A simultaneous successive linear estimator and a guide for hydraulic tomography analysis

Jianwei Xiang,[1] Tian-Chyi J. Yeh,[1,2] Cheng-Haw Lee,[2] Kuo-Chin Hsu,[2] and Jet-Chau Wen[3]

Received 22 May 2008; revised 8 December 2008; accepted 23 December 2008; published 28 February 2009.

[1] In this study, a geostatistically based estimator is developed that simultaneously includes all observed transient hydrographs from hydraulic tomography to map aquifer heterogeneity. To analyze tomography data, a data preprocessing procedure (including diagnosing and wavelet denoising analysis) is recommended. A least squares approach is then introduced to estimate effective parameters and spatial statistics of heterogeneity that are the required inputs for the geostatistical estimator. Since wavelet denoising does not completely remove noise from observed hydrographs, a stopping criterion is established to avoid overexploitation of the imperfect hydrographs. The estimator and the procedures are then tested in a synthetic, cross-sectional aquifer with hierarchical heterogeneity and a vertical sandbox with prearranged heterogeneity. Results of the test indicate that with this estimator and preprocessing procedures, hydraulic tomography can effectively map hierarchical heterogeneity in the synthetic aquifer as well as in the sandbox. In addition, the study shows that using the estimated hydraulic conductivity and specific storage fields of the sandbox, the classic groundwater flow model accurately predicts temporal and spatial distributions of drawdown induced by an independent pumping event in the sandbox. On the other hand, the classic groundwater flow model yields less satisfactory results when equivalent homogeneous properties of the sandbox are used.

Citation: Xiang, J., T.-C. J. Yeh, C.-H. Lee, K.-C. Hsu, and J.-C. Wen (2009), A simultaneous successive linear estimator and a guide for hydraulic tomography analysis, *Water Resour. Res.*, *45*, W02432, doi:10.1029/2008WR007180.

1. Introduction

[2] Classical aquifer test which involves one pumping well and an observation well has been shown to yield ambiguously averaged hydraulic properties of an aquifer, which vary with the location of observation and pumping wells, and heterogeneity [*Wu et al.*, 2005; *Liu et al.*, 2007; *Straface et al.*, 2007; *Kuhlman et al.*, 2008]. To avoid obtaining the ambiguously averaged estimates and to provide high-resolution aquifer characterization, a new aquifer characterization technology, known as hydraulic tomography (HT), has recently been developed [e.g., *Tosaka et al.*, 1993; *Gottlieb and Dietrich*, 1995; *Vasco et al.*, 2000; *Yeh and Liu*, 2000; *Bohling et al.*, 2002; *Brauchler et al.*, 2003; Zhu and Yeh 2005, 2006]. Although the ability of HT remains to be fully assessed under field conditions, results from sandbox experiments by *Liu et al.* [2002], *Illman et al.* [2007, 2008], and *Liu et al.* [2007] are encouraging. These studies showed that the transient HT can identify not only the pattern of the heterogeneous hydraulic conductivity (K) field, but also the variation of specific storage (S_s) in the sandbox. Moreover, these estimated K and S_s fields from the HT sandbox experiments further predicted the drawdown evolution caused by a pumping test that was not used in the HT analysis. Likewise, a recent application of HT to a well field at Montalto Uffugo Scalo, Italy, produced an estimated transmissivity field that is deemed to be consistent with the geology of the site [*Straface et al.*, 2007]. *Bohling et al.* [2007] and *Li et al.* [2007] also showed promising results of HT in their field experiments.

[3] Most of HT analyses in the past have used the sequential successive linear estimator (SSLE) of *Yeh and Liu* [2000] or *Zhu and Yeh* [2005], which includes data sets from HT surveys sequentially. *Illman et al.* [2008] reported that the order of test data included in SSLE affected the final estimates. In addition, little efforts in the past have focused on investigation of effects of noise in well hydrographs on the analysis of HT and development of methods for removing the noise. In this paper, we thereby develop a geostatistically based method for identifying the subsurface heterogeneity pattern using all the data collected from a HT survey simultaneously, similar to the approaches by *Vesselinov et al.* [2001], *Li et al.* [2007], *Fienen et al.* [2008], or *Li et al.* [2008], which are built upon the quasi-linear geostatistical approach [*Kitanidis*, 1995]. Moreover, a criterion for the nonlinear estimator is developed to deter-

[1]Department of Hydrology and Water Resources, University of Arizona, Tucson, Arizona, USA.
[2]Department of Resources Engineering, National Cheng Kung University, Tainan, Taiwan.
[3]Department of Environmental and Safety Engineering, Research Center for Soil and Water Resources and Natural Disaster Prevention, National Yunlin University of Science and Technology, Touliu, Taiwan.

0043-1397/09/2008WR007180$09.00

mine the appropriate level of improvement of estimation when hydrographs are infested with noise. We also propose a procedure for preprocessing HT data before application of the estimator. This estimator and procedure were tested in a synthetic aquifer with hierarchical heterogeneity [*Barrash and Clemo*, 2002; *Ye et al.*, 2005], and then applied to a sandbox experiment, which involves unknown measurement errors and where our mathematical model may not correctly describe the flow process model errors.

2. Methodology

2.1. Governing Groundwater Flow Equations

[4] A strategy for modeling groundwater flow in saturated, heterogeneous, porous media with incomplete specification of aquifer characteristics is the stochastic conditional mean approach [e.g., *Yeh et al.*, 1996]. That is, the flow process is governed by the following partial differential equation involving conditional means:

$$\nabla \cdot [K(\mathbf{x})\nabla h] + Q(\mathbf{x}_p) = S_s(\mathbf{x})\frac{\partial h}{\partial t}, \tag{1}$$

subject to boundary and initial conditions:

$$h|_{\Gamma_1} = h_1,\ [K(\mathbf{x})\nabla h]\cdot \mathbf{n}|_{\Gamma_2} = q,\ \text{and } h|_{t=0} = h_0, \tag{2}$$

where, in equation (1), h is conditional effective total head (L), $\mathbf{x}$ is the spatial coordinate ($\mathbf{x} = \{x, y\}$, (L), and y represents the vertical coordinate and is positive upward) in the two-dimensional, cross-sectional aquifers examined in this paper, $Q(\mathbf{x}_p)$ is the pumping rate (1/T) at the location $\mathbf{x}_p$, $K(\mathbf{x})$ is the conditional effective saturated hydraulic conductivity (L/T), and $S_s(\mathbf{x})$ is the conditional effective specific storage (1/L). In equation (2), h_1 is the prescribed total head at Dirichlet boundary Γ_1, q is the specific discharge (L/T) at Neumann boundary Γ_2, $\mathbf{n}$ is a unit vector normal to Γ_2, and h_0 represents the initial total head. The definitions of variables in the conditional mean equations are identical to those in a deterministic groundwater flow equation if all the parameters, boundary and initial conditions are fully specified [*Yeh et al.*, 1996]. In this paper, these governing equations are used to simulate the flow field during the HT survey and are solved by a 2-D finite element model (VSAFT2) developed by *Yeh et al.* [1993].

2.2. Simultaneous Successive Linear Estimator

[5] Instead of incorporating data sequentially into the estimation as is done in SSLE, a simultaneous successive linear stochastic estimator (SimSLE) is developed to include all drawdown data from different pumping tests during a HT survey simultaneously to estimate hydraulic properties of aquifers. Simultaneous inclusion of the data offers some advantages over the SSLE approach (see discussion section). Below is a brief description of the SimSLE.

[6] With given unconditional mean and spatial covariance functions of the hydraulic properties (prior joint probability distribution, implicitly Gaussian), the SimSLE starts with cokriging (a stochastic linear estimator) to estimate the conditional expected value of the property conditioned on $f^*(\mathbf{x}_i)$ (i.e., perturbations of log hydraulic property, K and/or S_s) measured at ith location ($i = 1, \ldots N_f$, where N_f is the total number of f measurements) and the observed head at location $\mathbf{x}_j$ at time t_ℓ during kth pumping test, denoted by $h^*(k, \mathbf{x}_j, t_\ell)$. The linear estimator is

$$\hat{f}^{(1)}(\mathbf{x}_0) = \sum_{i=1}^{N_f}\lambda_{0i}f^*(\mathbf{x}_i) + \sum_{k=1}^{N_p}\sum_{j=1}^{N_h(k)}\sum_{\ell=1}^{N_t(k,j)} \cdot \mu_{0kj\ell}\left(h^*(k,\mathbf{x}_j,t_\ell) - h_e(k,\mathbf{x}_j,t_\ell)\right), \tag{3}$$

where $\hat{f}^{(1)}(\mathbf{x}_0)$ is the cokriged f value at location $\mathbf{x}_0$; $h_e(k, \mathbf{x}_j, t_\ell)$ is the simulated head at the observation location and time of the pumping test, based on effective properties of an equivalent homogeneous medium; N_p is the total number of pumping tests; $N_h(k)$ is the total number of observation locations for kth pumping test; $N_t(k, j)$ is the total number of head measurements in time at jth observation location during kth pumping test. The cokriging weight (λ_{0i}) represents contribution of measurement f^* at ith location to the estimate at location $\mathbf{x}_0$. The contribution to the estimate from the observed head $h^*(k, \mathbf{x}_j, t_\ell)$ is denoted by $\mu_{0kj\ell}$. These weights are obtained by solving the following system of equations:

$$\begin{aligned}
&\sum_{i=1}^{N_f}\lambda_{0i}R_{ff}(\mathbf{x}_m,\mathbf{x}_i) + \sum_{k=1}^{N_p}\sum_{j=1}^{N_h(k)}\sum_{\ell=1}^{N_t(k,j)} \cdot \mu_{0kj\ell}R_{hf}\left((k,\mathbf{x}_j,t_\ell),\mathbf{x}_m\right) = R_{ff}(\mathbf{x}_0,\mathbf{x}_m)\\
&\sum_{i=1}^{N_f}\lambda_{0i}R_{hf}\left((p,\mathbf{x}_r,t_q),\mathbf{x}_i\right) + \sum_{k=1}^{N_p}\sum_{j=1}^{N_h(k)}\sum_{\ell=1}^{N_t(k,j)} \cdot \mu_{0kj\ell}R_{hh}\left((p,\mathbf{x}_r,t_q),(k,\mathbf{x}_j,t_\ell)\right) = R_{hf}\left((p,\mathbf{x}_r,t_q),\mathbf{x}_0\right),
\end{aligned} \tag{4}$$

