Hydrol. Earth Syst. Sci. Discuss., 7, 2017–2052, 2010
www.hydrol-earth-syst-sci-discuss.net/7/2017/2010/

Hydrology and Earth System Sciences Discussions

This discussion paper is/has been under review for the journal Hydrology and Earth System Sciences (HESS). Please refer to the corresponding final paper in HESS if available.

Bayesian approach for three-dimensional aquifer characterization at the hanford 300 area

H. Murakami[1], X. Chen[2], M. S. Hahn[2], Y. Liu[2], M. L. Rockhold[3], V. R. Vermeul[3], J. M. Zachara[3], and Y. Rubin[2]

[1]Department of Nuclear Engineering, University of California, Berkeley, California, USA
[2]Department of Civil and Environmental Engineering, University of California, Berkeley, California, USA
[3]Pacific Northwest National Laboratory, Richland, WA, USA

Received: 6 March 2010 – Accepted: 15 March 2010 – Published: 23 March 2010

Correspondence to: H. Murakami (harukom@nuc.berkeley.edu)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

This study presents a stochastic, three-dimensional characterization of a heterogeneous hydraulic conductivity field within DOE's Hanford 300 Area site, Washington, by assimilating large-scale, constant-rate injection test data with small-scale, three-dimensional electromagnetic borehole flowmeter (EBF) measurement data. We first inverted the injection test data to estimate the transmissivity field, using zeroth-order temporal moments of pressure buildup curves. We applied a newly developed Bayesian geostatistical inversion framework, the method of anchored distributions (MAD), to obtain a joint posterior distribution of geostatistical parameters and local log-transmissivities at multiple locations. The unique aspects of MAD that make it suitable for this purpose are its ability to integrate multi-scale, multi-type data within a Bayesian framework and to compute a nonparametric posterior distribution. After we combined the distribution of transmissivities with depth-discrete relative-conductivity profile from the EBF data, we inferred the three-dimensional geostatistical parameters of the log-conductivity field, using the Bayesian model-based geostatistics. Such consistent use of the Bayesian approach throughout the procedure enabled us to systematically incorporate data uncertainty into the final posterior distribution. The method was tested in a synthetic study and validated using the actual data that was not part of the estimation. Results showed broader and skewed posterior distributions of geostatistical parameters except for the mean, which suggests the importance of inferring the entire distribution to quantify the parameter uncertainty.

1 Introduction

Hydrogeological characterization plays a key role in various projects involving groundwater flow and contaminant transport. A detailed three-dimensional (3-D) description of spatial variability in subsurface hydraulic properties is imperative for predicting water and solute movement in the subsurface (Rubin, 2003). Recent focus on geochemical

and microbiological reactions in field studies, for example, requires flow parameters to be fully characterized a priori for testing their research hypotheses (Scheibe, 2001, 2003; Fienen et al., 2004).

One of the main challenges in hydrogeological characterization is to integrate datasets of different types and scales. Typical field studies usually include two or more different complementary sources of information, which may include depth-discrete small-scale data such as core analysis, slug tests and electromagnetic borehole flowmeter (EBF) tests and large-scale data such as pumping tests and tracer tests. With stochastic modeling of flow and transport becoming increasingly common, it is important not only to combine best-fitted values from each dataset, but also to correctly quantify and weigh errors and uncertainty associated with different datasets, and to transfer the uncertainty to the final prediction (Maxwell et al., 1999; Hou and Rubin, 2005; De Barros et al., 2009).

To tackle this challenge, various researchers have applied Bayesian approaches to the problem of subsurface characterization (Copty et al., 1993; Woodbery and Rubin, 2000; Chen et al., 2001). Within a Bayesian framework, the probability density function of a parameter vector can be updated sequentially to include more datasets in a consistent manner. In addition, the resulting predictive distribution can properly account for the parameter uncertainty inherent in estimating the parameter values from the data (Diggle and Ribeiro, 2002). Two recent developments in particular have increased the potential of the Bayesian approach for subsurface characterization: (1) Bayesian model-based geostatistics, and (2) the method of anchored distributions (MAD).

Bayesian model-based geostatistics, introduced by Kitanidis (1986) and Handcock and Stein (1993), assumes a parametric model for a spatial stochastic process and infers geostatistical structural parameters based on small-scale datasets or point measurements (Diggle and Ribeiro, 2006). While the traditional variogram approach determines best-fitted estimates of geostatistical structural parameters and their asymptotic confidence interval, the Bayesian model-based approach yields a posterior distribution of the parameters. Diggle and Ribeiro (2006) showed that correlation parameters such

as variance and scale follow non-Gaussian and skewed distributions, which suggests that the first two moments are not enough to characterize the distribution.

The method of anchored distributions (MAD) is a general Bayesian method for inverse modeling of spatial random fields that addresses complex patterns of spatial variability, multiple sources and scales of data available for characterizing the fields, and the complex relationships between observed and target variables (Zhang and Rubin, 2008a, b; Rubin et al., 2009). The central element of MAD is a new concept called "anchors." Anchors are devices for localization of large-scale data: they are used to convert large-scale, indirect data into local distributions of the target variables. The goal of the inversion is to determine the joint distribution of the anchors and structural parameters, conditioned on all of the measurements. The structural parameters describe large-scale trends of the target variable fields, whereas the anchors capture local heterogeneities. Following the inversion, the joint distribution of anchors and structural parameters can be directly used to generate random fields of the target variable(s). Different from most of the other inversion methods that determine a single best estimate of the field and asymptotic uncertainty bounds (Kitanidis, 1995; Zhu and Yeh, 2005; Ramarao et al., 1995), MAD yields a posterior distribution of the parameters.

In this paper, we assimilate EBF tests and constant-rate injection tests for characterizing a 3-D hydraulic conductivity K field at the Integrated Field Research Challenge (IFRC) site in DOE's Hanford 300 Area (http://ifchanford.pnl.gov). Since the EBF tests yield only relative K values along each of the EBF test wells, we need a local transmissivity T value at each of the EBF test wells to convert the relative values to absolute K values (Javandel and Witherspoon, 1969; Molz et al., 1994; Young et al., 1998; Fienen et al., 2004). The local T values can be determined by inverting the large-scale constant-rate injection tests. This assimilation requires us to quantify the uncertainty in T values based on the injection tests and to combine that uncertainty with the one from the EBF data.

The particular difficulty in inverting injection-test or pumping-test data is the computational effort associated with a long time series. Li et al. (2004) and Zhu and Yeh (2005)

have applied temporal moments of drawdowns to estimate T and storativity S fields. The sandbox experiment by Liu et al. (2007) has shown that the moment approach can successfully characterize a T field. The advantage of employing temporal moments is that we can compute them using steady-state flow equations, which can reduce the computational burden significantly. In addition, when the interest is limited to T, the zeroth-order temporal moment can eliminate the effects of the uncertainty in S, since it does not depend on S (Zhu and Yeh, 2005).

In the following sections, we first describe the site and the experimental procedure. We then present our approach, including the geostatistical inversion framework and the inference of the 3-D geostatistical parameters. After presenting the inversion results in a synthetic study to demonstrate and verify our approach, we discuss the results using the actual data at the site.

2 Site and experiment description

The Hanford 300 Area is located at the southern part of the Hanford Nuclear Reservation one mile north of Richland, Washington. The IFRC site is located within the footprint of a former disposal facility for uranium-bearing liquid wastes known as the South Process Pond, approximately 250 m west from the Columbia River. As is shown in Fig. 1, the triangular well field consists of 25 wells fully screened through the saturated portion of the Hanford formation, ten wells partially screened at different depths, and one deep characterization well (Bjornstad et al., 2009).

In this study, we focus on the saturated portion of the highly permeable and coarse-grained Hanford formation, which is a shallow unconfined aquifer. The main lithology is a poorly sorted mixture, dominated by gravel up to boulder size, with lesser amounts of sand and silt (Bjornstad et al., 2009). It overlies the Ringold formation, the upper portion of which is a continuous low-permeability layer consisting of cohesive and compact, well-sorted fine sand to silty sand. The saturated thickness is variable over the site, ranging from about 5 m to 8 m, with daily and seasonal fluctuations of the

water table in response to changes in the river stage. The prior estimates of hydraulic conductivity are 1000–100 000 m/day for the Hanford formation and 0.01–3.00 m/day for the Ringold formation (Meyer et al., 2007).

Fourteen constant-rate injection tests were conducted, each of which had one injection well and 7 to 10 observation wells. All the wells used in the tests are fully screened over the saturated portion of the Hanford formation. The distance between the injection and observation wells ranges between 8 and 60 m. The injection duration and rate are approximately 20 minutes and 315–318 gpm (1.19–1.20 $m^3\,min^{-1}$), respectively. The preliminary analysis of the late-time curve data, using the Cooper-Jacob straight-line method, has shown that most of the observation wells yield similar estimates for the T values in each test, which is considered to be the geometric mean of T, T_G, over the entire well field, as is mathematically proved by Sánchez-Villa et al. (1999). It suggests that the zone-of-influence expands very rapidly and the conventional pumping test analysis yields only an effective property, smoothing out the local heterogeneity at the well field.

