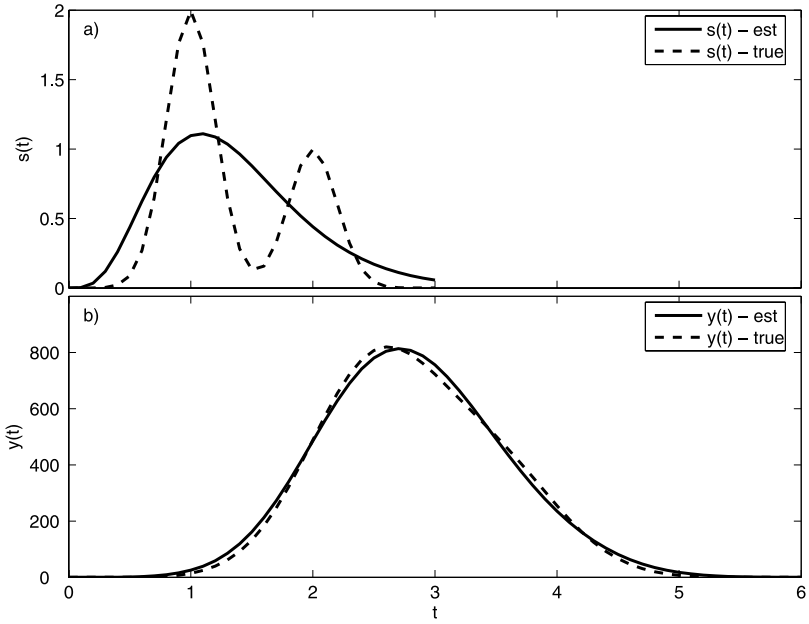

Figure 8. (a) Results of deconvolution for bimodal example without errors constraining $s(t)$ to be a gamma function. (b) Reproduction of $y(t)$ using the estimated transfer function constrained to be a gamma in the bimodal example case.

ation Research (NABIR) Program at the Y-12 National Security Complex in Oak Ridge, Tennessee. A nested recirculation system consisting of four pumping wells is in use to stimulate in situ microbial reduction of uranium from a mobile to a less mobile form (Figure 9) [*Wu et al.*, 2006]. The nested cell was implemented to create a recirculation zone (the inner loop) that is relatively isolated from the flow and geochemical conditions outside the outer loop [*Luo et al.*, 2006b]. Two injection wells, FW104 and FW24, were injecting at 0.4 and 0.8 L/min, respectively, and two extraction wells, FW103 and FW26, were each extracting at 0.4 L/min The outer loop injection rate (in FW24) was intentionally higher than the extraction rate to ensure that any leakage from the outer loop into the inner loop was