in which $m = 1, \ldots N_f$, $p = 1, \ldots N_p$, $r = 1, \ldots N_h(p)$ and ℓ and $q = 1, \ldots N_t(p, r)$. Our prior knowledge of the spatial structure (the unconditional covariance function) of f is given by R_{ff}. R_{hh} and R_{hf} are the unconditional covariance of h and the unconditional cross covariance of f and h, respectively, which are determined by a first-order analysis with the given R_{ff}. That is,

$$\begin{aligned}
&R_{hf}((k,\mathbf{x}_i,t_\ell),\mathbf{x}_m) = \sum_{j=1}^{N_e}J\left((k,\mathbf{x}_i,t_\ell),\mathbf{x}_j\right)R_{ff}\left(\mathbf{x}_j,\mathbf{x}_m\right)\\
&k = 1,\ldots N_p; i = 1,\ldots N_h(k);\ \ell = 1,\ldots N_t(k,i);\ m = 1,\ldots N_e
\end{aligned} \tag{5}$$

$$\begin{aligned}
&R_{hh}\left((u,\mathbf{x}_i,t_q),(k,\mathbf{x}_j,t_\ell)\right) = \sum_{m=1}^{N_e}R_{hf}\left((u,\mathbf{x}_i,t_q),\mathbf{x}_m\right)J\left((k,\mathbf{x}_j,t_\ell),\mathbf{x}_m\right)^T\\
&u \text{ and } k = 1,\ldots N_p;\ i \text{ and } j = 1,\ldots N_h(u \text{ or } k);\\
&q \text{ and } \ell = 1,\ldots N_t(u \text{ or } k, i),
\end{aligned} \tag{6}$$

where $J((k, \mathbf{x}_j, t_\ell), \mathbf{x}_m)$ is the sensitivity of head at jth location at time t_ℓ for kth pumping test with respect to the change of parameter at mth location; N_e is the number of elements in the study domain. The sensitivity matrix is evaluated using an adjoint state approach (see *Zhu and Yeh* [2005] for details).

[7] After obtaining the new estimate for all the elements using cokriging, the conditional covariance of f, $\varepsilon_{ff}^{(1)}(\mathbf{x}_m, \mathbf{x}_q)$, is then determined by

$$\varepsilon_{ff}^{(1)}(\mathbf{x}_m, \mathbf{x}_q) = R_{ff}(\mathbf{x}_m, \mathbf{x}_q) - \sum_{k=1}^{N_f} \lambda_{mk} R_{ff}(\mathbf{x}_k, \mathbf{x}_q) - \sum_{k=1}^{N_p} \sum_{j=1}^{N_h(k)} \sum_{\ell=1}^{N_t(k,j)} \mu_{mkj\ell} R_{hf}\big((k, \mathbf{x}_j, t_\ell), \mathbf{x}_q\big), \quad (7)$$

where m and $q = 1, \ldots N_e$. The conditional covariance reflects the effect of data on the reduction of uncertainty in the estimated parameter field. Subsequently, the estimated log perturbations of the property fields are added to the log of the effective properties, $F(\mathbf{x})$, then converted to the arithmetic scale, and used to solve equation (1) with boundary and initial conditions for the conditional effective head fields, $h^{(1)}(k, \mathbf{x}_j, t_\ell)$, of each pumping test.

[8] Following cokriging, a linear estimator of the following form,

$$\hat{f}^{(r+1)}(\mathbf{x}_0) = \hat{f}^{(r)}(\mathbf{x}_0) + \sum_{k=1}^{N_p} \sum_{j=1}^{N_h(k)} \sum_{\ell=1}^{N_t(k,j)} \cdot \omega_{0kj\ell}^{(r)} \Big[h^*(k, \mathbf{x}_j, t_\ell) - h^{(r)}(k, \mathbf{x}_j, t_\ell)\Big], \quad (8)$$

is used to improve the estimate for iteration $r > 1$, where $\omega_{0kj\ell}^{(r)}$ is the weight term, representing the contribution of the difference between the observed and simulated conditional heads (i.e., $h^*(k, \mathbf{x}_j, t_\ell)$ and $h^{(r)}(k, \mathbf{x}_j, t_\ell)$, respectively) at iteration r at location $\mathbf{x}_j$ of the kth pumping test at time t_ℓ to the estimate at location $\mathbf{x}_0$. The weights are determined by solving the following system of equations:

$$\sum_{k=1}^{N_p} \sum_{j=1}^{N_h(k)} \sum_{\ell=1}^{N_t(k,j)} \omega_{0kj\ell}^{(r)} \Big[\varepsilon_{hh}^{(r)}\big((p, \mathbf{x}_m, t_q), (k, \mathbf{x}_j, t_\ell)\big) + \Theta^{(r)} \delta_{kj\ell}\Big] = \varepsilon_{hf}^{(r)}\big((p, \mathbf{x}_m, t_q), \mathbf{x}_0\big), \quad (9)$$

where $p = 1, \ldots N_p$, $m = 1, \ldots N_h(p)$ and $q = 1, \ldots N_t(p, m)$. The terms $\varepsilon_{hh}^{(r)}$ and $\varepsilon_{fh}^{(r)}$ are the conditional covariance and the conditional cross covariance at iteration (r), which are evaluated using equations (5) and (6) using the conditional covariance of f (i.e., $\varepsilon_{ff}^{(r)}$ which is obtained from equation (7) for the first iteration). A dynamic stabilizer, $\Theta^{(r)}$, is added to the diagonal elements of $\varepsilon_{hh}^{(r)}$ (δ_{kjl} is a Dirac delta, equal to 1 when $k = j = l$ and zero otherwise) to stabilize the solution to equation (9). The dynamic stabilizer at iteration, r, is the maximum value of the diagonal elements of $\varepsilon_{hh}^{(r)}$ at that iteration times a user-specified multiplier [see *Yeh et al.*, 1996]. After completion of the estimation using equation (8) for all elements in the domain, the conditional covariance of f is updated subsequently as given below:

$$\varepsilon_{ff}^{(r+1)}(\mathbf{x}_m, \mathbf{x}_n) = \varepsilon_{ff}^{(r)}(\mathbf{x}_m, \mathbf{x}_n) - \sum_{k=1}^{N_p} \sum_{j=1}^{N_h(k)} \sum_{\ell=1}^{N_t(k,j)} \omega_{mkj\ell} \varepsilon_{hf}^{(r)}\big((k, \mathbf{x}_j, t_\ell), \mathbf{x}_n\big), \quad (10)$$

where n and $m = 1, \ldots N_e$.

[9] The iteration steps of SimSLE are the same as those in the SLE algorithm used by *Yeh et al.* [1996]. For noise-free hydrographs, the convergence is achieved if (1) change in variances that represent spatial variability of the estimated hydraulic properties between current and last iterations is smaller than a specified tolerance (i.e., the spatial variance of the estimates stabilizes), implying that the SimSLE cannot improve the estimation any further and (2) change of simulated heads between successive iterations is smaller than the tolerance, indicating that the estimates will not significantly improve the head field. If one of the two criteria is met, the estimates are considered to be optimal and the iterations are terminated.

[10] Head observations from laboratory or field experiments often contain noise (i.e., signals caused by processes not modeled by the governing flow equation, such as Earth tide and others including measurement errors) in addition to effects of hierarchical heterogeneity. Such unresolved noises can lead to divergence of inverse solutions (i.e., unrealistic estimates). As a consequence, an important issue that ought to be addressed is to what degree should the observed head be used to improve estimates of the hydraulic properties. Stabilization of mean square error of the simulated head during iteration should address the issue. Consider the mean square error of the head:

$$E\left[\left(h^*(k, \mathbf{x}, t) - \hat{h}^{(r)}(k, \mathbf{x}, t)\right)^2\right] = E\left[\left((h(k, \mathbf{x}, t) + \tau) - \hat{h}^{(r)}(k, \mathbf{x}, t)\right)^2\right] = E\left[\left(h(k, \mathbf{x}, t) - \hat{h}^{(r)}(k, \mathbf{x}, t)\right)^2\right] + E[\tau^2] = \varepsilon_{hh}^{(r)}(k, \mathbf{x}, t) + \sigma_\tau^2, \quad (11)$$

in which the observed head at location $\mathbf{x}$ and time t during kth pumping event is denoted by $h^*(k, \mathbf{x}, t)$; the corresponding simulated head based on the estimated parameters at rth iteration is given by $\hat{h}^{(r)}(k, \mathbf{x}, t)$. Equation (11) assumes that the observed head consists of the noise free head, $h(k, \mathbf{x}, t)$, and random noise, τ, with variance σ_τ^2. The term $\varepsilon_{hh}^{(r)}(k, \mathbf{x}, t)$ denotes the theoretical residual head variance for the noise free case at rth iteration (i.e., the diagonal term in equation (6)), which should decrease and approach zero with iterations because of improvement of the parameter estimates. Therefore, the mean square error for cases with noise should asymptotically converge to σ_τ^2 as the number of iterations increases. More importantly, equation (11) suggests that once $\varepsilon_{hh}^{(r)}(k, \mathbf{x}, t)$ becomes smaller than σ_τ^2 (i.e., the mean square error stabilizes), the iteration should stop to avoid overusage of imperfect head data (i.e., updating the estimate with noise). Consequently, for cases where the observed head is not noise free, we use the stabilization of the L2 norm of the conditional heads to terminate the iteration, i.e.,

$$L2_{cond}(r) = \frac{1}{N} \sum_{i=1}^{N} \left(h_i^* - \hat{h}_i^{(r)}\right)^2, \quad (12)$$

where h_i^* and $\hat{h}_i^{(r)}$ are observed and simulated heads, respectively; i is the index denoting the observation in a given time and location from a pumping test; N is the total number of head observations from all the pumping tests.