The EBF test data were obtained at 19 fully screened wells, which yielded 283 depth-discrete relative hydraulic conductivities with 0.3–0.6 m depth intervals. The pumping rate was 1.04–1.55 gpm (3.94×10^{-3}–$3.94\times10^{-3}\,m^3\,min^{-1}$), and kept constant during the test at each well. The vertical profiles indicated that the hydraulic conductivity over the central third of the Hanford formation was lower than its top and bottom thirds at many of the wells. Although the thickness and contact depths for this lower permeability material vary across the site, this general pattern was observed to some extent at most of the monitoring well locations.

The more detailed description of the site and data is available in Bjornstad et al. (2009) and Rockhold et al. (2010).

3 Methodology

For the 3-D characterization, we employ a two-step approach. First, we invert the large-scale injection tests to characterize the T field. We apply MAD to invert the zeroth-order moments of pressure build-up curves at multiple observation wells. As a result, we obtain a joint posterior distribution of T at the EBF test wells. Second, we combine this distribution with the EBF data for determining the absolute K values. Instead of a single K value at each of the EBF data point, we obtain the distribution of K at each point. Based on the distribution of the absolute K, we infer the 3-D geostatistical parameters, using the Bayesian model-based geostatistics.

Compared to direct coupling of the EBF and pumping tests used in Li et al. (2008), this two-step approach has a significant computational advantage. This approach is possible because we can model the flow process during the injection tests as 2-D planar flow in the horizontal plane, due to the particular site conditions as the following. The coarse-grained and highly permeable nature of the aquifer caused the elastic response and drainage effect to occur very rapidly (less than 30 s after the injection started), so that the radial flow regime dominated the pressure buildup responses (Neuman, 1975). In addition, despite the large injection rate, the maximum pressure buildup at the nearest observation wells was less than several centimeters, which is much smaller than the aquifer thickness (5–8 m). Although the EBF tests suggested vertical heterogeneity in the saturated zone, Dagan et al. (2009) recently showed that Dupuit's assumption is still valid – when the aquifer thickness is not large compared to the vertical integral scale, and the ratio between the vertical and horizontal integral scale is large, which is the case at this site.

3.1 Geostatistical inversion for transmissivity field

3.1.1 MAD framework

In this section, we summarize the Bayesian inversion framework, called (as indicated above) method of anchored distributions (MAD), which we use to invert the injection test data. This method was introduced by Zhang and Rubin (2008a, b) and Rubin et al. (2009) and is summarized here for completeness.

We denote a spatial random process by $Y(\boldsymbol{x})$, where $\boldsymbol{x}$ is the space coordinate. We further denote an entire field of Y by $\tilde{\boldsymbol{Y}}$, and denote a realization of the field by $\tilde{\boldsymbol{y}}$. The field $\tilde{\boldsymbol{Y}}$ is defined through the vector of model parameters $\{\boldsymbol{\theta},\boldsymbol{\vartheta}\}$. The $\boldsymbol{\theta}$ part of this vector, called the structural parameter vector, includes a set of parameters designed to capture the global features of $\tilde{\boldsymbol{Y}}$, such as the mean of the field and correlation structures. The $\boldsymbol{\vartheta}$ component of this vector consists of the anchored distributions, or anchors in short. Anchors are devices used to capture local features of $\tilde{\boldsymbol{Y}}$ that cannot be captured by the global parameters $\boldsymbol{\theta}$. In their simplest form, an anchor would be error free measurements of Y. Other forms of anchors include measurements coupled with error distributions and/or anchors that are obtained by inversion.

The data $\boldsymbol{z}$ is a vector of multiple observations of a physical process. The data can be described by the following equation:

$$\boldsymbol{z} = \mathcal{M}\left(\tilde{\boldsymbol{y}}\right) + \boldsymbol{\varepsilon}, \tag{1}$$

where $\mathcal{M}$ is a known function, or a set of functions, numerical or analytical, of the spatial field, and $\boldsymbol{\varepsilon}$ is a vector of zero-mean errors. The goal of the inversion is, first, to derive a posterior distribution of the model parameters conditioned on the data $\boldsymbol{z}$, $p(\boldsymbol{\theta},\boldsymbol{\vartheta}|\boldsymbol{z})$. This distribution then allows us to generate multiple realizations of the field $\tilde{\boldsymbol{Y}}$ for prediction.

Using Bayes' rule, we can define the posterior distribution of parameters as:

$$p\left(\boldsymbol{\theta},\boldsymbol{\vartheta}|\boldsymbol{z}\right) \propto p\left(\boldsymbol{z}|\boldsymbol{\theta},\boldsymbol{\vartheta}\right) p\left(\boldsymbol{\vartheta}|\boldsymbol{\theta}\right) p(\boldsymbol{\theta}). \tag{2}$$

where $p(\boldsymbol{\theta})$ is the prior distribution, $p(\boldsymbol{\vartheta}|\boldsymbol{\theta})$ is the anchor distribution given a structural parameter vector $\boldsymbol{\theta}$, and $p(z|\boldsymbol{\theta},\boldsymbol{\vartheta})$ is the likelihood of data z given a parameter set $\{\boldsymbol{\theta},\boldsymbol{\vartheta}\}$.

We estimate the likelihood $p(z|\boldsymbol{\theta},\boldsymbol{\vartheta})$ using the Monte Carlo simulations. Since the model parameters $\{\boldsymbol{\theta},\boldsymbol{\vartheta}\}$ and the data z are connected through the field, we generate multiple conditional realizations of the field $\tilde{\boldsymbol{Y}}$ for any given $\{\boldsymbol{\theta},\boldsymbol{\vartheta}\}$; with each realization, the forward model provides a prediction of z in the form of $\tilde{z}$, according to Eq. (1). In other words, z is viewed as a measured outcome from a random process, whereas $\tilde{z}$ is one of many possible realizations, given a particular parameter set of $\{\boldsymbol{\theta},\boldsymbol{\vartheta}\}$. The ensemble of $\tilde{z}$ is used for estimating $p(z|\boldsymbol{\theta},\boldsymbol{\vartheta})$.

Since increasing dimension of z increases the computational burden of the likelihood estimation significantly, we may divide the vector z into L segments as $z=\{z_1,z_2,\ldots,z_L\}$. We can decompose the likelihood into each segment as,

$$\begin{aligned} p\left(z|\boldsymbol{\theta},\boldsymbol{\vartheta}\right) &= p\left(z_1,\ldots,z_L|\boldsymbol{\theta},\boldsymbol{\vartheta}\right) \\ &= p\left(z_L|z_1,\ldots,z_{L-1},\boldsymbol{\theta},\boldsymbol{\vartheta}\right) p\left(z_{L-1}|z_1,\ldots,z_{L-2},\boldsymbol{\theta},\boldsymbol{\vartheta}\right)\ldots p\left(z_2|z_1,\boldsymbol{\theta},\boldsymbol{\vartheta}\right) p\left(z_1|\boldsymbol{\theta},\boldsymbol{\vartheta}\right) \\ &\approx \prod_{l=1}^{L} p\left(z_l|\boldsymbol{\theta},\boldsymbol{\vartheta}\right). \end{aligned} \tag{3}$$

In Eq. (3), we assume that the data segments z_1, z_2,..., z_L are conditionally independent for a given $\{\boldsymbol{\theta},\boldsymbol{\vartheta}\}$, since we consider that $\{\boldsymbol{\theta},\boldsymbol{\vartheta}\}$ contains information equivalent to the data. This equality strictly holds when the data segments are independent of each other – for example, when the data locations are beyond the zone-of-influence or zone-of-correlation. It approximately holds when the data segments are only weakly correlated, such as with different types of data at the same site. As Hou and Rubin (2005) pointed out, assuming independence leads to higher entropy and makes the estimation less informative.

3.1.2 Specification of a 2-D geostatistical model

Here we specify the geostatistical model for the 2-D T field. Let $Y(\boldsymbol{x})$ be natural-log transmissivity, $\ln T(\boldsymbol{x})$, at the location $\boldsymbol{x}=(x_1, x_2)$ in the 2-D domain. We assume that a vector $\boldsymbol{Y}$, containing Y at multiple locations, follows a multivariate Gaussian distribution with exponential covariance. We define a structural parameter vector as $\boldsymbol{\theta}=\{\mu,\sigma^2,\phi\}$, including uniform mean μ, variance σ^2, and integral scale ϕ, which are used at a geologically similar site (i.e. unconsolidated glacial materials) (Rubin, 2003; McLaughlin et al., 1993).