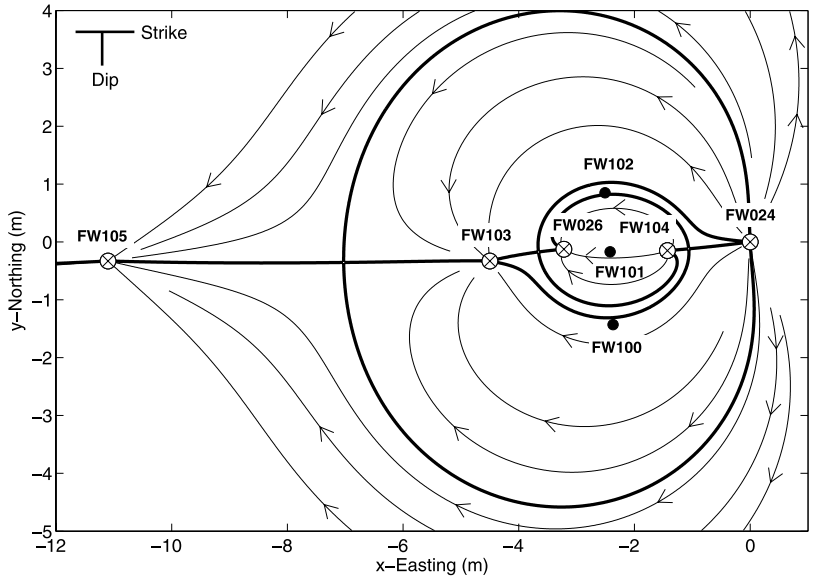

Figure 9. Schematic plan view of the nested dipole recirculation well field at the NABIR FRC in Oak Ridge, Tennessee. FW024 and FW104 are injection wells, and FW026, FW103, and FW105 are extraction wells. Tracer was injected in FW104 and observed in the multilevel sampler (MLS) wells, FW100, FW101, and FW102, and in the two main extraction wells, FW026 and FW103. Extraction from FW105 is for operational reasons related to the aboveground treatment system but is shown here only because it impacts the flow field. Lines with arrows indicate streamlines, and thick solid lines indicate separation streamlines delineating areas of recirculation, capture, and release.