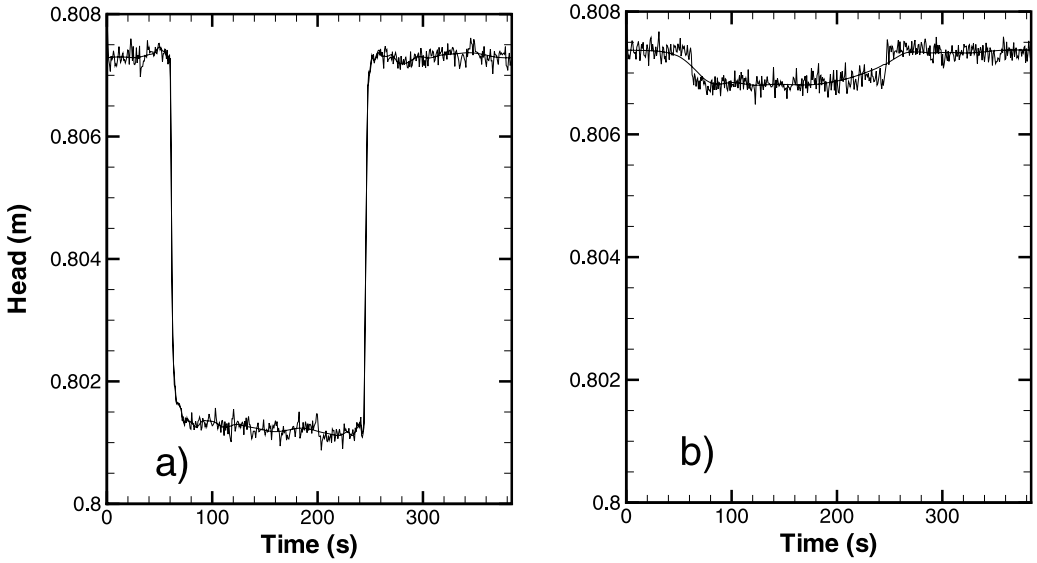

Figure 1. Typical hydrographs and corresponding denoised hydrographs (solid black line) in the sandbox experiments [*Liu et al.*, 2007]: (a) SNR 17.75 and (b) SNR 4.41. Notice vertical exaggeration.

Hereafter, we will refer to equation (12) as to the conditional L2 norm.

2.3. Preprocessing Data for SimSLE

2.3.1. Diagnosis of Bias

[11] The first step in preprocessing HT data for the SimSLE analysis is to qualitatively check for any bias or inconsistency in the data. We suggest this be accomplished by a simple rule-of-thumb approach: plotting hydrographs (drawdown time data) for each pumping test. For example, arrival time of a given drawdown should increase with distance from the pumping well unless there are physically explainable anomalies (e.g., fracture zones or other fast flow paths). Plots of evolution of contour surface of the drawdown induced by each pumping test generally should follow behaviors of drawdown in homogeneous aquifers [e.g., *Bakr et al.*, 1978] although details are different. Data with significant anomalies should be examined for possible operational causes (e.g., mislabeling monitoring ports or wells, leakage between packers, malfunctioning of equipment, or other factors). If the operational causes can be identified, the data sets should be corrected or excluded. Repeating the HT experiment also facilitates a viable diagnosis. Drawdown-log time plots should also reveal possible boundary effects and they are useful in assigning the type of boundary conditions of the modeling domain.

2.3.2. Wavelet Denoising

[12] Next, we tackle data noise or signals or perturbations caused by factors other than aquifer heterogeneity. To eliminate the noise, a denoising method based on wavelet analysis [e.g., *Mallat*, 1999; *Zhang et al.*, 2006] was developed. A wavelet analysis is similar to Fourier analysis in the sense that it breaks a signal down into its constituent parts for analysis. In contrast to the Fourier transform, the wavelet transform allows exceptional localization in both the time domain via translations of the mother wavelet, and the scale (frequency) domain via dilations, when analyzing signals of a nonstationary nature. Software from http://www.mathworks.com was used. Different families of wavelets were tested in this study and we found that the Daubechies 4 wavelet is effective for our cases.

[13] The wavelet denoising procedure used in this study comprises the following steps: (1) applying the wavelet transform to the noisy hydrograph to produce wavelet coefficients, (2) selecting an appropriate threshold limit and a threshold method to remove the noise, (3) inversing the wavelet transform of thresholded wavelet coefficients to generate a denoised hydrograph, and (4) calculating the variance of the difference between the denoised hydrographs and those of the equivalent homogenous medium. This variance should be approximately equal to the theoretical head variance from the first-order analysis. If this criterion is not met, repeat steps 2 and 3. Generally speaking, distinguishing noise from the effects of heterogeneity in a hydrograph can be highly subjective unless characteristics of noise or heterogeneity are known a priori.

2.4. Signal-to-Noise Ratio

[14] Following wavelet denoising, we examined the signal-to-noise ratio (SNR) of each well hydrograph. The SNR is an electrical engineering concept, which is defined as the ratio of signal amplitude to noise amplitude. In our study, it is defined as the ratio between the maximum drawdown of the denoised hydrograph and the standard deviation of noise. That is,

$$\mathrm{SNR} = \frac{\text{signal}}{\text{noise}} = \frac{\left(|h_d(\mathbf{x},0) - h_d(\mathbf{x},t)|\right)_{\max}}{\Sigma_n}, \qquad (13)$$

where $|\ |$ denotes the absolute value; h_d denotes the head after denoising; and Σ_n represents the standard deviation of noise, which is estimated from the wavelet denoising procedure. In signal processing, signals with SNR of 100 or greater are considered to be good signals. For hydrologic processes, we found that hydrographs with an average SNR of 7.13 or greater are effective for our synthetic case. Figure 1a shows an original hydrograph from *Illman et al.* [2007, 2008] and *Liu et al.* [2007] and the corresponding denoised hydrograph; the SNR is 17.75. The denoised hydrograph

therefore is considered useful and it improves the estimates. On the other hand, a hydrograph with low-SNR data (4.41) (Figure 1b), especially near the falling and rising limbs of the graph (reflecting effects of pumping and recovery, respectively), was discarded since it did not improve the estimates. Low-SNR data at early time and recovery can lead to erroneous estimates of the S_s field. As a rule of thumb, the SNR should be much greater than 1 such that the trend of drawdown induced by pumping is evident after denoising. However, a more in-depth analysis is needed to establish a rigorous criterion.

2.5. Estimation of Effective Properties and Variances

[15] Afterward, desirable hydrographs from the HT were selected to estimate unconditional effective K and S_s of an equivalent homogeneous medium with the known pumping rates. This task was accomplished by minimizing the squared difference between the observed head values and those obtained from simulations based on the homogeneity assumption using the VSAFT2:

$$\sum_{i=1}^{n_p}\sum_{j=1}^{n_h}\sum_{k=1}^{n_t}\left(h^*_{ijk}-\hat{H}_{ijk}\right)^2=\min_{K,S_s}. \tag{14}$$

In the above equation, h^*_{ijk} denotes observed head at kth time, jth observation location, and ith pumping test, and the corresponding simulated head using the unconditional effective K and S_s is indicated by $\hat{H}_{ijk}$. Minimization of equation (14) was performed using a standard nonlinear least squares (i.e., Gauss-Newton) approach in conjunction with the Levenberg-Marquardt algorithm [*Press et al.*, 1992]. Sensitivity matrices were evaluated by solving the sensitivity equation (i.e., differentiation of equation (1) with respect to the parameter).

[16] Subsequently, square of the difference between the head observed and the head in the equivalent homogeneous medium is used to compute the sample variance of the observed head at the given location and time. This sample head variance is in turn used to estimate the variance of the hydraulic properties by minimizing the following objective function:

$$\sum_{i=1}^{n_p}\sum_{j=1}^{n_h}\sum_{k=1}^{n_t}\left(\sigma^2_{ijk}-\hat{\sigma}^2_{ijk}\right)^2=\text{minimum}, \tag{15}$$

where σ^2_{ijk} is the sample head variance and $\hat{\sigma}^2_{ijk}$ is the theoretical head variance at kth time, jth observation location, and ith pumping test. This theoretical head variance is evaluated using the first-order approximation (i.e., equation (6)) with given variance of the hydraulic properties. The Levenberg-Marquardt algorithm was used to seek the variances of the properties that minimize equation (15). Relative change of equation (15) between successive iterations is used as convergence criterion. These estimated effective hydraulic properties and variances are used as prior information required by SimSLE for estimating the spatially varying K and S_s fields.

3. Applications to a Synthetic Aquifer

3.1. Description of the Aquifer and HT Test

[17] To test our SimSLE and the data preprocessing procedure for the HT analysis, a two-dimensional, cross-sectional, synthetic aquifer of the same length and height as the sandbox experiment conducted by *Illman et al.* [2007, 2008] and *Liu et al.* [2007] was used. The dimensions of the sandbox were 193.0 cm in length, 82.6 cm in height, and 10.2 cm in depth. Twenty four locations were selected (solid circles in Figure 2a) to serve as pressure monitoring ports during four pumping tests at four locations (open circles in Figure 2a). Both sides of the aquifer were set to the same constant boundary condition with a total head of 200 cm, while the bottom and top boundaries were set to be no-flux boundaries. The initial total head was assigned to be 200 cm. This aquifer was discretized into 741 elements and 800 nodes with element dimensions of 4.10 cm $\times$ 4.13 cm.

[18] The synthetic aquifer was created to imitate a geologic formation of hierarchical heterogeneity, consisting of four units with a bedding dip angle of 20°. The ln K and ln S_s fields within each unit were assumed to be normally distributed random fields with different means and variances (see Table 1). An exponential model was used as the spatial covariance functions of the fields with correlation scales of 200 cm in the bedding direction and 12 cm in the direction perpendicular to bedding. Different random seed numbers were used to create the ln K and ln S_s fields (Figures 2a and 2e, respectively) so that they are independent of each other.

[19] Four pumping tests in this aquifer were simulated using a time step of 0.25 s for a period of 15 s at which the drawdown reached a steady state condition because of the small size of the aquifer. Each pumping test had a pumping rate of 0.3 cm^3/s and drawdown at all 24 wells were recorded. These simulated drawdown-time data sets were regarded as noise-free hydrographs. These noise-free hydrographs were subsequently corrupted by adding normally distributed white noise with a standard deviation of 0.07 cm to represent measurement errors and these corrupted hydrographs are denoted as noisy hydrographs. Finally, the wavelet denoising procedure was applied to these noisy hydrographs to obtain the so-called denoised hydrographs. The SNR values of the 96 corrupted hydrographs were found much greater than 1 and all the hydrographs were included in the following HT analysis.

[20] To investigate the ability of SimSLE to estimate the heterogeneous K and S_s fields, drawdowns at 5 times (four early times at 0.5 s, 1.75 s, 2.25 s and 3 s; and one later time at 15 s) from the noise-free, noisy, and denoised HT hydrographs were used. Such a choice of sampling times stems from the finding by *Wu et al.* [2005] that the drawdown at early time is highly correlated with the S_s field and only weakly and negatively correlated with the K field in the area between the pumping and the observation locations. At large time, the drawdown is correlated at various degrees with K values in the area within the cone of depression but not the S_s field. In the estimation, K and S_s values at the top observation port in the left column in Figure 2a were assumed to be known and they were used as the hard data for conditioning the estimation in all the cases.