We define a vector $\boldsymbol{\vartheta}(\boldsymbol{x}_\vartheta)$ to represent a set of anchors. Since the anchors are a subset of the field, $p(\boldsymbol{\vartheta}|\boldsymbol{\theta})$ is a multivariate Gaussian distribution with mean μ and covariance $\sigma^2\mathbf{R}(\boldsymbol{x}_\vartheta,\boldsymbol{x}_\vartheta)$, where $\mathbf{R}(\boldsymbol{x}_\vartheta,\boldsymbol{x}_\vartheta)$ is an auto-correlation matrix as a function of ϕ and the locations of $\boldsymbol{\vartheta}$, $\boldsymbol{x}_\vartheta$. The distribution of $\boldsymbol{Y}$ conditioned on the structural parameters and anchors $p(\boldsymbol{y}|\boldsymbol{\theta},\boldsymbol{\vartheta})$ is a multivariate Gaussian distribution with conditional mean $\mu_{Y|\vartheta}$ and conditional covariance $\sigma^2\mathbf{R}_{Y|\vartheta}$, where the mean and covariance conditioned on the anchors are defined as

$$\mu_{Y|\vartheta} = \mu + \mathbf{R}(\boldsymbol{x},\boldsymbol{x}_\vartheta)\mathbf{R}(\boldsymbol{x}_\vartheta,\boldsymbol{x}_\vartheta)^{-1}(\boldsymbol{\vartheta}-\mu),$$

$$\mathbf{R}_{Y|\vartheta} = \mathbf{R}(\boldsymbol{x},\boldsymbol{x}) - \mathbf{R}(\boldsymbol{x},\boldsymbol{x}_\vartheta)\mathbf{R}(\boldsymbol{x}_\vartheta,\boldsymbol{x}_\vartheta)^{-1}\mathbf{R}(\boldsymbol{x}_\vartheta,\boldsymbol{x}) \quad (4)$$

where $\mathbf{R}(\boldsymbol{x},\boldsymbol{x})$ is the auto-correlation matrix for $\boldsymbol{Y}$, and $\mathbf{R}(\boldsymbol{x},\boldsymbol{x}_\vartheta)$ is the cross-correlation matrix between $\boldsymbol{Y}$ and $\boldsymbol{\vartheta}$.

3.1.3 Specification of likelihood

We consider the data z consisting of L injection tests (l=1, 2, ..., L). We divide z into L segments as $z=\{z_1,z_2,\ldots,z_L\}$, where z_l is the vector containing the zeroth-order moments at multiple observation wells in the l-th injection test. The governing equation and the temporal moment formulation are shown in Appendix A.

In order to determine the likelihood $p(z|\boldsymbol{\theta},\boldsymbol{\vartheta})$, we first compute the likelihood in each injection test $p(z_l|\boldsymbol{\theta},\boldsymbol{\vartheta})$. Since we have observed that the zeroth-order moments are

approximately Gaussian, we use a multivariate Gaussian distribution for the likelihood estimation. Although nonparametric density estimation is available, the Gaussian-likelihood estimation is computationally advantageous as the dimension of the data increases.

Using the ensemble of $\tilde{z}_l$ simulated on the multiple fields conditioned on each parameter set $\{\boldsymbol{\theta}, \boldsymbol{\vartheta}\}$, we calculate the mean and covariance to determine $p(z_l|\boldsymbol{\theta}, \boldsymbol{\vartheta})$ (Robert and Casella, 1999). When we include the multiple injection tests in the inversion, we multiply the likelihoods of multiple tests, according to Eq. (3), to obtain the likelihood for the entire data $p(\boldsymbol{z}|\boldsymbol{\theta}, \boldsymbol{\vartheta})$. We have observed that the zeroth-order moments from the different injection tests are not strongly correlated, so that we may use Eq. (3).

3.1.4 Placement of anchors

The success of MAD depends on placement of the anchors from two reasons. First, careful placement of anchors will maximize their ability to extract information from observations. Second, the computational burden is linked to the number of anchors, and a smaller number would improve computational efficiency. Hence, we need to place anchors (1) at sensitive locations to the data, (2) to capture local features of the field, and (3) according to the goal of the inversion. A detailed discussion is available in Rubin et al. (2009). Here we discuss only the issues relevant to our inversion.

First, to find sensitive locations, we refer to the sensitivity analysis. Li et al. (2005) has formulated the sensitivity of zeroth-order moments to a ln$T$ value at a specific location, using the adjoint-state method (Sun, 1994). In their formulation, sensitivity is high around observation well locations, which is consistent with findings by Castagna and Bellin (2009) and Vasco et al. (2000).

Second, to capture heterogeneity, we would ideally have more than one anchor per integral scale. Although the real integral scale is not known in advance, we may consider the minimum possible integral scale at the site. Anchors outside the well plot, far from any of the observation wells, are not effective in resolving spatial heterogeneity, so that we need fewer anchors outside the well plot.

Third, to achieve our goal, which is to obtain the ln$T$ values at the EBF well locations, we need anchors at all EBF well locations. All the EBF wells are used as observation wells during the injection tests, so that we do not need additional anchors for this purpose.

3.2 3-D geostatistical model for hydraulic conductivity field

The 2-D inversion of the injection tests yielded a joint distribution of ln$T$ at the EBF well locations. Since we placed anchors at all those locations, we can use the anchor distribution directly. Let us denote the ln$T$ values at the EBF well location by a vector $\boldsymbol{\vartheta}_{\mathrm{EBF}}$, which is a subset of $\boldsymbol{\vartheta}$. Marginalizing the other parameters leads to the posterior distribution of $\boldsymbol{\vartheta}_{\mathrm{EBF}}$ conditioned on the injection test data z as $p(\boldsymbol{\vartheta}_{\mathrm{EBF}}|z)$.

Let $K(x_1, x_2, x_3)$ and $k(x_1, x_2, x_3)$ be the absolute and relative K values at the location $\boldsymbol{x}=(x_1, x_2, x_3)$ in the 3-D domain, respectively. Based on Javandel and Witherspoon (1969), we have the correlation between the absolute and relative K values as $K(x_1, x_2, x_3) = T(x_1, x_2)k(x_1, x_2, x_3)/b(x_1, x_2)$ (Moltz et al., 1994; Fienen et al., 2004), where $b(x_1, x_2)$ is the aquifer thickness at the horizontal location (x_1, x_2). We can determine the natural log-conductivity $u=\ln K$ at (x_1, x_2, x_3) by

$$u(x_1,x_2,x_3) = \vartheta_{\mathrm{EBF}}(x_1,x_2) - \ln b(x_1,x_2) + \ln k(x_1,x_2,x_3). \quad (5)$$

We use a N-vector $\boldsymbol{k}$ containing all the relative conductivity values from the EBF data at N locations, and a N-vector $\boldsymbol{u}$ containing all the ln$K$ values at the same locations as $\boldsymbol{k}$. Equation (5) allows us to combine $\boldsymbol{k}$ and $p(\boldsymbol{\vartheta}_{\mathrm{EBF}}|z)$ into $p(\boldsymbol{u}|\boldsymbol{k},z)$, which is the distribution of $\boldsymbol{u}$ conditioned on both the injection test data and the EBF data.

We construct a 3-D geostatistical model, assuming that $\boldsymbol{u}(\boldsymbol{x})$ follows a multivariate Gaussian distribution with mean β and covariance $(\eta^2\mathbf{R}(\boldsymbol{x},\boldsymbol{x}) + \nu^2\mathbf{I})$, where η^2 is the variance of variability in ln$K$, $\mathbf{R}(\boldsymbol{x}, \boldsymbol{x})$ is the auto-correlation matrix for $\boldsymbol{u}(\boldsymbol{x})$ as a function of the locations $\boldsymbol{x}$, the horizontal integral scale λ_h and the vertical integral scale λ_v, $\mathbf{I}$ is the identity matrix of order N, and ν^2 is the nugget, which represents the EBF measurement errors. The structural parameter vector of the 3-D geostatistical model

is $\{\beta, \eta^2, \lambda_h, \lambda_v, \nu^2\}$. Our goal here is to obtain a joint posterior distribution of the parameters conditioned on both data $\{\boldsymbol{k}, \boldsymbol{z}\}$ through $\boldsymbol{u}$:

$$p\left(\beta, \eta^2, \lambda_h, \lambda_v, \nu^2 | \boldsymbol{k}, \boldsymbol{z}\right) = \int p\left(\beta, \eta^2, \lambda_h, \lambda_v, \nu^2 | \boldsymbol{u}\right) p\left(\boldsymbol{u} | \boldsymbol{k}, \boldsymbol{z}\right) du. \quad (6)$$

For the prior distribution of the parameters, we assume the Jeffreys prior for the mean and variance (Jeffreys, 1946), which is the least informative prior for those two parameters. The prior distribution of all the structural parameters is defined as

$$p\left(\beta, \eta^2, \lambda_h, \lambda_v, \nu^2\right) \propto \frac{1}{\eta^2} \pi\left(\lambda_h, \lambda_v, \nu^2\right). \quad (7)$$

For the rest of the prior distribution $\pi(\lambda_h, \lambda_v, \nu^2)$, we use an independent uniform distribution for each of $\{\lambda_h, \lambda_v, \nu^2\}$ bounded by each set of the minimum and maximum possible values. Following Diggle and Ribeiro (Chapter 6, 2006), we obtain an analytical expression for $p(\beta, \eta^2, \lambda_h, \lambda_v, \nu^2 | \boldsymbol{u})$ (Appendix B).