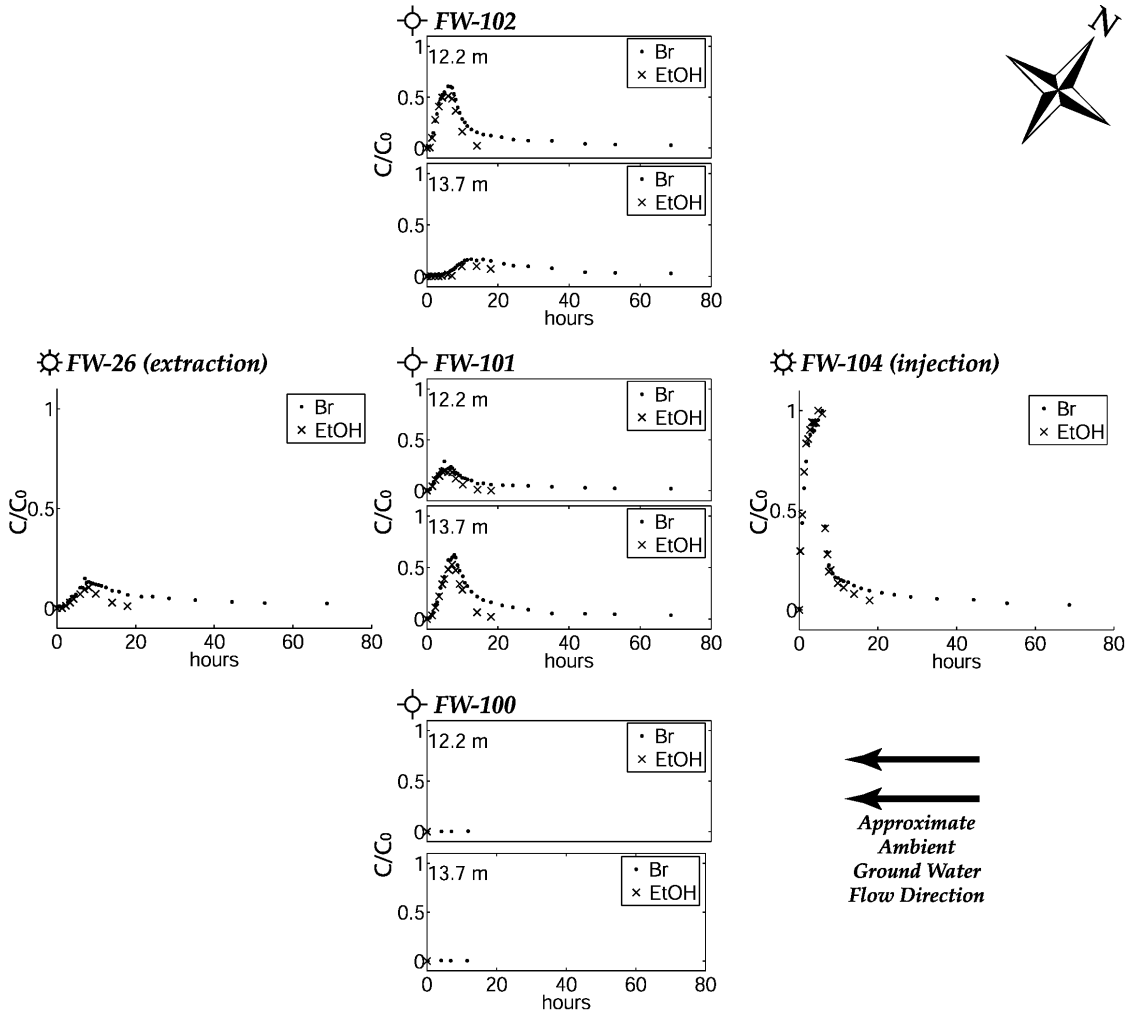

Figure 10. Injection history and breakthrough curves for bromide (conservative) and ethanol (reactive) during the NABIR FRC tracer test in October 2004.

injected water rather than native formation water. Flow occurs in a narrow band of highly fractured saprolite that dips at approximately 30° [*Solomon et al.*, 1992; *Fienen*, 2005].

6.1. Joint Tracer Test

[58] We conducted a joint tracer test injecting bromide and ethanol to investigate the transport of bromide and the first-order degradation (consumption) rate of ethanol. Both tracers were injected into FW104 for 360 min. The injection histories are presented in Figure 10. The tracers were fully mixed along the screened interval with a recirculation pump and the rest of the well was isolated using a packer. Samples were collected in the injection well to record the input function ($x(t)$) and in eight observation and pumping wells to provide breakthrough curves ($y(t)$). Measurements were collected for 10,700 min for bromide and 1300 min for ethanol. The tracer response was most pronounced in multilevel sampling wells FW101-2 and FW102-3 and much weaker in FW101-3 and FW102-2. Several other multilevel sampling locations are available but were measured very infrequently due to negligible response. These results are consistent with a previous tracer test, and the breakthrough curves for the four monitoring wells and extractions well FW26 are shown in Figure 10 for bromide and ethanol.

[59] Water extracted from the inner loop (FW26) was reinjected in FW104 after passing through the aboveground treatment system. The treatment system was expected to remove ethanol through stripping but not bromide, so the tail of the bromide injection history of FW104 is fatter. This is accounted for in the deconvolution, as would a more extreme case in which multiple peaks could be observed as recirculating pulses were reinjected. Water extracted from

Table 1. Reaction Rates and Discretization Parameters[a]

Well	k	Δt *Br*	Δt EtOH	Maximum t Br	Maximum t EtOH
FW 101-2	0.12 hr^{-1}	20	20	1,000	800
FW 101-3	0.13 hr^{-1}	20	20	1,100	600
FW 102-2	N/A	40	N/A	3,000	N/A
FW 102-3	0.10 hr^{-1}	20	20	1,100	700
FW 26	0.19 hr^{-1}	80	40	8,000	1,000
FW 103	N/A	100	N/A	10,000	N/A

[a]All times are in minutes except for the reaction rate, which is reported as 1/hr. N/A indicates ethanol data were incomplete, so only bromide analysis was conducted.