3.2. Performance Assessment

[21] Performance of our estimator for the synthetic aquifer case was evaluated using the standard correlation coefficient ($1 \geq |r| \geq 0$) which measures the similarity between the true and the estimated fields. A correlation coefficient close to 1 means the two fields are similar in pattern, even though the

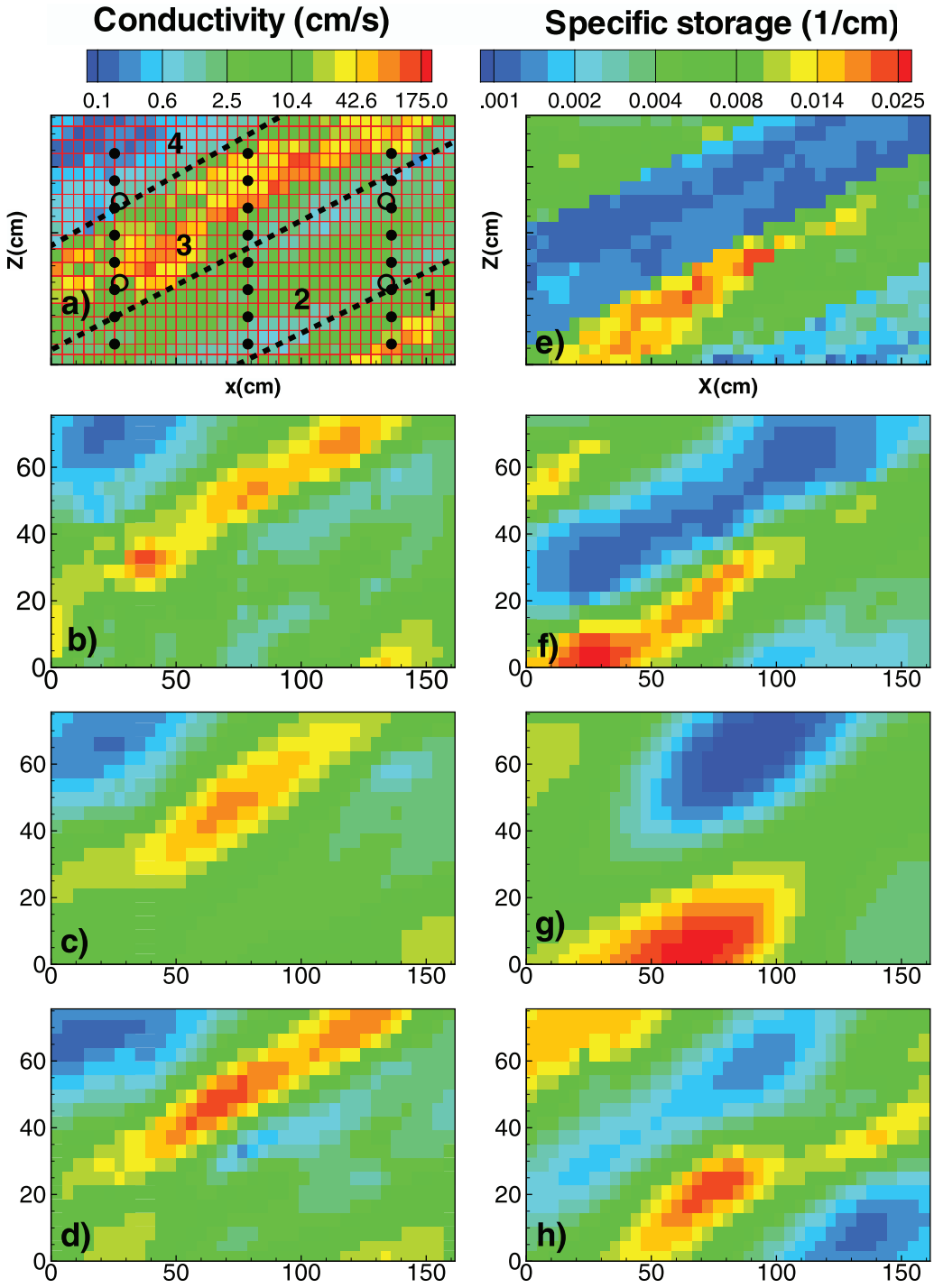

Figure 2. (a and e) True K (cm/s) and S_s (1/cm) fields of the synthetic aquifer. The dashed lines are the boundaries between units (1, 2, 3, and 4). Estimated K and S_s fields using (b and f) noise-free data, (c and g) noisy data, and (d and h) denoised data.

mean value of the two fields may be different. Thereby, in addition to the correlation, mean absolute error (L1 norm) and mean square error (L2 norm) of the estimated field were used to judge the performance.

[22] Besides the correlation analysis, L1 and L2 norms, similarity between the true and estimated hydraulic property fields was also determined using a fuzzy similarity comparison method, which has been applied to the task of comparing spatial patterns [*Hagen*, 2003]. This method allows a user to specify weights for both location and value matching. Specifically, the method computes similarity between the true property at a given element and the estimated property at the corresponding element and neighboring elements. The similarity is based on a locational and an error membership value, which are defined below.

Table 1. Means and Variances of the Generated ln K and ln Ss Field in Each Unit of the Synthetic Aquifer and Over the Entire Aquifer

	Unit 1	Unit 2	Unit 3	Unit 4	Overall
ln K					
Mean	2.03	0.85	2.46	−0.67	1.39
Variance	1.01	0.17	2.00	0.46	2.18
ln Ss					
Mean	−6.06	−4.82	−6.43	−5.06	−5.61
Variance	0.05	0.24	0.01	0.08	0.66

[23] An exponential function based on the statistical spatial model for the hydraulic properties was used to define a locational membership

$$V_l = \exp\left(-\sqrt{\left(\frac{l_x}{\lambda_x}\right)^2+\left(\frac{l_y}{\lambda_y}\right)^2}\right). \tag{16}$$

In which l_x and l_y are the separation distance between two compared elements in the x and y directions; λ_x and λ_y are the correlation scale of the true property field in the x and y directions, respectively. Equation (16) implies that the similarity decreases with an increase of the ratio of the separation distance to the correlation scale.

[24] The standard deviation of the true property (STD) is used as an error limit and to define the value of the error membership:

$$V_v = \begin{cases} 1.0 - \dfrac{|v|}{STD} & \dfrac{|v|}{STD} < 1 \\ 0 & \dfrac{|v|}{STD} \geq 1, \end{cases} \tag{17}$$

where v is the difference between the true property at a given element and the estimated property at a selected element. That is, if the difference of two properties is greater than the standard deviation, the two properties are judged to be completely different. Otherwise, a linear decay is used to determine the error membership value.

[25] The true property of an element in the domain is then compared with the estimated property of the same element and with the estimated property of neighboring elements (within the correlation scales). Equations (16) and (17) are employed to determine the locational and error membership values. For each pair of the true and the estimated property, a similarity value is then computed by multiplying the locational membership with the error membership. Only the maximum value among the similarity values of all pairs is retained for the selected element. Next, the maximum similarity between the estimated property at that element with the true property at the corresponding element and at neighboring elements is determined. The similarity value between the true and the estimated properties of this element is then defined by the average of the two maximum similarities. This procedure is applied to all the elements in the synthetic domain. Finally, a domain similarity is defined as the average element similarity of all elements in the domain.

3.3. Results

[26] Table 2 tabulates the estimated effective K and S_s values for an equivalent homogeneous and isotropic porous medium of the synthetic aquifer, using the noise-free, noisy, and denoised hydrographs. Estimated variances of ln K and ln S_s over the entire aquifer are also listed in Table 2, which were obtained using equation (15) with visually estimated correlation scales (100 cm and 33 cm, in the direction parallel and perpendicular to bedding, respectively) and the known dip angle. Table 2 shows that the estimated effective properties based on the noise-free, noisy, and denoised hydrographs are similar. Estimated effective K values are found to be slightly greater than the geometric mean while the effective S_s values are in agreement with the arithmetic mean of the corresponding values of the four units. Apparently, noise does not have significant effects on the estimates of the effective properties since this is an overdetermined inverse problem (i.e., more data than parameters need to be estimated and the problem is well posed [*Yeh et al.*, 2007]. The least squares approach is expected to be sufficient and effective.

[27] As shown in Table 2, estimated variances of ln K and ln S_s with equation (15) are smaller than the true ones, likely due to insufficient head data. The nonstationary nature of the drawdown (i.e., its variance depends on the spatially and temporally varying mean gradient) demands a large number of head data at the same radial distance from the pumping wells to obtain a representative sample head variance in equation (15).

[28] To illustrate the effectiveness of wavelet denoising, 480 pairs of observed heads before and after denoising at the 24 ports at 5 sampling times during the 4 pumping tests and corresponding simulated heads based on effective K and S_s are plotted as red circles in Figures 3a and 3b, respectively. Pluses in Figure 3 correspond to the simulated heads plus and minus one standard deviation of the theoretical head perturbation caused by heterogeneity, $\varepsilon^{(0)}_{hh}(k, \mathbf{x}, t)$. A comparison of Figures 3a and 3b shows that the wavelet denoising removed most of perturbations of the total heads between 199.5 cm and 200 cm, which are not significantly affected by the pumping test. Head data after denoising are generally within one standard deviation of the theoretical head perturbation caused by heterogeneity, suggesting

Table 2. Estimated Effective Hydraulic Properties and Estimated Spatial Variances of the Properties in the Synthetic Aquifer Using Noise-Free, Noisy, and Denoised Hydrographs[a]

Cases	Effective ln K (cm/s)	Variance ln K	Effective ln S_s (1/cm)	Variance ln S_s	Iterations
Synth (noise free)	1.422	0.679	−5.114	0.442	16
Synth noise	1.414	0.520	−5.181	0.154	17
Synth denoised	1.431	0.571	−5.084	0.329	23

[a]The number of iterations required to reach to the convergence of the solution is listed in the last column.