3.3 Implementation

3.3.1 Organization of constant-rate injection test data

To demonstrate our approach, we used four out of the fourteen constant-rate injection tests at the site, the locations of which are well balanced within the IFRC site. Figure 2 shows the configuration of the injection and observation wells for each of the four tests.

For each test, we calculated the zeroth-order moments at multiple observation wells by integrating the pressure buildup curves. Since the well field is located near the Columbia River, the water table changes according to the river stage fluctuation. Since the change was mostly linear within 20 min after the injection starts, we removed the ambient head contribution by linear interpolation. We determined the standard deviation of measurement error based on the resolution of the instrument, 0.003 ft (0.09 cm) by integrating it over the injection duration.

3.3.2 Prior distribution for MAD inversion

For the prior distribution of the 2-D structural parameters $\boldsymbol{\theta}$, we used three independent uniform distributions bounded by the minimum and maximum values, as are shown in Table 1. The prior distributions of each parameter sufficiently cover possible values from the historical data at the site (Meyer et al., 2007) or literature values for similar geological formations (Rubin, 2003). The uniform distributions are considered to be less informative than Gaussian distributions, which have been commonly used in the Bayesian geostatistical inversion (Li et al., 2005). Three thousand sets of $\boldsymbol{\theta}$ are generated from $p(\boldsymbol{\theta})$ using a quasi Monte-Carlo sampling method (Krommer and Ueberhuber, 1998).

As is shown in Fig. 3, we placed 44 anchors at all the well locations inside the well plot and at sparse locations outside the well plot, following the discussion in Sect. 3.1.4. For each set of $\boldsymbol{\theta}$, we generated 12 sets of anchors $\boldsymbol{\vartheta}$ from $p(\boldsymbol{\vartheta}|\boldsymbol{\theta})$, so that the number of prior parameter sets $\{\boldsymbol{\theta}, \boldsymbol{\vartheta}\}$ is 36 000. We have observed that those numbers are sufficient to obtain convergence both in the prior and posterior distributions.

3.3.3 Forward simulation in MAD

Figure 3 shows the computational domain used for the forward simulations. For simulating the zeroth-order temporal moments on multiple random fields, we followed the approach by Firmani et al. (2008), since the mathematical expression for the zeroth-order moments is the same as the one for steady-state flow toward a well.

Firmani et al. (2008) determined the grid and domain sizes according to the integral scale. Although the integral scale is unknown in our framework, we have the minimum or maximum possible integral scale ($\phi_{\min}$ and $\phi_{\max}$), which are the upper and lower bounds of the uniform prior distribution. We used these two values so that any possible integral scale can satisfy the requirement for the domain and grid sizes.

The computational grid size is uniform equal to $0.2\phi_{\min}$ in both x_1 and x_2 directions during the field generation. During the flow simulations, the grid blocks near the

injection well are divided into non-uniform grid from $0.04\phi_{\min}$ at the injection well location to $0.2\phi_{\min}$ at a distance of $0.8\phi_{\min}$, satisfying the condition that the ratio between the neighboring block size should not exceed 1.5 (Firmani et al., 2008). We determined the domain size such that the observation wells were $2\phi_{\max}$ away from the boundaries to reduce the boundary effect. During the flow simulation, we added another buffer zone with width $\phi_{\max}$ and uniform T equal to T_G between the field and the boundaries for further reducing the boundary effect. We intended to satisfy $3\phi_{\max}$ between any observation wells and the boundaries, based on the theory developed by Rubin and Dagan (1988).

We used the SGSIM program from GSLIB (Deutsch and Journel, 1992) to generate spatially correlated Gaussian random fields conditioned on each set of $\{\boldsymbol{\theta}, \vartheta\}$. We then simulated zeroth-order moments on each field, using a finite-element method with linear elements. We used 250 realizations of random fields and moments for each $\{\boldsymbol{\theta}, \vartheta\}$ to calculate the likelihood $p(z|\boldsymbol{\theta}, \vartheta)$. The 9 000 000 forward simulations took 60 000 computational hours in total. It was divided into several batches, and used up to 9000 cores on the Franklin supercomputer at the National Energy Research Scientific Computing Center (Berkeley, USA), each core of which is a 2.3 GHz single AMD Opteron processor.

3.3.4 3-D geostatistical model

After calculating $p(\boldsymbol{u}|\boldsymbol{k}, z)$ from the relative K values and $\ln T$ values at the EBF well locations, we generated a thousand sets of $\boldsymbol{u}$ from $p(\boldsymbol{u}|\boldsymbol{k}, z)$. For each set of $\boldsymbol{u}$, we computed a posterior distribution $p(\beta, \eta^2, \lambda_h, \lambda_v, v^2|\boldsymbol{u})$, based on the uniform prior distribution of λ_h, λ_v and v^2 bounded by the values shown in Table 2. We then integrate the distribution numerically to determine $p(\beta, \eta^2, \lambda_h, \lambda_v, v^2|\boldsymbol{k}, z)$.

4 Results and discussion

We first tested the MAD and numerical setting in a synthetic study for inverting the injection test data, and then we applied it to the actual data at the Hanford site.

4.1 Synthetic study

We generated a synthetic reference ln$T$ field with a geostatistical parameter set $\{\mu, \sigma^2, \phi\}=\{-1.8, 1.5, 20\}$, shown in Fig. 4. We obtained maximum likelihood estimates of the true parameters as $\{-1.76, 1.46, 20.0\}$, with near-exhaustive sampling (one out of every five points) (GeoR package by Ribeiro and Diggle, 2001). We then calculated the zeroth-order moments on the reference field and superposed a zero-mean independent Gaussian measurement error, which has the same variance as the actual data from the study site.

Our inversion process is based on the same sets of injection and observation wells as the actual experiments conducted at the IFRC site (Fig. 3). We also combined the different number of injection tests in the inversion: one injection test (injection at Well 2-18), two tests (injection at Well 2-09 and 2-34), three tests (injection at Well 2-09, 2-24, and 3-24), and four tests (injection at Well 2-09, 2-18, 2-24, and 3-24). They are compared to show the effect of additional information from the multiple tests.

Figure 5 shows the marginal posterior distributions of the geostatistical structural parameters $\{\mu, \sigma^2, \phi\}$ based on the various number of tests, with their corresponding true values. While the mean has a symmetric Gaussian-like distribution, the variance and scale has broad and skewed distributions. The results are improved with increasing number of tests up to three tests, i.e. the posterior distributions become narrower and biases are reduced. The improvement by additional tests is more significant for the variance and scale, which suggests that the estimation of variance and scale requires more observations. The improvement, however, is not significant from three to four tests, which would suggest that the effect of increasing number of tests could be saturated due to the measurement errors and redundancy of information in the data.

To examine the effect of anchors and evaluate the random fields generated based on the posterior distributions, we generated 200 000 fields from the posterior distribution of parameters (5000 posterior sample parameter sets with 40 fields per parameter set), and compared the ensemble with the true field. Two cases are studied: one based on a single test (injection at Well 2-18) and the other based on three tests (injection at Well 2-09, 2-24, 3-24).

Figure 6 shows the mean and 98% confidence interval of the predicted ln$T$ fields along the centerline of the well field as shown in Fig. 4 (Line A–B). The mean field along the line follows a general trend of the true field, especially so with more tests assimilated. If there were no anchors, the mean field would be a flat line at the global mean. Therefore the deviation from such a straight line is attributed to anchors that capture local heterogeneity. The uncertainty bounds are found to be tighter near the center, where there are more observation wells and more anchors. Increasing the number of tests not only reduces the uncertainty, but also reduces the bias by moving the mean field closer to the true field.

4.2 IFRC data analysis

After we gained confidence from the synthetic study, we applied the same scheme to the data from the IFRC site. Since the true values are unknown in this case, we validated the posterior distribution by predicting a subset of the data and comparing it to the actual data. The data used for validation was not used for inversion.

Figure 7 shows the marginal posterior distributions of the three structural parameters for the ln$T$ field at the IFRC site. The three plots show a similar feature as the synthetic study in Fig. 5 such as a Gaussian-like distribution for the mean, broad and skewed distributions for the variance and scale, and the effect of increasing the number of injection tests in the inversion.

For the first validation, we generated 200 000 fields (5000 posterior sample parameter sets with 40 fields per parameter set) based on one test (injection at Well 2-09), two tests (injection at Well 2-09 and 3-24) and three tests (injection at Well 2-09, 2-24

and 3-24). We then predicted the zeroth-order temporal moments in the injection test at Well 2-18, which was not the part of the estimation. Figure 8 shows the marginal predictive distributions of the observed moments at two observation wells, based on one, two and three tests, compared with the actual data. The true value is contained within the range of values defining the predictive distributions, and the combination of the three tests improves the prediction by narrowing the distributions.