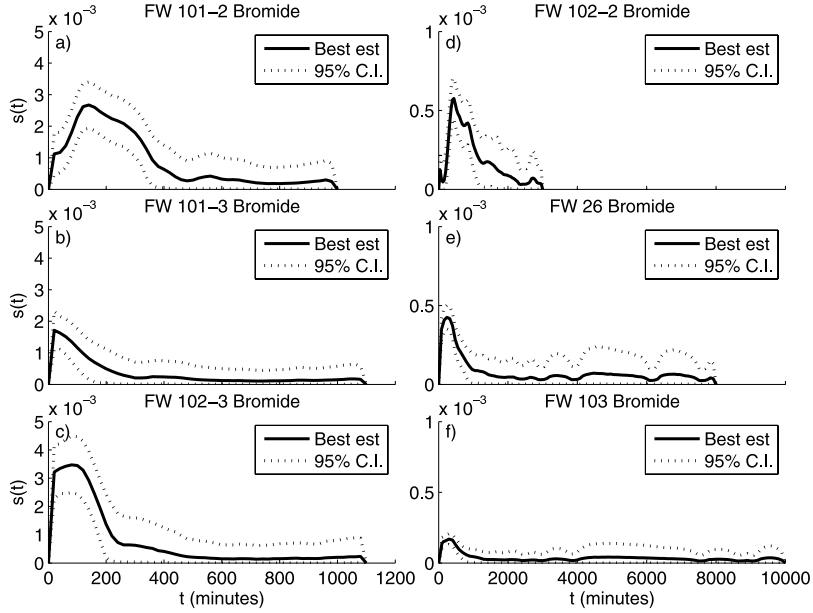

Figure 11. Transfer functions $s(t)$ for bromide as a function of time. Note different scales for $s(t)$ and time in Figures 11a–11c versus Figures 11d–11f. The solid line is the best estimate of the transfer function, and the dashed lines indicate the upper and lower 95% confidence intervals.

the outer loop (FW103) was also reinjected into the outer loop at FW24. Prior to reinjection, this water was mixed in a large tank that diluted any remaining bromide, so bromide was not injected in FW24.

[60] The discretization, memory and first-order reaction rates of ethanol are presented in Table 1. The transfer functions are presented in Figures 11 and 12 for bromide and ethanol, respectively. Discretization was selected to be close to the highest sampling frequency. For each well pair, the unconstrained transfer function was calculated with memory equal to the maximal time for which **x** or **y** measurements were available. Memory for the constrained case was selected as a point beyond which the unconstrained transfer function was flat and near zero. In either case, the transfer functions do not reach a concentration of zero. The laboratory analytical methods never truly report a value of zero, and the washout from the injection well results in an initially steep, then asymptotic decline in concentration on the trailing limb of the injection history. Matrix diffusion and bromide recirculation may also contribute to continued nonzero concentrations even after the main pulse of tracers has passed.

[61] The strength of the transfer functions is highest in monitoring wells FW101-2 and FW102-3, lower in FW101-3 and FW102-2, and lowest in the extraction wells. This is to be expected since the extraction wells pull tracer-free water that dilutes the signal. Furthermore, FW26 (the inner loop extraction well) captures more tracer than FW103 (the outer loop extraction well) indicating that the tracer is largely contained within the inner loop, although there is some leakage to the outer loop. The signal in the extraction wells is also smeared more than the other wells due to distance and greater number of flow paths traversed. Both FW26 and FW103 show potential modes far out in their tails. While it is possible that these are the result of noise (as suggested by the wider confidence intervals in these regions), it is also likely that delayed flow paths are encountered by the extracted water as it is sampling a greater diversity of flow paths.

6.2. First-Order Reaction Rate

[62] A two-tracer test can be conducted in which a nonreactive tracer (C_T) and a reactive tracer (C_R) are injected under the same flow conditions and their breakthrough curves are analyzed. For example, in our application to a field tracer study, first-order consumption of ethanol was assumed while bromide was a conservative tracer. Assuming 1-D flow along a streamline, equation (4) can be restated in terms of input concentration (C_T^{in}) and breakthrough curve (C_T)

$$C_T(t) = \int_0^t s(\tau) C_T^{in}(t-\tau) d\tau \tag{29}$$

[63] For the reactive tracer, if a first-order reaction rate coefficient k is assumed, C_R is calculated as

$$C_R(t) = \exp(-kt) C_T(t) \tag{30}$$

and the reaction rate is found as

$$k = -\frac{d}{dt} \ln\left(\frac{C_R}{C_T}\right). \tag{31}$$

[64] The concentration histories of C_R and C_T are represented by the corresponding transfer functions obtained by deconvolution. The results, shown in Table 1, are comparable to results determined using a parametric transfer function approach [*Luo et al.*, 2006a]. These rates indicate an estimation of the consumption of substrate only. To model a more complicated nonlinear biotransformation