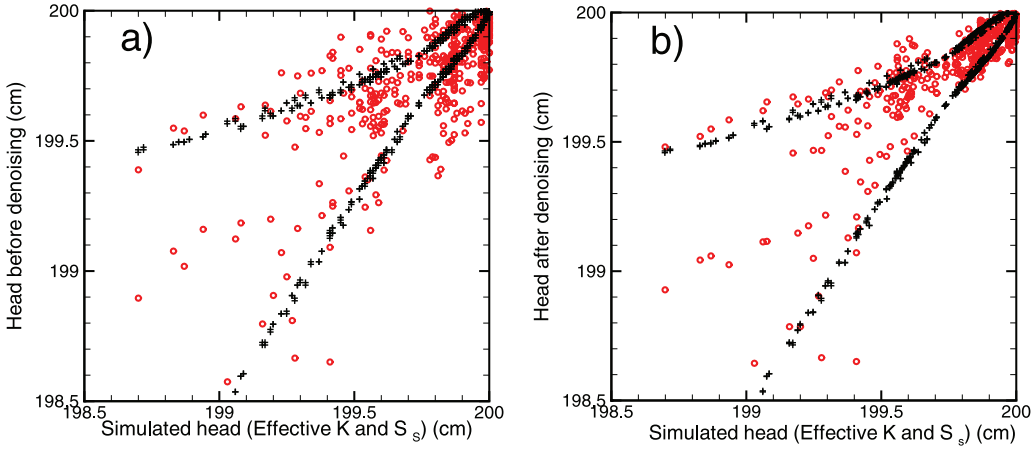

Figure 3. Effectiveness of wavelet denoising. A plot of 480 pairs of heads (a) before and (b) after denoising and simulated heads (red circles) based on effective K and S_s of the synthetic aquifer; pluses denote the simulated heads ±1 standard deviation of head variation induced by heterogeneity.

the remaining perturbations are likely effects of aquifer heterogeneity.

[29] Table 3 lists values of unconditional L1 and L2 norms of the hydraulic head, which are defined as

$$L1_{un} = \frac{1}{N}\sum_{i=1}^{N}\left|h_i - \hat{h}_i\right|$$
$$L2_{un} = \frac{1}{N}\sum_{i=1}^{N}\left|h_i - \hat{h}_i\right|^2, \qquad (18)$$

where $\hat{h}$ is the simulated head in the equivalent homogeneous aquifer using the effective parameter values derived from the noise-free hydrographs; h is the noise-free, noisy, or denoised heads of the heterogeneous aquifer, at the 5 selected times during the four pumping tests (i.e., N = 480). These statistics aim to measure effects of the hierarchical heterogeneity (i.e., spatial variability or structured noise), measurement noise, and noise residuals in hydrographs after wavelet denoising. As expected, the noise-free hydrograph has the smallest unconditional L1 and L2 norms, reflecting the effect of heterogeneity only. L1 and L2 norms are highest for the noisy hydrographs where head perturbations are results of both heterogeneity and the white noise. Values of L1 and L2 norms in the last row of Table 3 for the denoised hydrographs are between those of noise-free and noisy hydrographs, indicative of only partial removal of noise by the wavelet denoising procedure. On the basis of the L2 values, 92% of the perturbations in the noisy hydrographs are effects of hierarchical heterogeneity and the remaining 8% are random noise. The wavelet denoising procedure removed 80% of the random noise.

Table 3. Unconditional L1 and L2 Norms of the Head Based on Equation (18) for the Noise-Free, Noisy, and Denoised Cases of the Synthetic Aquifer Experiments

	L1	L2
Noise free	0.102	0.0374
Noisy	0.124	0.0405
Denoised	0.109	0.0380

[30] Estimating 741 pairs of K_s and S_s of the synthetic aquifer on the basis of the 480 drawdowns and one pair of K_s and S_s hard data set is an underdetermined (i.e., over-parameterized or ill-posed [see *Yeh et al.*, 2007]) inverse problem. That is, the number of the parameters to be estimated is larger than the number of data sets available. While SLE aims to seek the estimate of the conditional effective parameters [*Yeh et al.*, 1996] for underdetermined problems, the noise or unresolved noise may lead to anomalously high or low estimates. To circumvent this problem, the criterion based on the stabilization of the conditional L2 norm of the head (i.e., equation (12)) was employed.

[31] The conditional L2 norms for noise-free, noisy, and denoised cases are shown in Figure 4 as a function of the number of iterations, while corresponding behaviors of the ln K and ln S_s estimates in terms of their spatial variances are illustrated in Figures 5a and 5b, respectively. As expected, in the noise-free case, the L2 decreased from 0.0374 cm^2 (the unconditional L2, representing effects of heterogeneity only, see Table 3) exponentially as expected for the theoretical head variance (equation (11) with σ_τ^2 = zero).

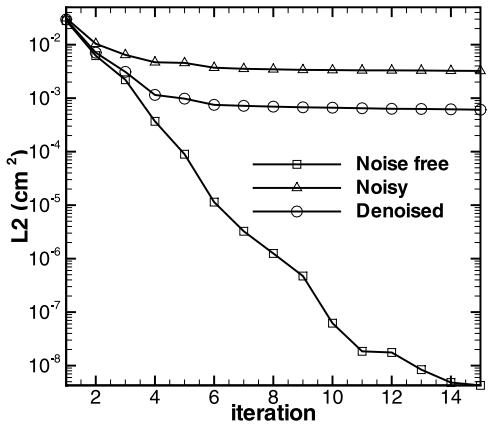

Figure 4. Conditional L2 norms of the head as a function of iteration for noise-free, noisy, and denoised data sets for the synthetic aquifer case.

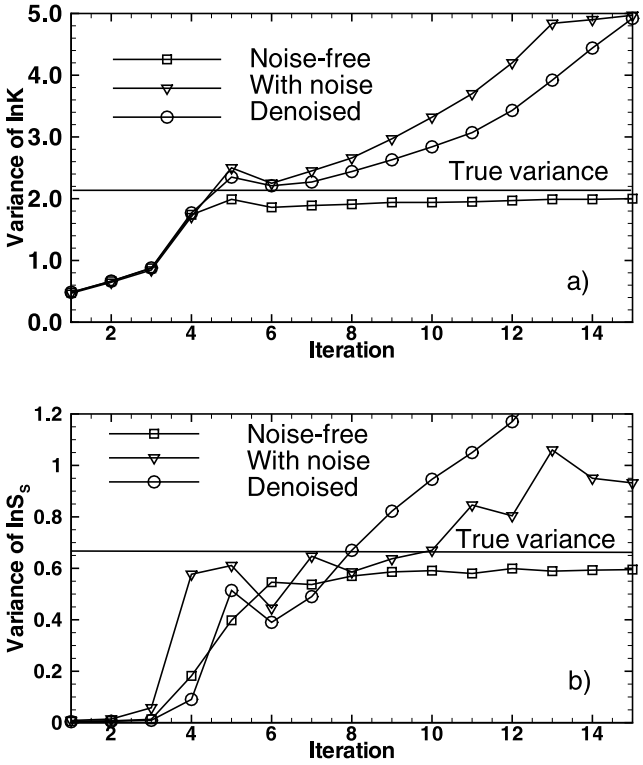

Figure 5. Variances of the estimated (a) ln K and (b) ln S_s fields versus the number of iterations for different scenarios associated with the synthetic aquifer.

Such a decrease suggests continuous improvements of the estimates and thus the predicted heads during iteration. Improvement of the estimates can also be seen in Figures 5a and 5b. According to Figures 5a and 5b, the variances, based on noise-free hydrographs, started from zero since the first guesses were the effective K and S_s, and then increased as point measurements of K and S_s and head information were included via cokriging. These variances continued to increase after the first iteration because of the successive linear estimation, which successively approximates the nonlinear relation between the head information and the hydraulic properties. Subsequently, these variances approached some stable values, indicating that usefulness of the observed head information was exhausted as the L2 decreased to a very small value. Notice that the spatial variances of the final estimates are smaller than the true variances, suggesting that the estimated fields are smoother than the true fields. This is expected since SimSLE seeks the conditional effective properties. The estimates at the tenth iteration were chosen as our final estimates on the basis of change in variances of the estimates.

[32] Figure 4 shows that for the case where hydrographs were noisy, the L2 norm decreased from 0.0405 cm^2 (representing effects of heterogeneity plus noise) and stabilized at the fourth iteration to the value of 0.0033 cm^2, which is close to the noise level we imposed (i.e., 0.0049 cm^2). The difference between the true noise level and the L2 norm is expected since L2 norm represents only a sample variance of the noise. Likewise, the variance of the estimated ln K (Figure 5a) fluctuated at the fourth iteration, then increased and exceeded the true variance for the noise hydrographs.

Table 4. Performance Assessment Statistics of Results From the Synthetic Aquifer

	L1	L2	Correlation	Similarity	Iterations
			Noise Free		
K	0.469	0.405	0.909	0.889	10
S_s	0.279	0.140	0.900	0.859	10
			Noised		
K	0.643	0.686	0.856	0.833	4
S_s	0.635	0.666	0.589	0.757	4
			Denoised		
K	0.635	0.678	0.857	0.851	6
S_s	0.536	0.421	0.722	0.762	6

The variance of the estimated ln S_s generally increased continuously and rapidly, indicative of divergence of the solution (Figure 5b) due to inclusion of noise in the estimation. Therefore, the final estimate was obtained at the fourth iteration where the L2 norm stabilized and before the estimates were "overimproved."

[33] As illustrated in Figure 4, the conditional L2 norm for the denoised case diminished from the unconditional L2 norm, 0.038 cm^2 (comprising the effects from both heterogeneity and noise residuals after wavelet denoising). Then, it stabilized around the fifth iteration at a value of 0.0006 cm^2, which is smaller than that for the noisy hydrographs since noise was partially removed by the wavelet denoising procedure. The variances of the estimated fields however still increased with iterations similar to those based on noisy hydrographs. According to the L2 stopping criterion, the final estimates were those at the sixth iteration. The variance of the estimated ln K for this case is slightly greater than its true variance and that of the estimated ln S_s is smaller than its true variance.