As another validation, we obtained the maximum a posteriori (MAP) estimate of T_G=exp(μ) in Fig. 7, which is 0.52 $m^2 s^{-1}$. We compared this value with T_G estimated from the Cooper-Jacob analysis (Sánchez-Villa et al., 1999), in which we fitted the late-time pressure-buildup curves at multiple observation wells in the injection test at Well 2-18. The 95% confidence bound of T_G was 0.38–0.57 $m^2 s^{-1}$. As we expected, our estimate of T_G corresponded to the estimates based on conventional analysis for a large-scale injection test.

Figure 9 shows the marginal distribution for the 3-D geostatistical structural parameters conditioned on the EBF data and injection test data. We can see that the marginal posterior distributions of the structural parameters are skewed except for the mean, which suggests that the entire distribution is necessary to quantify the parameter uncertainty.

We also compared the distributions based on the different number of tests included in the injection test inversion. As it turned out, increasing the number of injection tests did not reduce the parameter uncertainty in the 3-D structural parameters as significantly as in the 2-D parameters. This is because the uncertainty and sparseness of the EBF data obscures additional information of the increasing number of injection tests in the 3-D spatial inference. The horizontal scale especially has broad distribution, due to the relatively insufficient number of horizontal lags among the EBF wells.

HESSD
7, 2017–2052, 2010

5 Summary

In this paper, we presented a Bayesian approach for characterizing a 3-D K field by assimilating the EBF and constant-rate injection tests. We employed a two-step approach – first inverting the constant-rate injection test data for obtaining the local T values at the EBF well locations, and then converting the EBF data to local K values for the 3-D characterization.

For the injection test inversion, we used MAD, which is a newly developed Bayesian geostatistical inversion framework. This method enables us to obtain an entire posterior distribution for parameters, rather than best-fitted values, so that we properly quantify parameter uncertainty. Using MAD, we inverted zeroth-order moments of pressure buildups at multiple observation wells, which can eliminate uncertainty in a storage coefficient, as well as significantly reduce computational cost.

In a synthetic study, we first showed that MAD could successfully infer the geostatistical parameters and predict the ln$T$ field. As we included more tests, we could further reduce the uncertainty, and better capture the local heterogeneity.

By applying the method to the actual data, we obtained the posterior distribution of geostatistical structural parameters and the anchor values of the ln$T$ field for the Hanford 300 Area IFRC site. We validated the result using the predictive distribution of the zeroth-moments in the injection test that were not part of the inversion. In addition, the MAP estimate of the mean ln$T$ coincided with the T_G value obtained from the Cooper-Jacob analysis, which confirmed our method's consistency with conventional pumping test analysis.

We then combined the relative K values from the EBF data with the distribution of ln$T$ at EBF wells so that we obtain the distribution of the depth-discrete absolute K in the 3-D domain. The uncertainty in T is consistently carried on into the K values as a probability distribution. We thus constructed a 3-D geostatistical model for the K field using the model-based Bayesian geostatistical approach. The resulting distribution showed a skewed distribution except for the mean, which may not be captured by other optimization-based approaches.

Appendix A

Temporal moments

According to Li et al. (2004) and Zhu and Yeh (2005), the k-th-order temporal moment $m_k(\boldsymbol{x})$ for a pressure-buildup curve $s(\boldsymbol{x},t)$ is defined by

$$m_k(\boldsymbol{x}) = \int_0^\infty t^k s(\boldsymbol{x},t)\,dt. \tag{A1}$$

In this study, we use only the zero-order moment $m_0(\boldsymbol{x})$ for the inversion to characterize the T field, which can exclude uncertainty in the storage coefficient and avoid an alias effect of the storage-coefficient uncertainty to the T field. Under the constant injection condition, we obtain $m_0(\boldsymbol{x})$ from the equation:

$$\nabla\cdot(T\nabla m_0) + \tau Q\delta\left(\boldsymbol{x}-\boldsymbol{x}_p\right) = 0, \tag{A2}$$

with the boundary condition at the Drichlet boundary Γ_{dri} as,

$$m_0 = 0, \quad \text{at} \quad \Gamma_{\text{Dri}}, \tag{A3}$$

where $T(\boldsymbol{x})$ is the depth-integrated T value, τ is the injection duration, Q is the constant injection rate and $\boldsymbol{x}_p$ is the injection well location. The Drichlet boundary condition was imposed at the nearest observation well location, where the m_0 value is known, in the same manner that Firmani et al. (2006) imposed a boundary condition at the injection well location. Note that Eq. (A2) is the same as the one for determining hydraulic head under steady-state flow with a constant injection rate τQ.

HESSD
7, 2017–2052, 2010

Appendix B

Bayesian model-based geostatistics for 3-D structural parameters

According to Diggle and Ribeiro (Chapter 6, 2006), we calculate the posterior distributions of the 3-D geostatistical structural parameters conditioned on $\boldsymbol{u}(\boldsymbol{x})$. First, the scales λ_h and λ_v and nugget variance v^2 depend only on $\boldsymbol{u}$ as

$$p\left(\lambda_h,\lambda_v,v^2|\boldsymbol{u}\right) \propto \pi\left(\lambda_h,\lambda_v,v^2\right)|V_{\hat{\beta}}|^{\frac{1}{2}}|\left(\mathbf{R}+v^2\mathbf{I}\right)|^{-\frac{1}{2}}\left(S^2\right)^{-\frac{N-1}{2}}, \tag{B1}$$

where each term is defined as the follows:

$$V_{\tilde{\beta}} = \left[\mathbf{1}^T\left(\mathbf{R}+v^2\mathbf{I}\right)^{-1}\mathbf{1}\right]^{-1},$$

$$\hat{\beta} = V_{\tilde{\beta}}\mathbf{1}^T\left(\mathbf{R}+v^2\mathbf{I}\right)^{-1}\boldsymbol{u},$$

$$S^2 = \frac{\boldsymbol{u}^T\left(\mathbf{R}+v^2\mathbf{I}\right)^{-1}\boldsymbol{u}-\hat{\beta}^T V_{\tilde{\beta}}^{-1}\hat{\beta}}{N-1}, \tag{B2}$$

where N is the dimension of $\boldsymbol{u}$, $\mathbf{R}$=$\mathbf{R}(\boldsymbol{x},\boldsymbol{x})$ is the auto-correlation matrix for $\boldsymbol{u}$, $\mathbf{1}$ is the N-vector with all the elements equal to one, and $\mathbf{I}$ is the identity matrix. The variance η^2 follows an inverse-scaled χ^2 distribution χ^2_{Scl} with $(N-1)$ degrees of freedom and a scale parameter equal to S^2:

$$p\left(\eta^2|\lambda_h,\lambda_v,v^2,\boldsymbol{u}\right) \sim \chi^2_{Scl}\left(N-1,S^2\right). \tag{B3}$$

The mean follows a normal distribution with mean $\hat{\beta}$ and variance $\eta^2 V_{\hat{\beta}}$:

$$p\left(\beta|\eta^2,\lambda_h,\lambda_v,v^2,\boldsymbol{u}\right) \sim N\left(\hat{\beta},\eta^2 V_{\hat{\beta}}\right). \tag{B4}$$

We multiply Eqs. (B1), (B3), (B4), and the prior distribution to determine $p(\beta,\eta^2,\lambda_h,\lambda_v,v^2|\boldsymbol{u})$ in Eq. (6).

HESSD
7, 2017–2052, 2010

Acknowledgements. This study has been funded by the US DOE Office of Biological and Environmental Research, Environmental Remediation Science Program (ERSP), through DOE-ERSP grant DE-FG02-06ER06-16 as part of Hanford 300 Area Integrated Field Research Challenge Project. This research used resources of the National Energy Research Scientific Computing Center, which is supported by the Office of Science of the US Department of Energy under Contract No. DE-AC02-05CH11231.

References

Bjornstad, B. N., Horner, J. A., Vermuel, V. R., Lanigan, D. C., and Thorne, P. D.: Borehole completion and conceptual hydrolgeologic model for the IFRC Well Field, 300 Area, Hanford Site PNNL-18340, Pacific Northwest National Laboratory, Richland, Washington, 2009.

Castagna, M. and Bellin, A.: A Bayesian approach for inversion of hydraulic tomographic data, Water Resour. Res., 45, W04410, doi:10.1029/2008WR007078, 2009.

Chen, J., Hubbard, S., and Rubin, Y.: Estimating the hydraulic conductivity at the South Oyester Site from geophysical tomographic data using Bayesian techniques on the normal linear regression model, Water Resour. Res. 37(6), 103–1613, 2001.

Copty, N., Rubin, Y., and Mavko, G.: Geophysical-Hydrological Identification of Field Permeabilities Through Bayesian Updating, Water Resour. Res., 29(8), 2813–2825, 1993.

Dagan, G., Lessoff, S. C., and Fiori, A.: Is transmissivity a meaningful property of natural formations? Conceptual issues and model development, Water Resour. Res., 45, W03425, doi:10.1029/2008WR007410, 2009.

De Barros, F. P. J., Rubin, Y., and Maxwell, R. M.: The concept of comparative information yield curves and its application to risk-based site characterization, Water Resour. Res., 45, W06401, doi:10.1029/2008WR007324, 2009.