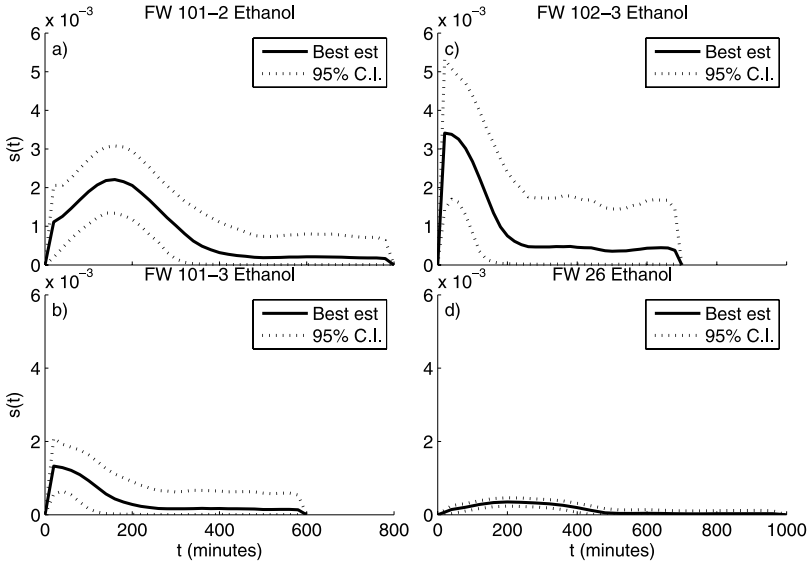

Figure 12. Transfer functions $s(t)$ for ethanol as a function of time. Note different timescales in Figures 12a and 12b versus Figures 12c and 12d. The solid line is the best estimate of the transfer function, and the dashed lines indicate the upper and lower 95% confidence intervals.

processes, the transfer function for the conservative tracer can be used with a stochastic-convective model.

7. Conclusions

[65] We show the robustness and utility of the Bayesian geostatistical inverse method for the deconvolution of tracer injection and breakthrough curves. The method presented can be applied to any linear time-invariant system where an input signal and output signal are both measured. The transfer function can be interpreted as a pdf and used in geochemical reactions to provide a more accurate solution than, for example, assuming a 1-D solution with bulk parameters. The shape of transfer functions can reveal characteristics of the hydrogeology of the site such as multiple major flow paths or channels expressed as multiple modes, or mass transfer limitations expressed in the tails. A comparison of transfer functions obtained with the same boundary conditions at different times may indicate changes in hydraulic conductivity due to biomass accumulation in bioremediation experiments.

[66] We showed through synthetic examples that the oscillatory errors plaguing previous transfer function approaches to hydrologic problems are handled well through the regularization provided by the prior covariance model in the Bayesian inverse method when the structural parameters are obtained optimally through the EM algorithm. Furthermore, the bimodal example illustrated the flexibility of the method. It is often difficult to guess the location or even the number of the modes based on inspection of data. A parametric transfer function approach may therefore provide an inappropriate solution whereas the nonparametric method presented here captures the multiple modes accurately. A hydrogeologic example in which multiple modes could be encountered is a braided channel environment in which anastomosing of streams may occur at different elevations such that even a 2-D numerical model would likely fail to predict travel paths correctly. The need for a highly resolved 3-D model to directly model the flow and transport accurately can, however, be mitigated using the transfer function method presented herein. Our method is also not tied to a single conceptual model of the hydrogeologic setting. This allows for flexibility and a more faithful representation of the actual flow and transport conditions than would likely be obtained through a numerical modeling effort with limited data. The data are pushed to extract as much information as possible without overextending their validity.