[34] A visual comparison between the true heterogeneous K and S_s fields (Figures 2a and 2e, respectively) and the final estimated fields based on the noise-free hydrographs at the tenth iteration (Figure 2b for K and Figure 2f for S_s) suggests that SimSLE depicts hierarchical spatial variation of hydraulic parameters (i.e., variation between units as well as that within a unit). The estimated K and S_s fields at the fourth iteration using the noisy head values are shown in Figures 2c and 2g, respectively, while the estimated K and

Table 5. Means and Variances of True and Estimated Hydraulic Properties of Each Zone Using Noise-Free, Noisy, and Denoised Hydrographs

	Log Hydraulic Conductivity				Log-Specific Storage			
		Noise				Noise		
Unit	True	Free	Noisy	Denoised	True	Free	Noisy	Denoised
				Mean				
1	2.03	1.86	1.95	2.13	−6.06	−5.78	−6.34	−6.59
2	0.85	1.04	1.43	1.28	−4.82	−4.78	−5.18	−5.08
3	2.46	2.60	2.64	2.60	−6.43	−6.20	−5.85	−6.07
4	−0.67	−0.40	−0.40	−0.64	−5.06	−5.05	−5.16	−4.53
				Variance				
1	1.01	0.82	0.31	0.67	0.05	0.06	1.60	0.44
2	0.17	0.59	0.50	1.04	0.24	0.38	0.58	0.94
3	2.00	1.47	1.22	1.55	0.01	0.09	0.27	0.46
4	0.46	0.73	0.71	0.57	0.08	0.15	1.46	0.39

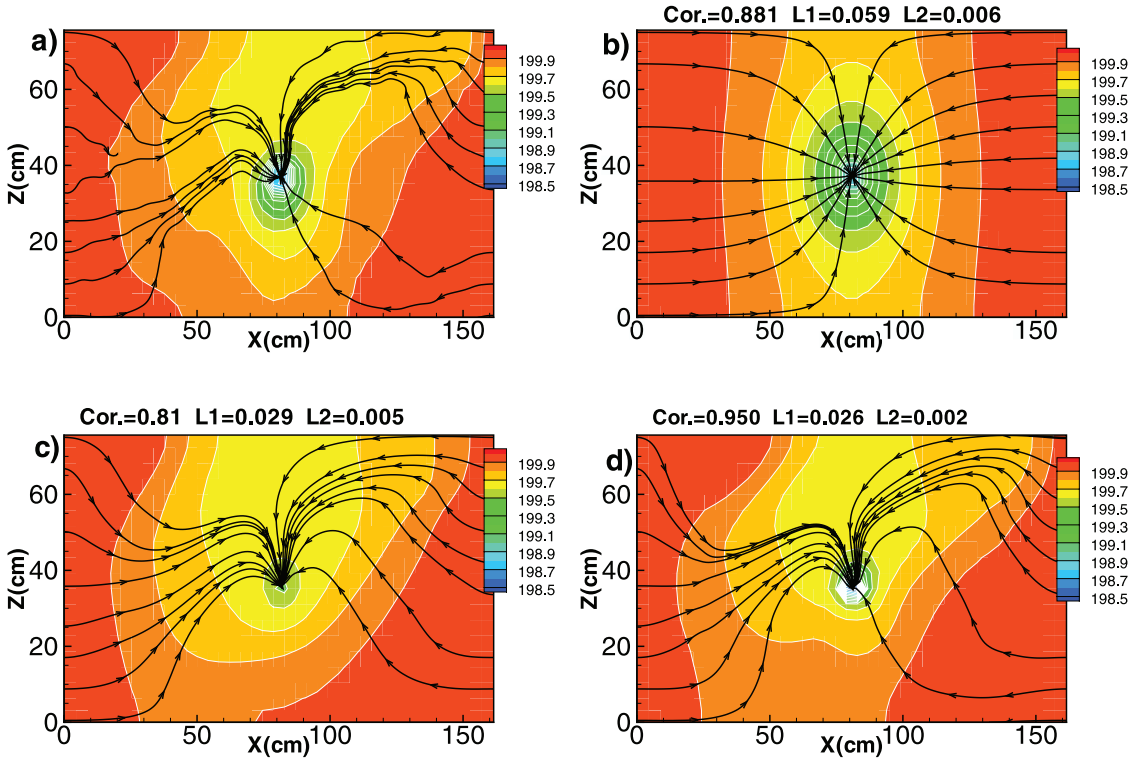

Figure 6. Comparison of head and streamlines at t = 1.5 s after pumping in the synthetic aquifer with (a) true and (b) effective parameters and estimated K and S_s fields based on (c) noisy and (d) denoised hydrographs. The statistical metrics are based on the true head field in Figure 6a.

S_s fields based on denoised hydrographs at the sixth iteration are shown in Figures 2d and 2f, respectively.

[35] Performance metrics for these cases are reported in Table 4. According to these metrics, the estimated S_s field can be as good as the estimated K field if the noise-free hydrographs were used. However, it is much worse than the estimated K field if the noisy hydrographs were used. This result confirms that estimation of the S_s field is more prone to effects of noise in hydrographs [e.g., *Li et al.*, 2007]. In comparison to results of the noise-free case, the estimated K and S_s fields using the noisy data and the conditional L2 norm as the stopping criterion are smoother but still retain the general pattern of heterogeneity. The performance metrics also indicate that denoised data can improve the estimates, revealing more details of heterogeneity. But the estimated fields are still inferior to those obtained from noise-free data, suggesting that the wavelet denoising procedure apparently is useful but cannot restore the true hydrograph. Difficulties in estimating S_s field can be attributed to the fact that early time drawdown needed for estimating the S_s field is often small in magnitude. Its SNR is very small once noise is imposed.

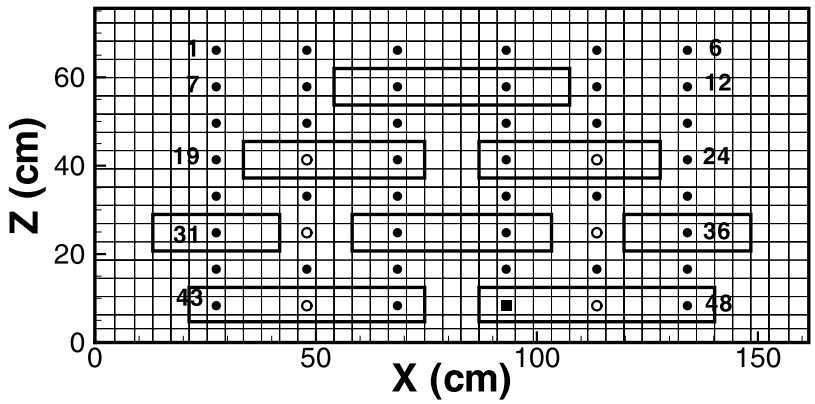

Figure 7. Schematic setup and discretization of the sandbox used in the experiment of *Liu et al.* [2007] and *Illman et al.* [2007]. Open circles are pumping ports; both solid and open circles are used as observation ports. The solid square denotes the pumping port for the validation purpose. The open rectangles are the low-permeability zones.

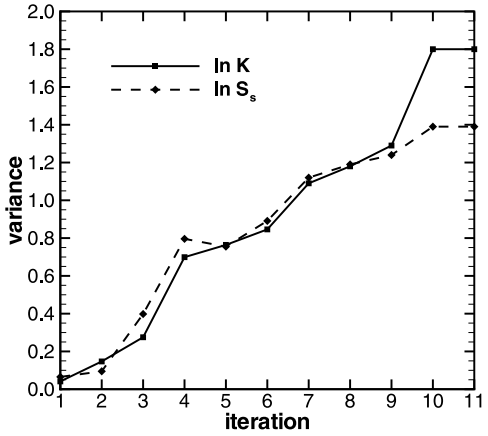

Figure 8. Variances of estimated ln K and ln S_s as a function of iteration for the sandbox experiment.

[36] Table 4 also shows that for evaluation of the estimates over the entire domain, the similarity analysis yields a similar result as other metrics. Nevertheless, we believe that the similarity analysis would have been useful had we targeted the analysis at some specific feature in the domain.

[37] Table 5 tabulates means and variances of true and estimated hydraulic properties of each zone using noise-free, noisy, and denoised hydrographs. It further corroborates the early conclusion that the hydraulic tomography and the SimSLE can depict the hierarchical (nonstationary) heterogeneity satisfactorily.

[38] Finally, we compare snapshots of simulated heads and streamlines in the aquifer with true and estimated parameter fields at 1.5 s after pumping (i.e., early time at which S_s plays an important role) at the center of the aquifer. These simulated fields using true K and S_s parameters are shown in Figure 6a; the fields resulting from the effective K and S_s fields are plotted in Figure 6b; Figure 6c illustrates those fields derived from the estimated K and S_s fields using noisy data; results based on the parameter fields derived from denoised hydrographs are demonstrated in Figure 6d. A visual comparison of Figures 6a–6c and the correlation, L1, and L2 values of the head field listed in Figure 6 further substantiate the usefulness of the HT analysis and denoising procedure.

4. Application to a Laboratory Sandbox Experiment

4.1. Preprocessing Data

[39] The laboratory sandbox experiment (see Figure 7) conducted by *Liu et al.* [2007] and *Illman et al.* [2007, 2008] involved 8 pumping tests. Drawdown data from the 47 ports excluding the pumping port were collected for each pumping test and they were denoised using the wavelet denoising procedure. Data from two pumping tests (at ports 2 and 5) which have small SNRs that yielded anomalous drawdown contours and did not improve the estimates but caused their divergence were discarded. Only the remaining 6 pumping test data sets were used for the HT analysis. To condition the estimation, one K and S_s values from in situ slug test measurements at port 1 (Figure 7) [*Liu et al.*, 2007] were used as the hard data; effective K and S_s (0.1268 cm/s and 8.73×10^{-4} 1/cm) were derived from minimization of equation (14); the variances of ln K and ln S_s were estimated using equation (15) to be 2.0 and 0.1, respectively. The correlation scales were assessed subjectively on the basis of the heterogeneity pattern of the laboratory sandbox (namely, 70 cm and 20 cm in the horizontal and vertical directions, respectively).

[40] During the HT experiment, the total head at the two side boundaries of the sandbox varied slightly (maximum 0.13 cm). To eliminate the effect of a time varying boundary condition, the actual drawdown at observation locations minus the observed boundary drawdown was used as correction to the observed drawdowns. These observed drawdowns were subtracted from an assigned initial total heads (200 cm) to obtain the observed total heads for the HT analysis. Five observed total heads at 0.75 s, 1.50 s, 2.25 s, 3.00 s, and 15.00 s from each observation port during each pumping test were selected for the HT analysis.

4.2. Results

[41] A plot of variances of the estimated ln K and ln S_s fields as a function of the iteration number is illustrated in Figure 8. The variances increased continuously indicating effects of unresolved noise. The final estimates were chosen on the basis of the stabilization of the conditional L2 norm of the head during iteration (Figure 9), which suggest that the best estimates are at the sixth iteration.

[42] Figure 10 shows the distributions of the estimated K and S_s fields, respectively. In Figure 10, six low-K zones in the sandbox (Figure 7) are vividly portrayed by the estimated K field, but the low-K zones close to the bottom are fuzzy. The low resolution at the bottom is due to the no-flux boundary at the bottom where the flow generally follows the boundary, and where the pressure excitations were not sampled because of absence of monitoring ports between the low-K zones and the bottom boundary–consistent with findings by *Illman et al.* [2007, 2008] and *Liu et al.* [2007].