Diggle, P. J. and Ribeiro, P. J.: Bayesian inference in Gaussian model-based geostatistics, Geographical and Environmental Modeling, 6(2), 129–146, doi:10.1080/1361593022000029467, 2002.

Diggle, P. J. and Ribeiro, P. J.: Model-based geostatistics, Springer, New York, U.S.A., 2006.

Fienen, M. N., Kitanidis, P. K., Walton, D., and Jardine, P.: An application of Bayesian inverse methods to vertical deconvolution of hydraulic conductivity in a heterogeneous aquifer at Oak Ridge National Laboratory, Math. Geol., 36(1), 101–126, 2004.

Firmani, G., Fiori, A., and Bellin, A.: Three-dimensional numerical analysis of steady state pumping tests in heterogeneous confined aquifers, Water Resour. Res., 42, W03422, doi:10.1029/2005WR004382, 2006.
Handcock, M. S. and Stein, M. L.: A Bayesian analysis of kriging, Technometrics, 35(4), 403–410, 1993.
Hou, Z. and Rubin, Y.: On minimum relative entropy concepts and prior compatibility issues in vadose zone inverse and forward modeling, Water Resour. Res., 41, W12425, doi:10.1029/2005WR004082, 2005.
Javandel, I. and Witherspoon, P. A.: A method of analyzing transient fluid flow in multilayered aquifers, Water Resour. Res., 5, 856–869, 1969.
Jeffreys, H.: An invariant form for the prior probability in estimation problems, Proc. R. Soc. Lon. Ser.-A, 186, 453–461, 1946.
Kitanidis, P. K.: Parameter uncertainty in estimation of spatial functions: Bayesian analysis, Water Resour. Res., 22(4), 499–507, 1986.
Kitanidis, P. K.: Quasi-linear geostatistical theory for inversing, Water Resour. Res., 31(10), 2411–2419, 1995.
Krommer, A. R. and Ueberhuber, C. W.: Computational integration, Society for Industrial and Applied Mathematics, Philadelphia, 1998.
Li, W., Nowak, W., and Cirpka, O. A.: Geostatistical inverse modeling of transient pumping tests using temporal moments of drawdown, Water Resour. Res., 41, W08403, doi:10.1029/2004WR003874, 2005.
Li, W., Englert, A., Cirpka, O. A., and Vereecken, H.: Three-dimensional geostatistical inversion of flowmeter and pumping test data, Ground Water, 46(2), 193–201, 2008.
Liu, X., Illman, W. A., Craig, A. J., Zhu, J., and Yeh, T.-C. J.: Laboratory sandbox validation of transient hydraulic tomography, Water Resour. Res., 43, W05404, doi:10.1029/2006WR005144, 2007.
McLaughlin, D., Reid, L. B., Li, S. G., and Hyman, J.: A stochastic method for characterizing groundwater contamination, Ground Water, 31(2), 237–249, 1993.
Maxwell, R., Kastenberg, W. E., and Rubin, Y.: A methodology to integrate site characterization information into groundwater-driven health risk assessment, Water Resour. Res., 35(9), 2841–2885, 1999.

Meyer, P. D., Ye, M., Rockhold, M. L., Neuman, S. P., and Cantrell, K. J.: Combined estimation of hydrogeologic conceptual model, parameter, and scenario uncertainty with application to uranium transport at the Hanford Site 300 Area, PNNL-16396, Pacific Northwest National Laboratory, Richland, WA, 2007.

Molz, F. J., Boman, G. K., Young, S. C., and Waldrop, W. R.: Borehole flowmeters: Field application and data analysis, J. Hydrol., 163(4), 347–371, 1994.

Neuman, S. P.: Analysis of pumping test data from anisotropic unconfined aquifers considering delayed gravity response, Water Resour. Res., 11(2), 329–342, 1975.

Ribeiro, P. J. and Diggle P. J.: geoR: a package for geostatistical analysis R-NEWS, 1(2), 15–18, 2001.

Rockhold, M. L., Vermeul, V. R., Mackley, R. D., Fritz, B. G., Mendoza, D. P., Newcomer, E. M., Newcomer, D. R., Murray, C. J., and Zachara, J. M.: Hydrogeologic characterization of the Hanford 300 Area Integrated Field Research Challenge site and numerical modeling of the first aquifer tracer Test, Ground Water, submitted, 2010.

Rubin, Y. and Dagan, G.: Stochastic analysis of the effects of boundaries on spatial variability in groundwater flows: 1. Constant head boundary, Water Resour. Res., 24(10), 1689–1697, 1988.

Rubin, Y: Applied Stochastic Hydrogeology, Oxford Univ. Press, Oxford, U.K., 2003.

Rubin, Y., Chen, X., Murakami, H., and Hahn, M. S.: A Bayesian approach for inverse modeling, data assimilation and conditional simulation of spatial random fields, Water Resour. Res, submitted, 2009.

Sánchez-Vila, X., Meier, P. M., and Carrera, J.: Pumping tests in heterogeneous aquifers: An analytical study of what can be obtained from their interpretation using Jacob's method, Water Resour. Res., 35(4), 943–952, 1999.

Scheibe, T. D., Chien, Y. J., and Radtke J.: Use of quantitative models to design microbial transport experiments in a sandy aquifer, Ground Water, 39(2), 210–222, 2001.

Scheibe, T. D. and Chien, Y. J.: An evaluation of conditioning data for solute transport prediction, Ground Water, 41(2), 128-141, 2003.

Sun, N.-Z.: Inverse Problems in Groundwater Modeling, Theor. Appl. Transp. Porous Media, vol. 6, Kluwer Acad., Dordrecht, Netherlands, 1994.

Vasco, D. W., Keers, H., and Karasaki, K.: Estimation of reservoir properties using transient pressure data: An asymptotic approach, Water Resour. Res., 36(12), 3447–3465, 2000.

Vermeul, V. R., Bjornstad, B. N., Fritz, B. G., Fruchter, J. S., Mackley, R. D., Newcomer, D. R., Mendoza, D. P., Rockhold, M. L., Wellman, D. M., and Williams, M. D.: 300 Area uranium stabilization through polyphosphate injection: Final Report, PNNL-18529, Pacific Northwest National Laboratory, Richland, WA, 2009.

Young, S. C., Julian, H. E., Pearson, H. S., Molz, F. J., and Boman, G. K.: Application of the electromagnetic borehole flowmeter., US Environmental Protection Agency Research Report EPA/600/R-98/058, Ada, Oklahoma, 1998.

Zhang, Z. and Rubin, Y.: MAD: a new method for inverse modeling of spatial random fields with applications in hydrogeology, Eos Trans. AGU, 89(53), Fall Meet. Suppl., Abstract H44C-07, 2008a.

Zhang, Z. and Rubin, Y.: Inverse modeling of spatial random fields, unpublished manuscript, 2008b.

Zhu, J. and Yeh, T.-C. J.: Analysis of hydraulic tomography using temporal moments of drawdown recovery data, Water Resour. Res., 42, W02403, doi:10.1029/2005WR004309, 2006.

Table 1. The lower and upper bounds of the prior distribution for the structural parameters of the 2-D transmissivity field.

	Minimum	Maximum	Reference
Mean μ, $m^2\,s{-}1$ ($\ln T$)	−4.82*	2.17*	Meyer et al. (2007)
Variance, σ^2	0.5	3.0	Rubin (2003), Table 2.1 and Table 2.2
Scale, ϕ, m	8	30	Rubin (2003), Table 2.1 and Table 2.2

* The upper bound and lower bounds of K multiplied by the average aquifer thickness 7.62 m.

HESSD

7, 2017–2052, 2010

Table 2. The lower and upper bounds of the prior distribution for the horizontal scale, vertical scale and nugget variance of the 3-D hydraulic conductivity field.

	Minimum	Maximum	Reference
Horizontal scale, λ_h, m	8	50	Rubin (2003), Table 2.1 and Table 2.2
Vertical scale, λ_v, m	0.5	10	* saturated thickness <10 m
Nugget variance, ν^2	10^{-3}	0.25	* less than 50% in standard deviation

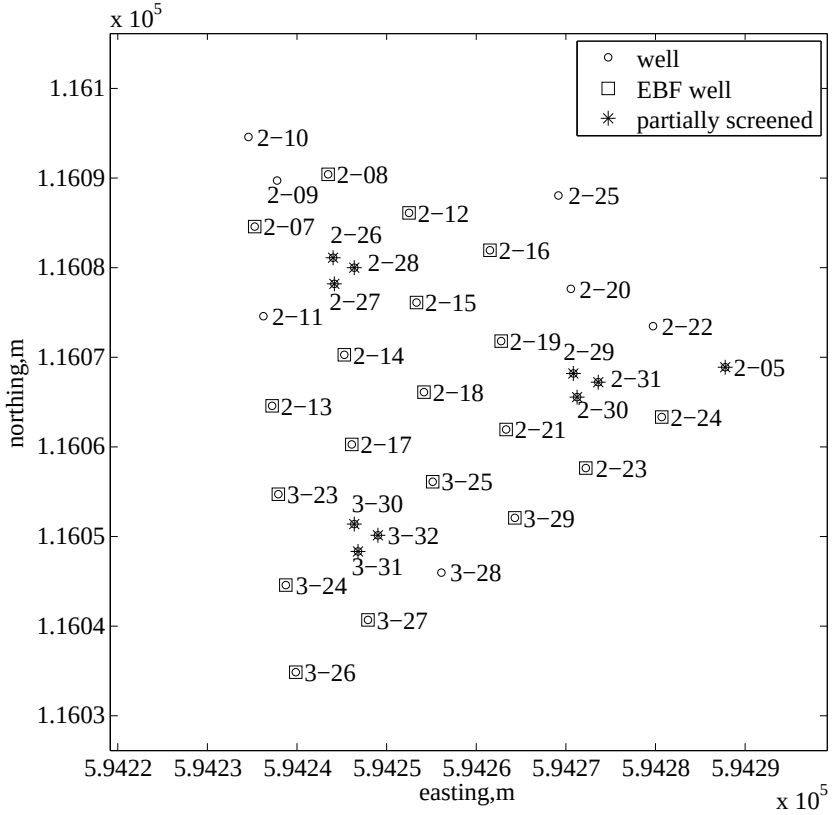

Fig. 1. Site map of the IFRC site (The coordinate system follows the convention used at the Hanford site).