[67] Using the MCMC Gibbs sampler with reflected Brownian motion requires some computational effort (CPU time up to several days on a typical desktop computer) but this is much less significant than the amount of human and computer resources required to parameterize a 3-D numerical flow and transport model, and pales in comparison to the time and expense associated with conducting a tracer test. There are always trade-offs between using a transfer function approach and a more complicated model. Epistemic uncertainty arises due to violations of the assumptions of steadiness and linearity, both of which are handled better by a more complicated model. The greater number of parameters in a numerical model, however, the greater the cost of obtaining sufficient measurements for calibration. We emphasize here, again, that the method in this work is not universally applicable, but is most useful in cases where the underlying physics are difficult to model, boundary conditions are steady state, and where measurements are expensive to obtain. In the field case presented in this work, the greatest expense is installing wells. While many measurements can be obtained over time at the existing wells, the spatial coverage afforded by each well is limited and cannot be enhanced without great expense. Therefore the method was developed to provide the best solutions consistent with available data and to suppress errors introduced through measurement error and sparsity, and the minor violations of time invariance and linearity.

[68] Using a reactive tracer in combination with the conservative tracer, the first-order reaction rate was also estimated. This is only possible when a linear reaction is assumed. In the more complicated nonlinear reactions characterizing other parts of the bioremediation project, a stochastic-convective model can be applied using the transfer functions of the conservative tracer to characterize convection.

[69] In the particular application presented (the tracer test at the NABIR FRC in Oak Ridge, Tennessee), conservative tracer was injected into a recirculating well field such that effectively multiple pulses were injected. The deconvolution resulted in a transfer function that distinguished late arrivals due to the initial injection from early arrivals of reinjection. The nature of the transfer functions enabled a direct comparison of the response for multiple sampling points which confirmed assumptions made based on the lithology and structural geology of the site. The stochastic nature of the method allows for uncertainty to be evaluated. Conditional realizations can be applied to Monte Carlo modeling of the geochemical or biological reactions.

[70] An implementation in MATLAB is provided at http://www.stanford.edu/group/peterk/software/lstran.html. The only requirements are files with the input and output function in the same time units, and the editing of a parameter file choosing a few options such as discretization and memory. The Bayesian inverse approach to the transfer function method is not limited to ground water applications, and could be extended to surface water hydrology, infiltration/recharge, or other linear systems.

[71] **Acknowledgments.** The authors wish to thank Craig Criddle, who, as Principal Investigator, has supported and motivated this work and more; Anna Michalak for sharing her code to base the nonnegativity approach upon; Philip Jardine and David Watson for FRC support; Jack Carley and Tonia Mehlhorn for field support and experiment planning at the FRC; and Jennifer Nyman, Margaret Gentile, and Carlos Diaz at Stanford and Kenneth Lowe at ORNL for analyzing bromide and ethanol data. Wei-min Wu has also provided guidance and input in all aspects of the project. This manuscript was improved by the reviews of three anonymous reviewers. This work was funded by the United States Department of Energy Natural and Accelerated Bioremediation Research Biological and Environmental Research grant DE-F603-00ER63046.

References

Besbes, M., and G. de Marsily (1984), From infiltration to recharge: Use of a parametric transfer function, *J. Hydrol.*, *74*, 271–293.

Blank, D., J. Delleur, and A. Giorgini (1971), Oscillatory kernel functions in linear hydrologic models, *Water Resour. Res.*, *7*, 1102–1117.

Box, G. E. P., G. M. Jenkins, and G. C. Reinsel (1994), *Time Series Analysis: Forecasting and Control*, 3rd ed., Prentice-Hall, Upper Saddle River, N. J.

Casella, G., and E. George (1992), Explaining the Gibbs sampler, *Am. Stat.*, *46*, 167–174.

Cirpka, O., and P. Kitanidis (2000), An advective-dispersive stream tube approach for the transfer of conservative-tracer data to reactive transport, *Water Resour. Res.*, *36*, 1209–1220.