[43] The estimated S_s field in Figure 10 on the other hand does not reflect the pattern of the K field. Rather, the field reflects an overall trend that the S_s values of the medium at the bottom are smaller than those on the top. This pattern appears to be physically correct: sands at the bottom are compressed more because of greater overlying materials. The result is also consistent with the estimated S_s field from the analysis of cross-hole aquifer tests by *Liu et al.* [2007].

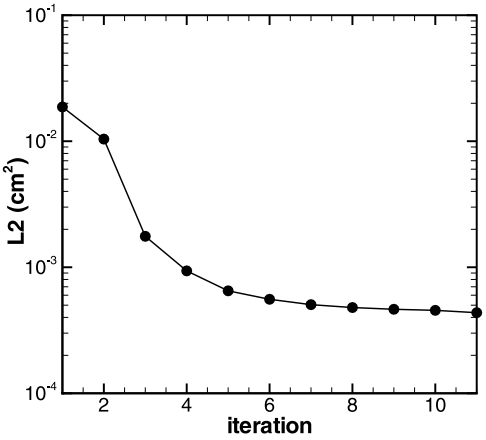

Figure 9. Conditional L2 norm of the head as a function of iteration for the sandbox experiment.

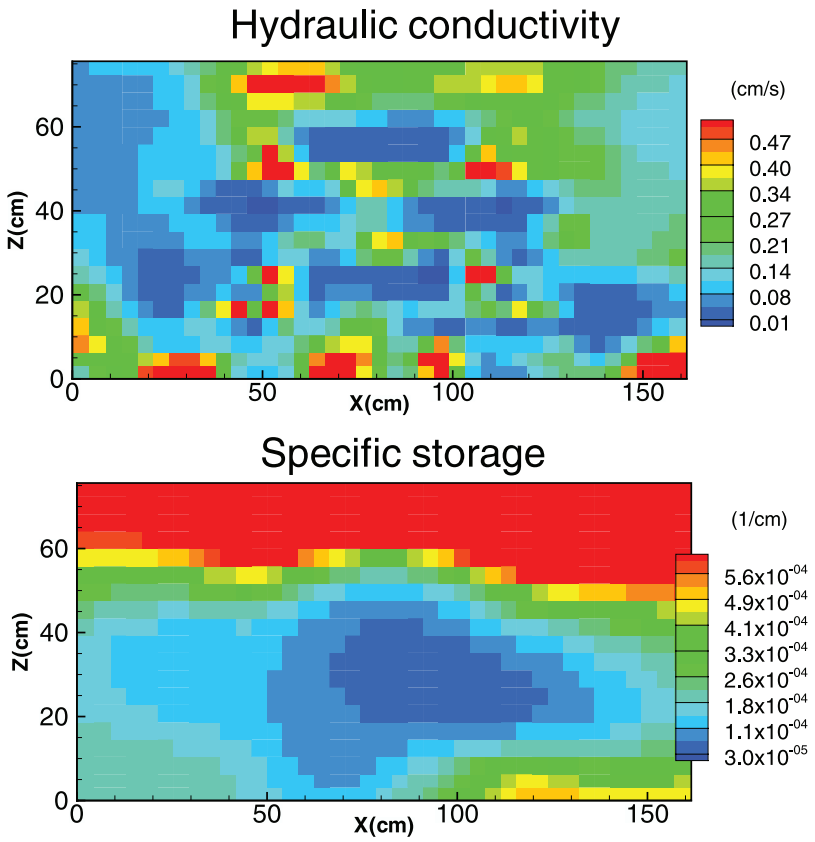

Figure 10. Estimated hydraulic conductivity (cm/s) and specific storage (1/cm) fields for the sandbox experiment.

[44] In this sandbox experiment, the true K and S_s fields are unknown and the performance metrics therefore cannot be evaluated. In order to validate these estimates, we followed the approach by *Liu et al.* [2007] which uses the estimated K and S_s fields as input to the forward flow model to simulate a pumping test conducted at port 46 that was not used in the HT experiment. If the estimated K and S_s fields are representative of the true fields in the sandbox, the temporal and spatial distributions of the simulated head due to the pumping should closely predict the observed ones.

[45] As illustrated in Figure 11, the estimated heterogeneous K and S_s fields yielded heads close to the observed ones at all of the 47 ports at 3.0 s, 6.75 s, 13.5 s, and 27.0 s (early to late time). The large discrepancy at 3.0 s may be attributed to the numerical discretization error as well as inaccuracy of the estimated S_s field, which controls the early time behavior of the drawdown. On the other hand, the effective homogeneous K and S_s fields produced biased heads at all times.

[46] This comparison suggests the following: (1) The effective parameters of the sandbox scale obtained by simultaneously fitting drawdowns from the pumping tests at 6 different locations failed to satisfactorily reproduce drawdowns caused by pumping at another location in the sandbox. Perhaps, they can do better for a stress area (e.g., a production well field) that covers an area much greater than many correlation scales of the heterogeneity. (2) HT in conjunction with SimSLE characterizes the heterogeneity of aquifers sufficiently such that drawdown evolution due to a different pumping event is predicted. We believe these results are significant. First, not only do the results reinforce the validity of HT, SimSLE and our data processing approach, but they also demonstrate a need for fine-resolution mapping of K and S_s fields to overcome the phenomenological nature associated with the domain-scale effective parameters. Besides, these results confirm that the classical governing groundwater flow equation can yield excellent predictions of drawdowns in a heterogeneous sandbox when the K and S_s fields are adequately characterized and the initial and boundary conditions as well as source/sink terms are fully prescribed. The result supports a similar conclusion reached by *Liu et al.* [2007].

5. Discussion

[47] Distinguishing noise from effects of heterogeneity in a hydrograph can be highly subjective unless characteristics of noise or heterogeneity are known a priori. Complete removal of noise from hydrographs is difficult and unresolved noise residuals can impact the estimation. The impact is manifested through a continuous increase in the variances of estimated hydraulic properties but their spatial pattern remains almost constant. As a result, the head field changes continuously but its L2 norm stabilizes. Stabilization of the conditional L2 norm of the head thus works well as the convergence criterion for our SimSLE when the data are infested with noise.

[48] Finally, although no explicit comparison with SSLE has been presented, we believe that SimSLE has several advantages over the SSLE. (1) SimSLE needs to evaluate

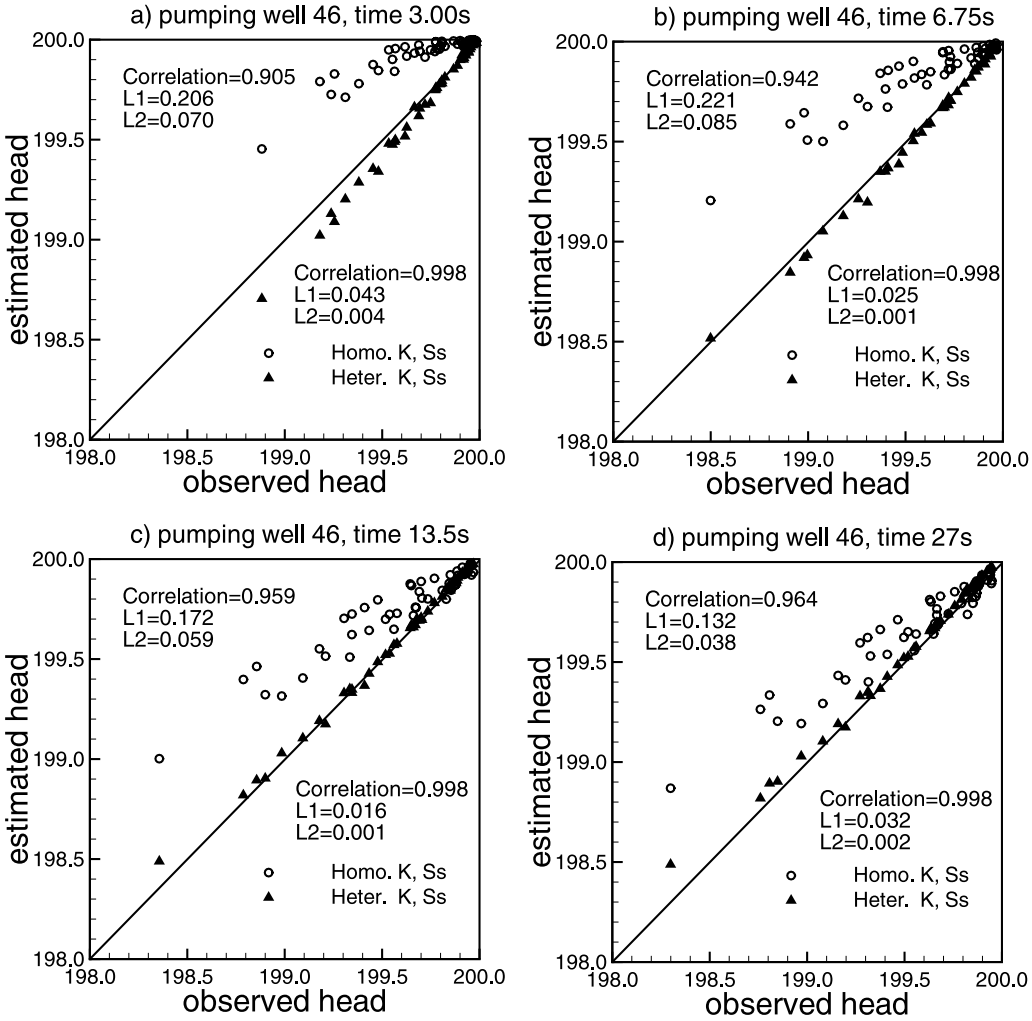

Figure 11. Validation results: observed versus estimated heads (cm) at the 47 observation ports at four different times in the sandbox. Two simulated heads were considered: one using estimated effective K and S_s fields from the equivalent homogeneous domain and the other using the estimated heterogeneous K and S_s fields derived from the analysis of HT.

the adjoint state equation only once for a given observation location using newly estimated hydraulic property fields from all pumping tests since the adjoint state equation is independent of the pumping rate and pumping location. On the other hand, using SSLE, one must solve the adjoint state equation for each pumping test because the parameters in the adjoint state equation are modified for each pumping test. (2) SimSLE avoids the loop iteration of SSLE, and the computational effort is thus reduced. (3) Adding data sets in different sequences in SSLE may lead to a slightly different final result, which is more sensitive to the last data set [see *Illman et al.*, 2008]. This problem does not exist in SimSLE. (4) SimSLE uses all observations simultaneously, providing more constrains for the inverse problem and thus converges faster than SSLE.