HESSD
7, 2017–2052, 2010

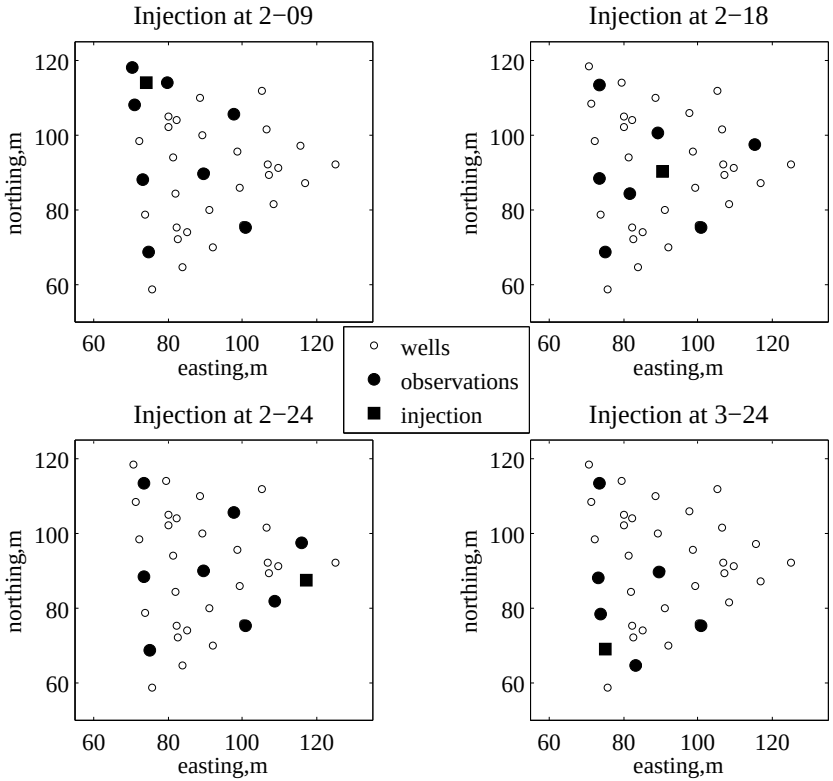

Fig. 2. Configuration of injection and observation wells in each test used in this paper. The reference point of local coordinates is at (594 164 m, 115 976 m) in the Hanford coordinates.

HESSD

7, 2017–2052, 2010

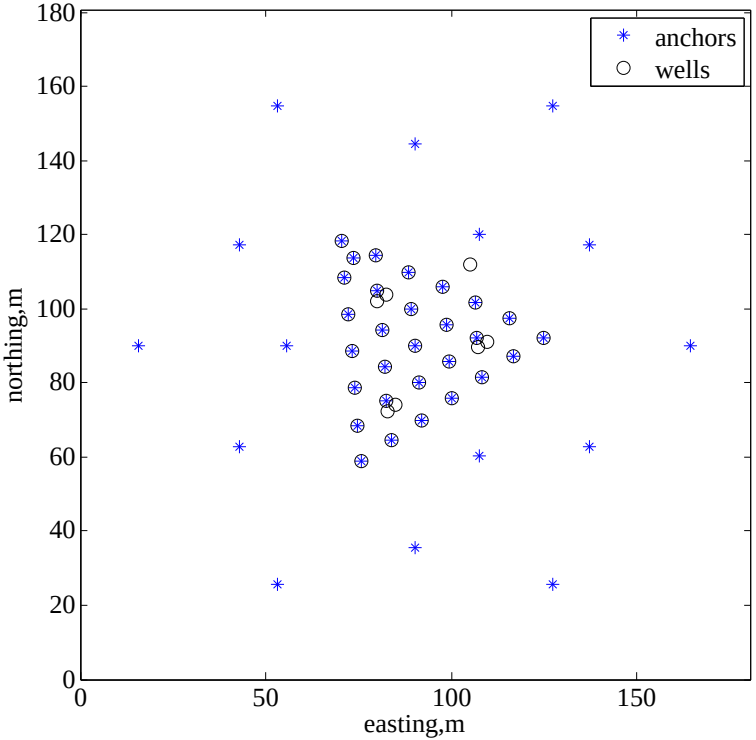

Fig. 3. Anchor locations in the domain for the constant-rate injection test inversion. The reference point of local coordinates is at (594 164 m, 115 976 m) in the Hanford coordinates.

HESSD

7, 2017–2052, 2010

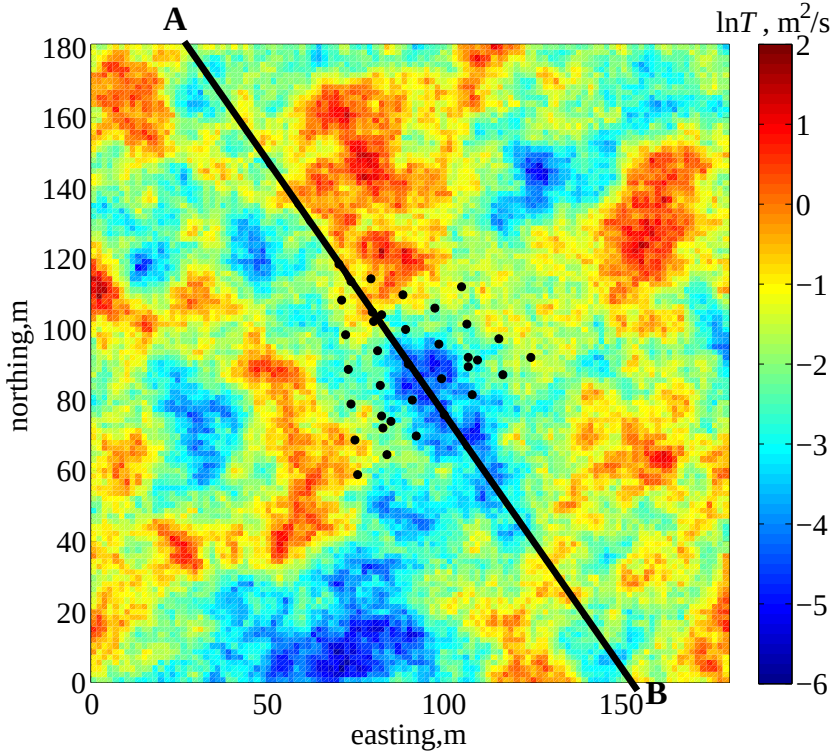

Fig. 4. Reference field for the synthetic study. Line A-B is used for the transect.

HESSD
7, 2017–2052, 2010

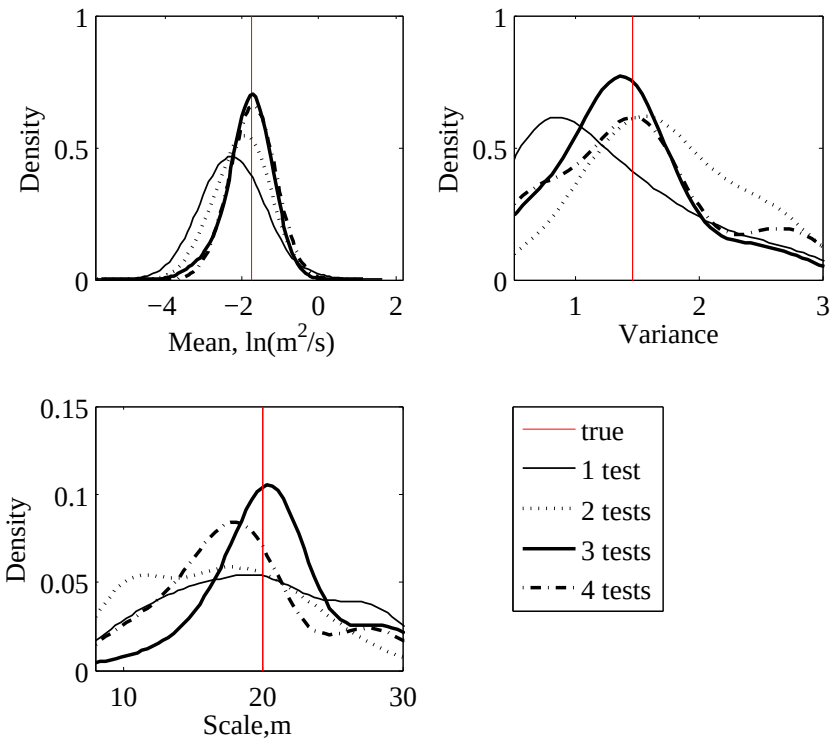

Fig. 5. Marginal posterior distributions of the structural parameters (mean, variance and scale) in the synthetic study, with their corresponding true values. The ones based on the different number of tests are compared.