Cirpka, O., and P. Kitanidis (2001), Travel-time based model of bioremediation using circulation wells, *Ground Water*, *39*, 422–432.

Crane, M., and M. Blunt (1999), Streamline-based simulation of solute transport, *Water Resour. Res.*, *35*, 3061–3078.

Cvetkovic, V., and G. Dagan (1994), Transport of kinetically sorbing solute by steady random velocity in heterogeneous porous formations, *J. Fluid Mech.*, *265*, 189–215.

Daboczi, T. (1998), Nonparametric identification assuming two noise sources: A deconvolution approach, *IEEE Trans. Instrum. Measure.*, *47*, 828–832.

De Josselin De Jong, G. (1958), Longitudinal and transverse diffusion in granular deposits, *Eos Trans. AGU*, *39*, 67.

Delleur, J., and R. Rao (1971a), Noise in analysis of linear hydrologic systems, *Eos Trans. AGU*, *52*, 198.

Delleur, J. W., and R. A. Rao (1971b), Linear systems analysis in hydrology—The transform approach, the kernel oscillations and the effect of noise, in *Systems Approach to Hydrology: Proceedings of the First Bilateral U.S.-Japan Seminar in Hydrology*, edited by V. M. Yevjevich, pp. 116–142, Water Resour. Publ., Highlands Ranch, Colo.

Dempster, A., N. Laird, and D. Rubin (1977), Maximum likelihood from incomplete data via EM algorithm, *J.R. Stat. Soc., Ser. B*, *39*, 1–38.

Dreiss, S. J. (1982), Linear kernels for karst aquifers, *Water Resour. Res.*, *18*, 865–876.

Eidsvik, J., H. Omre, T. Mukerji, G. Mavko, and P. Avseth (2002), Seismic reservoir prediction using Bayesian integration of rock physics and Markov random fields: A North Sea example, *Leading Edge*, *21*, 290–294.

Fienen, M. N. (2005), The three-point problem, vector analysis and extension to the n-point problem, *J. Geosci. Educ.*, *53*, 357–362.

Geman, S., and D. Geman (1984), Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images, *IEEE Trans. Pattern Anal. Mach. Intell.*, *6*, 721–741.

Gill, P. E., W. Murray, and M. H. Wright (1981), *Practical Optimization*, Elsevier, New York.

Hastings, W. (1970), Monte-Carlo sampling methods using Markov chains and their applications, *Biometrika*, *57*, 97–109.

Jury, W. A., and K. Roth (1990), *Transfer Functions and Solute Movement Through Soil: Theory and Applications*, Springer, New York.

Kaluarachchi, J., V. Cvetkovic, and S. Berglund (2000), Stochastic analysis of oxygen- and nitrate-based biodegradation of hydrocarbons in aquifers, *J. Contam. Hydrol.*, *41*, 335–365.

Kitanidis, P. K. (1995), Quasi-linear geostatistical theory for inversing, *Water Resour. Res.*, *31*, 2411–2419.

Kitanidis, P. K. (1997), *Introduction to Geostatistics: Applications in Hydrogeology*, Cambridge Univ. Press, New York.

Kitanidis, P. K., and E. G. Vomvoris (1983), A geostatistical approach to the inverse problem in groundwater modeling (steady state) and one-dimensional simulations, *Water Resour. Res.*, *19*, 677–690.

Leij, F. J., E. Priesack, and M. G. Schaap (2000), Solute transport modeled with Green's functions with application to persistent solute sources, *J. Contam. Hydrol.*, *41*, 155–173.

Liu, C. L., and J. W. S. Liu (1975), *Linear Systems Analysis*, McGraw-Hill, New York.

Long, A. J., and R. G. Derickson (1999), Linear systems analysis in a karst aquifer, *J. Hydrol.*, *219*, 206–217.