[49] The disadvantages of SimSLE are as follows: (1) Since all head data sets are used simultaneously and the same convergence criteria are applied to reach the final estimate in SimSLE, one bad data set may affect the overall quality of the estimate. Note that both SimSLE and SSLE result in the same estimate if the data sets are free of errors. (2) The memory requirement is greater because the sizes of covariance matrix of h and the cross-covariance matrix of f and h are larger than the corresponding matrices in SSLE.

6. Conclusion

[50] Results of this study show the following: (1) in spite of noise in hydrographs from a HT survey, the least squares approach can satisfactorily estimate effective hydraulic properties of the synthetic aquifer with hierarchical heterogeneity because the inverse problem is well posed. (2) Accurate estimation of spatial variances of K and S_s from HT data is difficult because of the nonstationary nature of the flow field, which demands a large amount of head data to obtain representative sample head variances. (3) For ill-posed problems, HT in conjunction with our SimSLE yield satisfactory estimates of the hierarchical K and S_s fields of the synthetic aquifer. If hydrographs are corrupted with noise, the SNR is a useful measure of reliability of a corrupted hydrograph, and wavelet denoising the hydrographs is a viable means to improve the estimate. In addition, the use of stabilization of the conditional L2 norm of head as a

convergence criterion in SimSLE avoids overexploitation of noisy data. (4) HT surveys delineate detailed hydraulic heterogeneity in aquifers, which can be used to predict different flow scenarios. That is, the estimate hydraulic properties do not suffer from the phenomenological nature associated with the domain-scale effective properties. Finally, simultaneous inclusion of hydrographs from all pumping tests in the analysis offers some advantages over the previous sequential approach but it suffers from the requirement of huge computational resources.

[51] **Acknowledgments.** The second author would like to acknowledge support from the National Cheng-Kung University during his visits in 2006, 2007, and 2008. Support from the Strategic Environmental Research and Development Program (SERDP) subcontracted through the University of Iowa as well as funding from the National Science Foundation (NSF) through grants EAR-0229717, IIS-0431079, and EAR-0450388 are acknowledged. We thank constructive and insightful suggestions by Anna Michalak, Walter Illman, and two anonymous reviewers as well as Resta Cheng for proofreading the manuscript.

References

Bakr, A. A., L. W. Gelhar, A. L. Gutjahr, and J. R. MacMillan (1978), Stochastic analysis of spatial variability in subsurface flows: 1. Comparison of one- and three-dimensional flows, *Water Resour. Res.*, *14*(2), 263–271, doi:10.1029/WR014i002p00263.

Barrash, W., and T. Clemo (2002), Hierarchical geostatistics and multifacies systems: Boise Hydrogeophysical Research Site, Boise, Idaho, *Water Resour. Res.*, *38*(10), 1196, doi:10.1029/2002WR001436.

Bohling, G. C., X. Zhan, J. J. Butler Jr., and L. Zheng (2002), Steady shape analysis of tomographic pumping tests for characterization of aquifer heterogeneities, *Water Resour. Res.*, *38*(12), 1324, doi:10.1029/2001WR001176.

Bohling, G. C., J. J. Butler Jr., X. Zhan, and M. D. Knoll (2007), A field assessment of the value of steady shape hydraulic tomography for characterization of aquifer heterogeneities, *Water Resour. Res.*, *43*, W05430, doi:10.1029/2006WR004932.

Brauchler, R., R. Liedl, and P. Dietrich (2003), A travel time based hydraulic tomographic approach, *Water Resour. Res.*, *39*(12), 1370, doi:10.1029/2003WR002262.

Fienen, M. N., T. Clemo, and P. K. Kitanidis (2008), An interactive Bayesian geostatistical inverse protocol for hydraulic tomography, *Water Resour. Res.*, *44*, W00B01, doi:10.1029/2007WR006730.

Gottlieb, J., and P. Dietrich (1995), Identification of the permeability distribution in soil by hydraulic tomography, *Inverse Probl.*, *11*, 353–360, doi:10.1088/0266-5611/11/2/005.

Hagen, A. (2003), Fuzzy set approach to assessing similarity of categorical maps, *Int. J. Geogr. Inf. Sci.*, *17*(3), 235–249, doi:10.1080/13658810210157822.

Illman, W. A., X. Liu, and A. Craig (2007), Steady-state hydraulic tomography in a laboratory aquifer with deterministic heterogeneity: Multimethod and multiscale validation of K tomograms, *J. Hydrol.*, *341*(3–4), 222–234, doi:10.1016/j.jhydrol.2007.05.011.

Illman, W. A., A. J. Craig, and X. Liu (2008), Practical issues in imaging hydraulic conductivity through hydraulic tomography, *Ground Water*, *46*(1), 120–132, doi:10.1111/j.1745-6584.2007.00374.x.

Kitanidis, P. K. (1995), Quasi-linear geostatistical theory for inversing, *Water Resour. Res.*, *31*(10), 2411–2420, doi:10.1029/95WR01945.

Kuhlman, K. L., A. C. Hinnell, P. K. Mishra, and T.-C. J. Yeh (2008), Basin-scale transmissivity and storativity estimation using hydraulic tomography, *Ground Water*, *46*(5), 706–715, doi:10.1111/j.1745-6584.2008.00455.x.

Li, W., A. Englert, O. A. Cirpka, J. Vanderborght, and H. Vereecken (2007), Two-dimensional characterization of hydraulic heterogeneity by multiple pumping tests, *Water Resour. Res.*, *43*, W04433, doi:10.1029/2006WR005333.

Li, W., A. Englert, O. A. Cirpka, and H. Vereecken (2008), Three-dimensional geostatistical inversion of flowmeter and pumping test data, *Ground Water*, *46*(2), 193–201, doi:10.1111/j.1745-6584.2007.00419.x.

Liu, S., T.-C. J. Yeh, and R. Gardiner (2002), Effectiveness of hydraulic tomography: Sandbox experiments, *Water Resour. Res.*, *38*(4), 1034, doi:10.1029/2001WR000338.

Liu, X., W. A. Illman, A. J. Craig, J. Zhu, and T.-C. J. Yeh (2007), Laboratory sandbox validation of transient hydraulic tomography, *Water Resour. Res.*, *43*, W05404, doi:10.1029/2006WR005144.

Mallat, S. (1999), *A Wavelet Tour of Signal Processing*, 577 pp., Academic, San Diego, Calif.

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery (1992), *Numerical Recipes in FORTRAN 77: The Art of Scientific Computing*, 933 pp., Cambridge Univ. Press, New York.

Straface, S., T.-C. J. Yeh, J. Zhu, S. Troisi, and C. H. Lee (2007), Sequential aquifer tests at a well field, Montalto Uffugo Scalo, Italy, *Water Resour. Res.*, *43*, W07432, doi:10.1029/2006WR005287.

Tosaka, H., K. Masumoto, and K. Kojima (1993), Hydropulse tomography for identifying 3-D permeability distribution, in *Proceedings of the 4th Annual International Conference on High Level Radioactive Waste Management; Apr 26–30 1993; Las Vegas, NV, USA, High Level Radioactive Waste Management*, edited by B. M. Cole, pp. 955–959, Am. Soc. of Civ. Eng., New York.

Vasco, D. W., H. Keers, and K. Karasaki (2000), Estimation of reservoir properties using transient pressure data: An asymptotic approach, *Water Resour. Res.*, *36*(12), 3447–3465, doi:10.1029/2000WR900179.

Vesselinov, V. V., S. P. Neuman, and W. A. Illman (2001), Three-dimensional numerical inversion of pneumatic cross-hole tests in unsaturated fractured tuff: 1. Methodology and borehole effects, *Water Resour. Res.*, *37*(12), 3001–3017, doi:10.1029/2000WR000133.

Wu, C.-M., T.-C. J. Yeh, J. Zhu, T. H. Lee, N.-S. Hsu, C.-H. Chen, and A. F. Sancho (2005), Traditional analysis of aquifer tests: Comparing apples to oranges?, *Water Resour. Res.*, *41*, W09402, doi:10.1029/2004WR003717.

Ye, M., R. Khaleel, and T.-C. J. Yeh (2005), Stochastic analysis of moisture plume dynamics of a field injection experiment, *Water Resour. Res.*, *41*, W03013, doi:10.1029/2004WR003735.

Yeh, T.-C. J., and S. Liu (2000), Hydraulic tomography: Development of a new aquifer test method, *Water Resour. Res.*, *36*(8), 2095–2105, doi:10.1029/2000WR900114.

Yeh, T.-C. J., R. Srivastava, A. Guzman, and T. Harter (1993), A numerical model for two dimensional flow and chemical transport, *Ground Water*, *31*(4), 634–644, doi:10.1111/j.1745-6584.1993.tb00597.x.

Yeh, T.-C. J., M. Jin, and S. Hanna (1996), An iterative stochastic inverse method: Conditional effective transmissivity and hydraulic head fields, *Water Resour. Res.*, *32*(1), 85–92, doi:10.1029/95WR02869.

Yeh, T.-C. J., C.-H. Lee, K.-C. Hsu, and Y.-C. Tan (2007), Fusion of active and passive hydrologic and geophysical tomographic surveys: The future of subsurface characterization, in *Data Integration in Subsurface Hydrology*, *Geophys. Monogr. Ser.*, vol. 171, edited by D. W. Hyndman, F. D. Day-Lewis, and K. Singha, pp. 109–120, AGU, Washington, D. C.

Zhang, Q., R. Aliaga-Rossel, and P. Choi (2006), Denoising of gamma-ray signals by interval-dependent thresholds of wavelet analysis, *Meas. Sci. Technol.*, *17*, 731–735, doi:10.1088/0957-0233/17/4/019.

Zhu, J., and T.-C. J. Yeh (2005), Characterization of aquifer heterogeneity using transient hydraulic tomography, *Water Resour. Res.*, *41*, W07028, doi:10.1029/2004WR003790.

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

K.-C. Hsu and C.-H. Lee (corresponding author), Department of Resources Engineering, National Cheng Kung University, Tainan, 701, Taiwan. (leech@mail.ncku.edu.tw)

J.-C. Wen, Department of Environmental and Safety Engineering, Research Center for Soil and Water Resources and Natural Disaster Prevention, National Yunlin University of Science and Technology, Touliu, Yunlin, 64045, Taiwan.

J. Xiang and T.-C. J. Yeh, Department of Hydrology and Water Resources, University of Arizona, Tucson, AZ 85721-0001, USA.