HESSD

7, 2017–2052, 2010

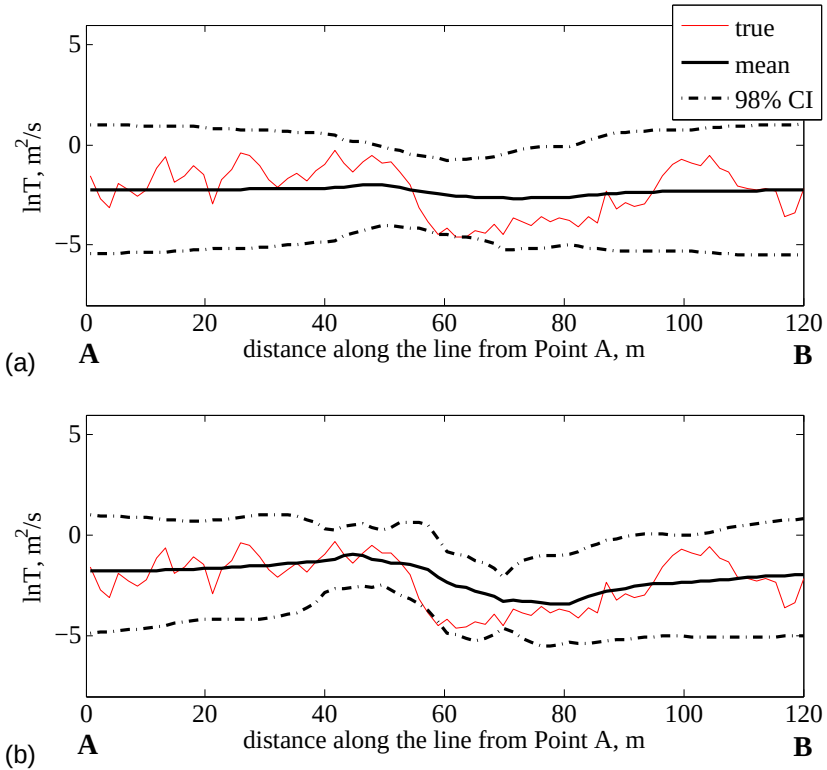

Fig. 6. Comparison among the reference field, the mean field and the 98% confidence interval of the generated fields, along the center line of the IFRC well field (Line A-B in Fig. 4), for the inversion based on **(a)** one test (injection at Well 2-18) and **(b)** three tests (injection at Wells 2-09, 2-24 and 3-24).

HESSD
7, 2017–2052, 2010

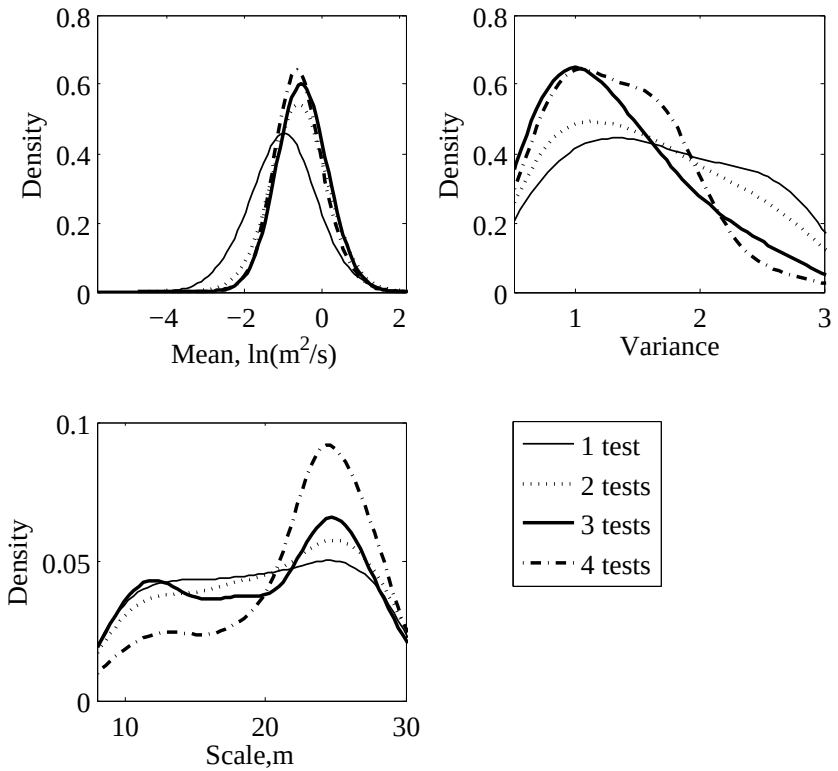

Fig. 7. Marginal posterior distributions of the structural parameters (mean, variance and scale) for the Hanford IFRC site data.

HESSD

7, 2017–2052, 2010

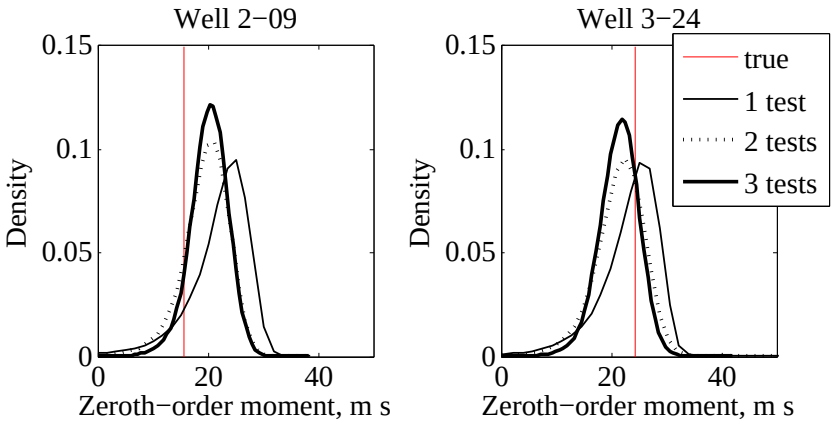

Fig. 8. Comparison between the zeroth-order moments observed at Well 2-09 and 3-24 in the injection test at Well 2-18, and predictive posterior distributions from the inversion, including the different number of injection test.

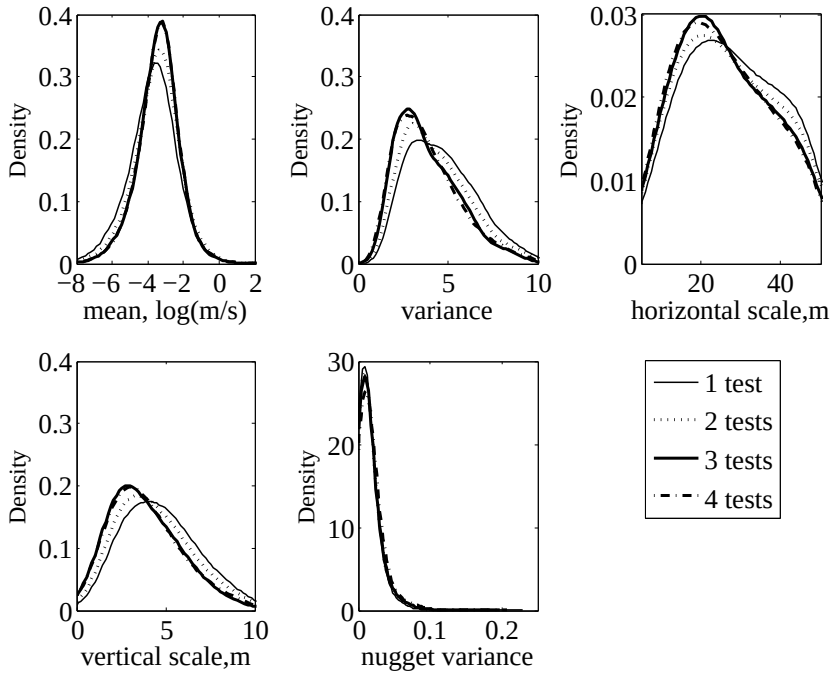

Fig. 9. Marginal posterior distribution of 3-D geostatistical structural parameters of K values at the Hanford IFRC site, based on the different number of injection test.