Luo, J., O. A. Cirpka, M. N. Fienen, W. Wu, T. L. Mehlhorn, J. Carley, P. M. Jardine, C. S. Criddle, and P. K. Kitanidis (2006a), A parametric transfer function concept for analyzing reactive transport in nonuniform flow, *J. Contam. Hydrol.*, *83*, 27–41.

Luo, J., W.-M. Wu, M. Fienen, P. M. Jardine, T. L. Mehlhorn, D. B. Watson, O. Cirpka, C. S. Criddle, and P. Kitanidis (2006b), A nested-cell approach for in situ remediation, *Ground Water*, *44*(2), 266–274.

Metropolis, N., A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller (1953), Equation of state calculations by fast computing machines, *J. Chem. Phys.*, *21*, 1087–1092.

Michalak, A. (2003), Application of Bayesian inference methods to inverse modeling for contaminant source identification, Ph.D. dissertation, Stanford Univ., Stanford, Calif.

Michalak, A., and P. Kitanidis (2003), A method for enforcing parameter nonnegativity in Bayesian inverse problems with an application to contaminant source identification, *Water Resour. Res.*, *39*(2), 1033, doi:10.1029/2002WR001480.

Michalak, A., and P. Kitanidis (2004), A method for the interpolation of nonnegative functions with an application to contaminant load estimation, *Stochastic Environ. Res. Risk Assess.*, *19*, 8–23.

Mosegaard, K., and A. Tarantola (1995), Monte Carlo sampling of solutions to inverse problems, *J. Geophys. Res.*, *100*, 12,431–12,447.

Neuman, S., and G. de Marsily (1976), Identification of linear systems response by parametric programming, *Water Resour Res*, *12*, 253–262.

Ogata, A., and R. B. Banks (1961), A solution of the differential equation of longitudinal dispersion in porous media, *U.S. Geol. Surv. Prof. Pap.*, *411-A*.

Oliver, D., L. Cunha, and A. Reynolds (1997), Markov chain Monte Carlo methods for conditioning a permeability field to pressure data, *Math. Geol.*, *29*, 61–91.

Omre, H., and O. Lodoen (2004), Improved production forecasts and history matching using approximate fluid-flow simulators, *SPE J.*, *9*, 339–351.

Provencher, S. (1982), A constrained regularization method for inverting data represented by linear algebraic or integral equations, *Comput. Phys. Commun.*, *27*, 213-27.

Rubin, Y. (2003), *Applied Stochastic Hydrogeology*, Oxford Univ. Press, New York.

Sauty, J.-P. (1980), An anlalysis of hydrodispersive transfer in aquifers, *Water Resour. Res.*, *16*, 145–158.

Sherman, L. K. (1932), Streamflow from rainfall by the unit-graph method, *Eng. News Rec.*, *108*, 501–505.

Simmons, C., T. Ginn, and B. Wood (1995), Stochastic-convective transport with nonlinear reaction: Mathematical framework, *Water Resour. Res.*, *31*, 2675–2688.

Skaggs, T., Z. Kabala, and W. Jury (1998), Deconvolution of a nonparametric transfer function for solute transport in soils, *J. Hydrol.*, *207*, 170–178.

Solomon, D. K., G. K. Moore, L. E. Toran, R. B. Dreier, and W. M. McMaster (1992), A hydrologic framework for the Oak Ridge Reservation, *Status Rep. ORNL/TM-12026*, Environ. Sci. Div., Oak Ridge Natl. Lab., Oak Ridge, Tenn.

Vasco, D., and A. Datta-Gupta (1999), Asymptotic solutions for solute transport: A formalism for tracer tomography, *Water Resour. Res.*, *35*, 1–16.

Wu, W. M., et al. (2006), Pilot-scale bioremedation of uranium in a highly contaminated aquifer. I: Conditioning of a treatment zone, *Environ. Sci. Technol*, in press.

M. N. Fienen and P. K. Kitanidis, Department of Civil and Environmental Engineering, Stanford University, Stanford, CA 94305-4020, USA. (fienen@stanford.edu)

J. Luo, School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA.