Robustness of joint interpretation of sequential pumping tests: Numerical and field experiments

Shao-Yang Huang,¹ Jet-Chau Wen,² Tian-Chyi J. Yeh,³ Wenxi Lu,⁴ Hsiang-Lan Juan,⁵ Chung-Min Tseng,⁵ Ju-Huang Lee,⁵ and Kuo-Chyang Chang⁵

Received 31 March 2011; revised 30 August 2011; accepted 19 September 2011; published 28 October 2011.

[1] Three conceptual models are evaluated for estimating transmissivity (T) fields using data from sequential pumping tests at a field site and data from similar tests simulated in a synthetic aquifer. The three approaches are (1) an equivalent homogeneous approach, (2) a heterogeneous approach based on a single pumping test, and (3) a heterogeneous approach based on joint interpretation of the sequential pumping tests (i.e., hydraulic tomography, HT). They are evaluated on the basis of their abilities to obtain representative estimates of the T field of the aquifer and, more importantly, on the ability of their estimates to predict drawdown distributions in the aquifer induced by independent validation pumping tests. Results show that the first approach yields scenario-dependent T estimates, which vary with the location of the pumping well. Independent validation tests show that the predicted drawdowns in both aquifers are biased and dispersed. While the second approach produces scenario-dependent T spatial distributions capturing the general pattern of the aquifer, the T fields consistently yield better drawdown predictions than those based on the first approach. Lastly, the joint interpretation approach reduces the scenario dependence of the T estimates and improves the quality of the T estimates as more data sets from sequential pumping tests are included. More importantly, the resultant T estimates lead to the best prediction of different flow events. The robustness of the joint interpretation is then elucidated.

Citation: Huang, S.-Y., J.-C. Wen, T.-C. J. Yeh, W. Lu, H.-L. Juan, C.-M. Tseng, J.-H. Lee, and K.-C. Chang (2011), Robustness of joint interpretation of sequential pumping tests: Numerical and field experiments, *Water Resour. Res.*, 47, W10530, doi:10.1029/2011WR010698.

1. Introduction

[2] Knowledge of hydraulic properties of aquifers is essential to groundwater resource management and groundwater contamination prevention and remediation. Numerous methods have been developed over the past several decades for determining aquifer hydraulic properties. One popular method is the so-called aquifer test with the use of the analytical solution by *Theis* [1935] and that by *Cooper-Jacob* [1946] for analysis. These solutions assume that the aquifer is homogeneous and has infinite lateral extents, although aquifers are inherently heterogeneous and bounded in nature. As a result, it is generally believed that the hydraulic properties estimated by these methods represent spatially averaged hydraulic properties over the cone of depression

0043-1397/11/2011WR010698

[e.g., *Meier et al.*, 1998; *Sanchez-Vila et al.*, 1999]. On the other hand, *Wu et al.* [2005] showed that the analysis yields ambiguous spatially averaged hydraulic properties of an aquifer, which vary with the location of observation. Analyses of field pumping tests by *Straface et al.* [2007] and *Wen et al.* [2010] confirm the finding. *Wu et al.* [2005] further suggested that in order to obtain representative hydraulic properties of a heterogeneous aquifer within the cone of depression using the Theis solution, hydrographs from many observation wells must be used simultaneously.

[3] To improve our ability to characterize aquifers, a hydraulic tomographic (HT) survey has been developed [Gottlieb and Dietrich, 1995; Vasco et al., 2000; Yeh and Liu, 2000; Bohling et al., 2002; Brauchler et al., 2003, 2010; Zhu and Yeh, 2005; Li et al., 2007; Liu et al., 2007; Illman et al., 2007, 2010; Fienen et al., 2008; Castagna and Bellin, 2009; Cardiff et al., 2009; Yin and Illman, 2009]. HT is a series of cross-well interference tests (or sequential pumping tests) using a well field to the maximum extent in terms of hydraulic tests.

[4] HT has been tested successively in synthetic porous aquifers [*Yeh and Liu*, 2000; *Zhu and Yeh*, 2005], and fractured aquifers [*Hao et al.*, 2008], laboratory sandboxes [*Liu et al.*, 2002, 2007; *Illman et al.*, 2007, 2010; *Xiang et al.*, 2009; *Liu and Kitanitdis*, 2011] plot-scale fields [*Bohling et al.*, 2007; *Straface et al.*, 2007; *Li et al.*, 2007, 2008; *Brauchler et al.*, 2010, 2011], and a fractured granite field site [*Illman et al.*, 2009]. In particular, *Liu et al.* [2007] and

¹Graduate School of Engineering Science and Technology, National Yunlin University of Science and Technology, Douliou, Yunlin, Taiwan.

²Department and Graduate School of Safety Health and Environment Engineering, Research Center for Soil and Water Resources and Natural Disaster Prevention, National Yunlin University of Science and Technology, Douliou, Yunlin, Taiwan. ³Department of Hydrology and Water Resources, University of Arizona,

³Department of Hydrology and Water Resources, University of Arizona, Tucson, Arizona, USA.

⁴Department of Groundwater Science and Engineering, College of Environment and Resources, Jilin University, Changchun, Jilin, China. ⁵Water Resource Agency, Ministry of Economic Affairs, Taipei, Taiwan.

Copyright 2011 by the American Geophysical Union.

Xiang et al. [2009] show that the transient HT not only identifies the pattern of the heterogeneous hydraulic conductivity (K) and the specific storage (S_s) field of the sandbox but also yields the estimates that can satisfactorily predict the drawdown caused by a pumping test that was not used in the HT analysis.

[5] Illman et al. [2010] and Berg and Illman [2011] evaluated different aquifer characterization methods (including the Theis aquifer test, kriging using core 48 samples, kriging with 48 single well test estimates, and HT) for characterizing a heterogeneous K distribution in a fully saturated sandbox. The performance of these methods was evaluated on the basis of their ability to predict steady state drawdowns at the 48 sampling locations of 16 independent pumping tests at different locations in the sandbox. Note, these drawdowns were not used in the inverse analysis of HT. They concluded that the K estimate from kriging using 48 core samples yielded satisfactory predictions. However, the estimate based on HT yielded the best predictions.

[6] While the robustness of HT has been substantiated by these sandbox experiments and some recent field applications, its ability for high-resolution aquifer characterization under field conditions remains to be fully assessed. Especially, the benefits of HT have not been demonstrated and evaluated in comparison with other approaches under field conditions. As a result, the objective of this paper is to quantitatively evaluate three approaches for estimating Tfields using data from sequential pumping tests at a field site and data from similar tests simulated in a synthetic aquifer. The synthetic aquifer case represents the ideal case where true T field, boundary conditions, head measurements, and pumping rates are exactly known and noisefree. These three approaches to be evaluated include (1) the equivalent homogeneous approach, (2) a heterogeneous approach based on a single pumping test, and (3) a heterogeneous approach based on joint interpretation of the sequential pumping tests (i.e., HT). For the synthetic aquifer case, the estimated T fields from the three approaches are appraised by comparing them with the true field. The robustness of these estimated T fields of the synthetic and field aquifers are subsequently validated in terms of their ability to predict drawdown distributions of the aquifers induced by independent validation pumping tests. This type of validation of T estimates from field pumping tests has rarely been carried out previously. Finally, insights to joint interpretation of sequential pumping tests are discussed with respect to other approaches.

2. Sequential Pumping Tests

2.1. Field Experiment

[7] Sequential pumping tests with multiple observation wells were conducted at an experimental field site on the north side of the campus of National Yunlin University of Science and Technology (NYUST), located in the western central part of Taiwan. The site is located at the Cho-shui River alluvial fan. A detailed description of the field site and geology is given by *Wen et al.* [2010] and only a brief description is given here. According to *Wen et al.* [2010], the alluvial aquifer at the site consists of 16.4 m of sand overlying a clay layer and is overlain by 3.4 m of silty clay. The aquifer is considered to be a confined aquifer. Eleven

boreholes (BH01 through BH11) were drilled 20 m in depth from the ground surface at the site. A schedule 40 PVC pipe of 4-inch diameter was used for the well screen and casing, and the borehole was screened with a 0.02-inch slotted screen. The screen length was 18.50 m and was placed 1.50 m below the ground surface and extended down to the 20.00 m depth. All boreholes were distributed throughout the 10×10 m area of the site with BH04 being almost at the center of the site. A plan view of the layout of these boreholes is shown in Figure 1.

[8] During each of the sequential pumping tests, water was pumped from a given well at a constant pumping rate and drawdown-time data were collected at the 10 boreholes, excluding the pumping well. A transducer with a data logger, with a precision of 1 mm, was installed in each well to measure the drawdown. The pumping rate ranged from $8 \times 10^{-5} \text{ m}^3 \text{ s}^{-1}$ to $2.06 \times 10^{-4} \text{ m}^3 \text{ s}^{-1}$. The different pumping rates were used to ensure observable drawdown during each pumping test. In order to record a rapid head change at the beginning of a pumping test, the drawdown was recorded each second for a period of 380 min from the beginning of pumping. Subsequently, the recording time step was increased to every minute until the pumping test terminated. Each pumping test was continued for at least 72 h to ensure that a steady flow condition was achieved. Afterward, the pump was shut off to allow the aquifer to recover and the head was recorded during groundwater recovery. The pump was then moved to another well, and the same procedure was repeated once the groundwater recovery was completed. Overall, 11 pumping tests were carried out, yielding 11 sets of drawdown-time data including 110 drawdown-time curves. Figure 2 illustrates the 10 drawdown-time curves collected at 10 observation wells induced by pumping at BH03. The steady state condition was reached after 50,000 s (13.8 h).



Figure 1. The layout of the well field at NYUST field site.



Figure 2. The drawdown-time curves of pumping test (BH03 pumping well).

2.2. Numerical Experiment

[9] Sequential pumping tests similar to those of the NYUST field experiment were also simulated in a synthetic 2-D plane aquifer. The synthetic aquifer had a dimension of 21 m \times 21 m and it was discretized into 21 \times 21 elements of $1 \text{ m} \times 1 \text{ m}$ in size. Each element was assigned a T value, using a random field generator [Gutjahr, 1989] assuming a lognormal joint probability distribution. The random InT (natural logarithm of T) field has a mean of -6.2, a variance of 1.67, and an isotropic exponential covariance function with a correlation scale equal to 5 m. These values were selected based on the results of the preliminary analysis of data from the NYUST site. The corresponding mean and variance of the T field (linear scale) are 0.00444 m² s⁻¹ and 4.1×10^{-5} m⁴ s⁻², respectively. The geometric and harmonic means of the *T* field are 0.00203 m² s⁻¹ and $0.000946 \text{ m}^2 \text{ s}^{-1}$, respectively, and the spatial distribution of T is shown in Figure 3a. All sides of the aquifer are assumed to have a constant head boundary condition with a head equal to 100 m. The number of wells and their locations in this synthetic aquifer are intentionally made to be identical to those of the field site at NYUST. Sequential pumping tests were simulated using a finite element model VSAFT2 [Yeh et al., 1993] (see section 3), which in principle mimics the field sequential pumping tests. The simulation of the test was carried out using a steady state approach with a pumping rate of $0.000231 \text{ m}^3 \text{ s}^{-1}$. The simulated steady state drawdowns at the observation wells are then sampled and used in the following inverse modeling exercises.

[10] The purpose of this numerical experiment is to provide a direct assessment of the performance of different aquifer characterization approaches under idealized situations where the true T distribution of the aquifer is known and data sets are error-free. Results of the analysis of the idealized data sets are to be used for comparison with those from the real-world aquifer tests at the NYUST field aquifer. The results from the field aquifer likely involve the effects of measurement errors, uncertainty in boundary conditions and pumping rates, errors in mathematical model (e.g., depth-averaged approach), scale differences, and processes which are not considered in the model. Any similarity between the two corroborates the results of the NYUST experiments; any differences may be attributed to the errors.

3. Analysis of the Sequential Pumping Tests

[11] Steady state drawdown data at the 10 observation wells during each of the seven pumping tests at boreholes (BH01, BH02, BH03, BH06, BH07, BH08, and BH09) in either the synthetic or the NYUST aquifer were used in the analysis. The remaining four tests were reserved for validation purposes (see section 5.3). The steady state drawdown of each observation well used in the inverse modeling of the NYUST aquifer was the average value of the drawdowns recorded over a 48-h time interval after the drawdown stabilized (see Figure 2). Three approaches were used to analyze the drawdown-time data sets from the sequential pumping test: (1) an equivalent homogeneous conceptual model, (2) a heterogeneous model based on



Figure 3. (a) The true T field of the synthetic aquifer. The others are the estimated T fields based on heads observed at 10 wells induced by pumping at a given well (the white square in each figure) (approach 2).

individual pumping test, and (3) a heterogeneous model based on joint interpretation of all of the sequential pumping tests (i.e., HT).

3.1. Approach 1: Equivalent Homogeneous Approach

[12] This approach conceptualizes a heterogeneous aquifer as an equivalent homogenous aquifer in the same way as a widely used traditional aquifer test analyses (e.g., Thiem, Jacob, and Theis approaches). The objective of this approach is to estimate the effective or apparent T for the equivalent homogeneous aquifer, which will mimic the average drawdown over the entire aquifer during a given pumping test. *Meier et al.* [1998] and *Sanchez-Vila et al.* [1999] showed that applications of the Jacob method to drawdown-time data at observation wells at different locations during pumping at a well in a synthetic heterogeneous aquifer yield a very similar transmissivity estimate. On the contrary, *Wu et al.* [2005] argued that the heads implied in the governing equation for flow through an equivalent homogeneous aquifer represent the spatial trend of the heads in a heterogeneous aquifer, while the head observed at a single borehole in the heterogeneous aquifer represents the perturbation around the trend. Analyzing heads observed at a single borehole in a heterogonous aquifer with the Theis solution, which is built upon the equivalent homogeneous conceptual model, is therefore comparing apples (perturbations) and oranges (trends). Wu et al. [2005] and Wen et al. [2010] advocated that a large number of spatial head observations in the aquifer must be used to obtain a representative effective parameter for the equivalent homogeneous aquifer. Following this concept, we estimated the effective T for the aquifer using steady state drawdowns at 10 observation wells during each of the seven pumping tests. Because the medium is assumed to be homogenous and drawdown and discharge (pumping rate) are known, uniqueness of the solution to this inverse problem under this condition is guaranteed [Nelson, 1960]. Moreover, the number of observations is greater than the number of parameters to be estimated.

W10530

Estimating T for each test therefore is an overdetermined and well-defined inverse problem [*Yeh et al.*, 2011], which has a unique solution. According to *Meier et al.* [1998], *Sanchez-Vila et al.* [1999], *Wu et al.* [2005], and *Wen et al.* [2010], this approach should yield the same T estimate in spite of different pumping locations. The estimate should predict head fields that capture the general flow pattern induced by any stress. That is, the predicted head field for other pumping events will be unbiased but scattered. The scatter reflects the effects of heterogeneity.

3.2. Approach 2: Heterogeneous Approach Based on Individual Pumping Test

[13] This approach adopts a heterogeneous conceptual model to estimate the T value at every location in the aquifer using each pumping test and its steady heads at 10 observation wells, excluding the head at the pumping well. As a result, this is a highly parameterized inverse problem. Since the number of parameters (T value in each element of a finite element model for the aquifer) to be estimated is much greater than the number of head observations for a given pumping test, the inverse problem thus is ill-defined (or underdetermined) [*Yeh et al.*, 2011]. For such an ill-defined inverse problem, as to be discussed in section 3.4, successive linear estimator algorithm (SLE) based on geostatistics is an appropriate method for obtaining a conditional mean solution with a specified unconditional mean, and a covariance function conditioned on observed heads [*Yeh et al.*, 1996].

3.3. Approach 3: Heterogeneous Approach Based on Joint Inversion of Pumping Tests

[14] Different from approach 2, this approach will estimate T at every location in the aquifer by including several pumping test data sets from the sequential pumping tests simultaneously (i.e., HT). According to the principle of HT, as more nonredundant pumping test data sets from the sequential pumping tests are included in the analysis, the resolution of the estimated T field will improve. This is true even though the problem is highly parameterized and ill-posed.

3.4. Mathematical Models

[15] VSAFT2 (available at http://www.hwr.arizona.edu/ yeh) was used for both forward and inverse modeling exercises in this study. The model is a two-dimensional, finite element model that solves governing equations for flow and transport in a geologic media [Yeh et al., 1993]. To conduct the numerical analysis, the NYUST aquifer was treated as a two-dimensional, depth-averaged aquifer. The horizontal dimension of the aquifer used in the analysis was 51 m \times 51 m and descritized into 2601 elements with $1 \text{ m} \times 1 \text{ m}$ in size. The initial total head (head hereafter) data observed at the 11 wells prior to the sequential pumping tests were interpolated onto each node of the grid and used as the initial steady head condition of the aquifer at the site. The extrapolated head at the boundary grids was assumed to be the constant boundary head condition of the aquifer. The extrapolation yields a value of 43.88 m in head for the north and east boundaries and 43.65 m for the head at the south and west boundaries. Note that in the following discussions of the NYUST aquifer, only an area of 21 m \times 21 m centered on the aquifer will be considered. For the synthetic aquifer, the inverse model setup is identical to those described in the numerical experiment section (2.2).

[16] The inverse algorithm used here is SimSLE [Xiang et al., 2009], which can simultaneously include results from several pumping tests to estimate the hydraulic property, and is based on the successive linear estimator algorithm (SLE), developed by Yeh et al. [1996]. This algorithm is included in VSAFT2. Here a brief description of the SLE is presented. Mathematically, SLE is a stochastic estimator that seeks the conditional approximate mean field of a stochastic parameter field. Specifically, it estimates the parameter value at any location successively utilizing the linear combination of weighted measurements of parameters and/or heads. The weights are determined from the spatial correlation between the head and the parameter, the spatial correlation between heads, and that between the parameters. With a given mean and covariance function of a stochastic field, and observed heads, SLE always yields a unique estimate since it is the conditional expectation (the best unbiased estimate) as opposed to conditional realizations. Conditional realizations are possible realizations of parameter fields that can yield head fields that honor observed heads and measured parameters if any (see e.g., Hanna and Yeh, 1998; Castanaga and Bellin, 2009; Bohling and Butler, 2010). On the other hand, the conditional expectation is the average of all conditional realizations. When the number of observations is greater than or equal to the number of parameters to be estimated and other necessary conditions are met [Yeh et al., 2011], only one conditional realization exists. This conditional realization is the conditional mean and is the true field. Conversely, if the necessary conditions are not met, a large number of conditional realizations exist, which have a common, unique conditional mean field. This conditional mean field is not the true field but it is the best unbiased estimate that honors all observations. Deviations from the true field are then quantified by the residual covariance.

[17] Inverse modeling using the SLE of VSAFT2 requires specification of values for the mean and variance of T, and a correlation structure. In approach 1, we used the mean and variance of the synthetic aquifer for either aquifer case. The effective T obtained in approach 1 for a given pumping test (Table 1) was then used as the mean for the analysis in approach 2 for the corresponding pumping test. The

Table 1. Homogeneous *T* Value Estimated Using a Single Pumping Test (Approach 1)^a

Pumping Wells	Synthetic Aquifer (10 Heads) $(m^2 s^{-1})$	Synthetic Aquifer (441 Heads) $(m^2 s^{-1})$	NYUST Field Site $(m^2 s^{-1})$
BH01	0.00579	0.00586	0.00334
BH02	0.00328	0.00326	0.00249
BH03	0.00326	0.00337	0.00188
BH06	0.00154	0.00114	0.00121
BH07	0.00264	0.00094	0.00209
BH08	0.00394	0.00431	0.00072
BH09	0.00374	0.00444	0.00137
Mean	0.00346	0.00333	0.00187
Variance	1.69E-06	3.18E-06	7.67E-07
Seven pumping wells	0.00300	0.00215	0.00164

^aShown are the *T* estimate variations between the locations of the pumping well.

variance is the same as that in approach 1. Approach 3 used the same mean and variance as those of approach 1. The correlation structure in all analyses was assumed to be an exponential function with a correlation scale equal to one finite element length in all directions. The stopping criterion for the iteration of SimSLE is the stabilization of the L2 norm of the drawdown [see *Xiang et al.*, 2009].

3.5. Performance Statistics

[18] The standard correlation coefficient (COR) $(1 \ge |r| \ge 0)$, mean absolute error (L1 norm), and mean square error (L2 norm) were the performance statistics used to measure the similarity between the true and the estimated T field for the cases involving the synthetic aquifer where the true T field is known. These metrics are also employed to evaluate the similarity between the predicted and the observed drawdown fields of the four validation tests in the synthetic and NYUST aquifers.

4. Estimation Results

4.1. Approach 1

[19] Table 1 shows seven estimated effective T values corresponding to the seven pumping tests of the synthetic and NYUST aquifers. The mean and variance of the seven estimates of each aquifer are also listed in the table. The last row of Table 1 lists the effective T estimate for each aquifer based on simultaneous consideration of all draw-downs of the seven pumping tests. According to the table, the effective T estimate from each pumping test changes with the pumping location. That is, the inverse solution is scenario-dependent, varying with the pumping location, even though the estimate is unique for a given pumping location. This result holds for both the synthetic and NYUST aquifers.

[20] To further ensure that this scenario-dependent issue is not caused by insufficient head observations used in the estimation, the head value of every node of the synthetic aquifer (a total of 441 head values excluding boundary nodes) was subsequently used to estimate the effective Tvalue for each pumping test. Estimates are also tabulated in Table 1, and again show their scenario dependence.

[21] According to the table, the mean of the seven estimates based on 441 heads is slightly larger than the geometric mean (0.00203) and smaller than the arithmetic mean (0.00444) of the true T field (see section 2.2). Likewise, the estimate of the effective T from each pumping test also varies around the arithmetic and geometric means of the true T field. These results suggest that each effective T estimate based on each pumping test is indeed some weighted average of all local T values over the entire domain, but the weight is affected by the pumping well location. As a matter of fact, this finding is implicit in the cross-correlation map (Figure 10b in the work of Wu et al., [2005]). This map shows that the head at any location within an aquifer during the late time of a pumping test at another given location is not only correlated with T's everywhere within the cone of depression, but is also highly correlated with T's near the pumping and observation wells. In other words, the effective T based on the head at one location or heads at many locations will be greatly influenced by the heterogeneity near the pumping well. This finding is a new insight that supplements the conclusion by *Wu et al.* [2005].

4.2. Approach 2

[22] The estimated T field for the aquifer using approach 2, with the heads sampled at 10 observation wells during each of the seven pumping tests, are shown in Figures 3b–3h, which correspond to the pumping well locations BH01, BH02, BH03, BH06, BH07, BH08, and BH09, respectively.

[23] At first glance of the seven figures, each apparently captures the general pattern of the true T field (Figure 3a): the zone of low T values (blue) extending from the top boundary to the bottom right and the zone of intermediate Tvalues (green) on the sides are consistently present in each figure. Noticeably, these patterns are not restricted to the area close to the wells where heads were observed and used in the estimation; they also cover areas near the boundaries of the aquifer and far away from the wells. These results indicate that the head at any point within an aquifer during the late time of a pumping test is influenced by T values even far away from the observation and near the boundary. This is also in agreement with the cross-correlation map of Wu et al. [2005], but in contrast to the results of Bohling and Butler [2010]. The explanation for the difference will be given in the discussion.

[24] A second look at these figures reveals that each estimated T field is different from the others even though all estimated fields bear some similarities. Table 2 tabulates the geometric mean, the arithmetic mean, and the variance of each estimated T field of the synthetic aquifer from each single pumping test. Since the true T field is known, the performance statistics (L1, L2, and COR) for the comparison between the true T field and each estimated field are listed in Table 2 as well. These figures, the means and variance, and performance statistics all indicate that the T field estimated using approach 2 based on each pumping test varies with the location of the pumping well. The mean value of the estimated T field for each pumping test is also different from the effective T value from approach 1 for the same test. Notice that the head data set of the synthetic aquifer is free of measurement and physical model errors, and there is no uncertainty in the boundary condition and pumping rates in the inverse modeling effort.

[25] The corresponding results for the NYUST aquifer are illustrated in Figures 4a–4g for pumping well locations BH01, BH02, BH03, BH06, BH07, BH08, and BH09, respectively. Similar to the results of the synthetic field experiment, each estimated T field of the NYUST aquifer bears some similar features (not as clear as those in the synthetic aquifer, possibly because of noise) and exhibits the same scenario dependence.

[26] The results from both approaches 1 and 2 as well as the cross correlation between the head observation and the T values in a heterogeneous aquifer [*Wu et al.*, 2005] suggest that unless the heterogeneity of the aquifer is completely known, any estimated T field will always vary with the pumping location.

4.3. Approach 3

[27] The scenario-dependent estimates from approach 2 (i.e., head data from each pumping test yields a heterogeneous T field of a similar general pattern but different details)

Pumping Well No.	Performance Statistics						
	Geometric Mean (m ² s ⁻¹)	Arithmetic Mean $(m^2 s^{-1})$	Variance (m ⁴ s ⁻²)	$L1 (m^2 s^{-1})$	$L2 (m^4 s^{-2})$	COR	
BH01	0.00378	0.00441	6.34E-06	3.38E-03	3.29E-05	0.445	
BH02	0.00282	0.00328	4.22E-06	2.93E-03	3.46E-05	0.454	
BH03	0.00238	0.00288	3.15E-06	2.86E-03	3.46E-05	0.525	
BH06	0.00158	0.00206	2.66E-06	3.03E-03	4.07E-05	0.412	
BH07	0.00191	0.00238	2.67E-06	2.96E-03	3.85E-05	0.448	
BH08	0.00249	0.00321	5.52E-06	2.86E-03	3.29E-05	0.501	
BH09	0.00266	0.00315	4.39E-06	2.83E-03	3.37-05	0.495	

Table 2. Approach 2 Statistics of the Estimated T Field of the Synthetic Aquifer^a

^aData are based on approach 2 with different pumping test locations and the statistics showing its comparison with the true field.

fortify the principle of hydraulic tomography. In other words, observed heads at 10 observation wells induced by pumping at different locations carry different information about the heterogeneity of the aquifer. Therefore, we successively incorporated the seven steady state head data sets from the sequential pumping tests into the inversion until the steady state head data sets of all seven pumping tests were completely utilized. During inclusion of additional tests into the estimation, data from all tests considered are simultaneously used to conduct the joint interpretation, rather than sequentially as in the work of *Zhu and Yeh* [2005]. In addition, the principle of reciprocity (see section 5), which can reduce data sets, is not considered. The estimated *T* fields using sequentially increasing number of tests are illustrated



Figure 4. The estimated *T* fields based on heads observed at 10 wells induced by pumping at a given well (the white square in each figure) at the NYUST site (approach 2).

in Figures 5b–5h for the synthetic aquifer and Figures 6a–6g for the NYUST aquifer.

[28] A comparison of the estimated T field of the svnthetic (Figure 5b) and the NYUST aquifer (Figure 6a), using the data due to pumping at BH01 only with the rest of Figures 5 and 6, shows that sequential incorporation of the steady state head data from different pumping tests into the inverse modeling effort stabilizes the general pattern of the estimated T field. After two or three additional pumping test data sets are used, the improvement of the estimate is very minor. These findings are also supported and quantified in terms of L1, L2, and COR in Table 3. As indicated in the Figures 5h and 6g as well as the statistic metrics in Table 3, the inclusion of all seven tests generally vields the best-estimated field. Again, improvements of the estimated T field are not restricted only to the region close to the wells where heads were observed and used in the estimation, but also cover areas near the boundaries of the aquifer and far away from the wells. These results also disagree with those by *Bohling and Butler* [2010].

5. Validation of Estimated T Fields

[29] One of the ultimate goals of aquifer characterization is to improve our ability to predict responses of the aquifer under any stress. Accurate mapping of hydraulic property distributions in aquifers is one of the necessary steps to accomplish this goal. In the case of the synthetic aquifer where the true hydraulic properties are known precisely, the estimated properties by any approach can be compared with their true values (as done in section 4). Alternatively, two possible approaches can be used to validate the estimates using the same well field for the estimation. One approach is to test the ability of the estimated hydraulic properties for predicting the drawdown observed at wells that are excluded from the inverse modeling efforts. The



Figure 5. (a) The true T field of the synthetic aquifer. The estimated T field based on data sets from a different number of pumping tests (approach 3) in the synthetic aquifer. The white square represents the pumping well location.



Figure 6. The estimated *T* field based on data sets from a different number of sequential pumping tests (approach 3) in the NYUST aquifer. The white square denotes the pumping well location.

Table 3. Approach 3 Statistics of the Estimated T Field of the Synthetic Aquifer^a

Pumping Well No.	Performance Statistics						
	Geometric Mean $(m^2 s^{-1})$	Arithmetic Mean $(m^2 s^{-1})$	Variance $(m^4 s^{-2})$	$(m^2 s^{-1})$	$(m^4 s^{-2})$	COR	
BH01	0.00378	0.00441	6.34E-06	3.38E-03	3.29E-05	0.445	
BH01, 02	0.00267	0.00360	1.22E-05	3.17E-03	3.61E-05	0.398	
BH01, 02, 03	0.00255	0.00425	2.75E-05	3.39E-03	4.56E-05	0.341	
BH01, 02, 03, 06	0.00204	0.00320	9.95E-06	2.61E-03	3.00E-05	0.556	
BH01, 02, 03, 06, 07	0.00218	0.00366	1.51E-05	2.90E-03	3.79E-05	0.378	
BH01, 02, 03, 06, 07, 08	0.00223	0.00392	1.96E-05	2.55E-03	2.80E-05	0.578	
Total 7 wells	0.00205	0.00346	1.48E-05	2.61E-03	2.99E-05	0.545	

^aData are based on approach 3 using different numbers of sequential pumping tests and the statistics showing its comparison with the true field.

other is to test the ability of the estimate for predicting aquifer responses at all observation wells induced by new stresses at locations which are not considered in the inverse modeling exercises, for example, *Illman et al.* [2007, 2009], *Liu et al.* [2007], and *Xiang et al.* [2009]. These two means are qualified as an independent validation of the estimated parameters.

[30] According to Wu et al. [2005], the drawdown at an observation well during the late time period of a pumping test is correlated with heterogeneity everywhere within the cone of depression and, in particular, is highly correlated with the T values near the observation and the pumping wells. This implies that the drawdown measured at an observation well during a pumping test in a heterogeneous aquifer will likely be different from the drawdown recorded at another observation well at a different location at the same distance from the pumping well. The claim of the independent validation of the first method is thus substantiated. Likewise, the cross-correlationship suggests that the drawdowns observed at the same observation well induced by pumping at different locations in a heterogeneous aquifer will be different. This corroborates the independence of the second validation method.

[31] Notice that the cross-correlation pattern of the cross-hole pumping test based on the groundwater flow model is the same regardless of which one of the pair of wells is the pumping well or the observation well. Specifically, for a given pair of wells, say, well A and well B, the drawdown-time curve observed at well A because of pumping at well B will be the same as that observed at well B when well A is pumped in any heterogeneous aquifer. This is the so-called principle of reciprocity [*Bruggeman*, 1972]. As a result, drawdown/pumping data sets of any pair of observation and pumping well that have been used in the inversion theoretically are redundant (i.e., perfectly correlated), and thus cannot be used to validate the estimated *T* field unless noise is considered.

[32] In both the synthetic and NYUST aquifers, there were 11 pumping events, and in each event, drawdowns at 10 wells were sampled, excluding the pumping well. Therefore, there are 110 steady drawdown data in total from the sequential pumping tests. Approaches 1 and 2 use drawdown data from 10 observation wells induced by one of the 11 pumping tests to the estimated T field. As a result, the 100 heads resulting from the remaining 10 pumping tests are nonredundant and are qualified and used for validating the estimate based on a given pumping test.

[33] On the other hand, for approach 3, only half of the 110 heads are nonredundant according to the principle of reciprocity. Figure 7 shows a scatterplot of the 55 normalized steady state observed drawdowns against the other 55 normalized drawdowns of the NYUST field experiment. These drawdowns are normalized with their corresponding pumping rates. For the synthetic aquifer, the scatterplot of the 55 pairs of drawdowns exactly forms a 45° line, indicating of the validity of the principle of reciprocity. The scatter of the normalized drawdown data of the pumping tests at the NYUST field site could be attributed to noise in the measurements of the heads or flow rates. On the other hand, it merely disproves the principle of reciprocity in the real-world scenarios. The failure of the principle in this situation may be attributed to possible physical incorrectness



Figure 7. Scatterplot of drawdowns from 55 pairs of pumping/observation wells during the 11 sequential pumping tests at the NYUST field site.

of the governing flow equation (e.g., turbulent flow, see *Delay et al.* [2011]). Nevertheless, proof of this principle for the real-world aquifers is beyond the scope of this paper.

[34] For the HT analysis of the study, results of seven out of the 11 pumping tests were used for the joint estimations, and results of pumping tests at boreholes BH04, BH05, BH10, and BH11, were reserved for validation purposes. Therefore, we have 16 pairs of an observation well and a pumping well that were not used in the joint inversion. Since drawdowns at the pumping well were not recorded during the pumping test, we have only 12 pairs of drawdown data left for validation. Only half of the 12 drawdowns (i.e., six drawdowns) are nonredundant if the principle of reciprocity is deemed to be valid.

5.1. Validation of Estimates of Approach 1

[35] Figures 8a-8h show scatterplots of observed versus predicted steady drawdowns at 10 observation wells of the 10 validation pumping tests in the synthetic aquifer. The predicted drawdowns in each figure represent the simulated drawdowns using the estimated effective T derived from the heads from all nodes of the synthetic aquifer associated with a specified pumping test (Table 1). The predicted drawdowns in Figure 8h denote those simulated with the effective T estimated from simultaneous inclusion of all data sets from the seven pumping tests. Because of using heads from the seven pumping test for calibration in this case, only 40 drawdowns are qualified as nonredundant data for validation and are shown in Figure 8h. Performance statistics (L1, L2, and COR) are also given in each figure. The scatterplots and the statistics indicate that the prediction based on the effective T derived from simultaneous inclusion of the seven pumping tests is the best. The predicted drawdowns using the effective T-based individual pumping test data are generally unsatisfactory: they are either much smaller or greater than those observed.

[36] Results of the same analysis for the NYUST field experiments are shown in Figures 9a-9h. Overall, the



Figure 8. Scatterplots of observed drawdowns versus predicted drawdowns of the 10 validation pumping tests. The predicted drawdowns in each plot are based on the homogeneous T field, which were estimated using heads at all nodes in the synthetic aquifer during pumping at a given pumping well (the well number is given at the top) (approach 1). The predicted drawdowns in Figure 8h are derived from the estimated T based on all heads during the seven pumping tests.

predictions are also unsatisfactory as are those of the synthetic aquifer. The scatter in these figures is more significant than that in figures for the synthetic aquifer, possibly because of noise or mathematical model errors, or differences in the scale of the observed and simulated drawdowns. Similar to the synthetic aquifer case, the prediction based on the effective T obtained using heads of all pumping tests is the best and unbiased. These best unbiased predictions corroborate the cross-correlation analysis by Wu et al. [2005]. The inclusion of more pumping tests brings forth more samples of heterogeneity and, thus, a more representative effective T has been obtained. The unsatisfactory predicted results also illustrate the limitations of traditional Theis and Jacob analyses.

5.2. Validation of Estimates of Approach 2

[37] Comparisons of observed versus predicted drawdowns using heterogeneous T estimates based on each one of the seven single pumping tests in the synthetic aquifer are shown in Figures 10a–10g. Similar results for the NYUST field experiment are illustrated in Figures 11a– 11g. Overall, the heterogeneous approach yields T fields that lead to improved predictions of drawdowns for the 10 validation pumping tests compared to those based on the



Figure 9. Scatterplots of observed drawdowns versus predicted drawdowns of the 10 validation pumping tests. The predicted drawdowns in each plot are based on the homogeneous T field, which were estimated using heads at all nodes in the NYUST field aquifer during pumping at a given pumping well (the well number is given at the top) (approach 1). The predicted drawdowns in Figure 9h are derived from the estimated T based on all heads during the seven pumping tests.

equivalent homogeneous approach. The improvement in the predicted drawdown is much more obvious in the cases associated with the synthetic aquifer than in the cases associated with the NYUST aquifer. Comparing Figures 8 and 10 of the synthetic aquifer, it is clear that the systematic bias in predicted drawdowns associated with the scenario dependence of the T estimate has been reduced significantly using the heterogeneous conceptual model. Likewise, the scatter, which is attributed to heterogeneity omitted by the conceptual model, also becomes much smaller. These improvements are, however, not very clear in the cases with the NYUST aquifer, possibly because of errors in the measurements in the heads or pumping rates, as well as mathematical model errors and other issues such as 3-D effects.

Nevertheless, the performance statistics of the cases show minor improvements over the results based on the homogeneous conceptual model. That is, even under the situation where data are noisy and infested with errors, a highly parameterized conceptual model is not a drawback to an inverse modeling effort.

5.3. Validation of Estimates of Approach 3

[38] Figures 12a–12g show the scatterplots of the 1764 pairs of true and simulated drawdowns at every node of the synthetic aquifer (21×21), excluding the boundary nodes during the four validation pumping tests. Each figure represents the comparison of the true and the predicted heads using the estimated heterogeneous *T* field based on a given







BH03

Figure 10. Scatterplots of observed drawdowns versus predicted drawdowns of the 10 validation pumping tests. The predicted drawdowns in each plot are based on the heterogeneous T field, which were estimated using heads at 10 observation wells in the synthetic aquifer during pumping at a given pumping well (the well number is given in each figure) (approach 2).

number of sequential pumping tests. The large circles in these figures represent the six nonredundant heads, based on the principle of reciprocity. As indicated in these figures, the predicted drawdown at every node of the aquifer (not just the nodes in the area surrounded by the wells) improves, as data of additional sequential pumping tests are included in the estimation. However, after incorporating the data of four sequential pumping tests, data of additional tests do not improve the bias of the predicted head significantly although scatter is reduced. Notice that the improvement of the drawdown recorded at the three wells during the four tests (the six nonredundant drawdowns) is noticeably better than the drawdowns at many other locations. This can be attributed to their locality close to all the observation wells. Again, we emphasize that improvements are not limited to locations near observation wells and in between observation wells but also far away from the wells and near the boundary. That is, improvement (in terms of scatter) progressively extends to the entire aquifer, as more sequential test data are included. This result again disagrees with that by *Bohling and Butler* [2010].

[39] Comparison between the observed and predicted drawdowns at the NYUST field aquifer during the four validation pumping tests are illustrated in Figures 13a–13g. All 40 pairs of observed and simulated drawdowns (ignoring the principle of reciprocity) are used in the comparison because of the presence of noise in the observed data. The scatterplots and the performance statistics again show that drawdown prediction progressively becomes better as data from more pumping tests are included. More importantly, the *T* estimates from the joint inversion of the sequential pumping tests in the NYUST aquifer yield the best predictions of





Figure 11. Scatterplots of observed drawdowns versus predicted drawdowns of the 10 validation pumping tests. The predicted drawdowns in each plot are based on the heterogeneous T field estimated using heads at 10 observation wells in the NYUST field aquifer during pumping at a given pumping well (the well number is given in each figure) (approach 2).

drawdowns induced by the validation pumping tests in comparison with those based on either approach 1 or 2.

6. Discussion and Conclusions

[40] Results of our analysis show that the equivalent homogeneous approach (approach 1) yields scenario-dependent effective T estimates. The estimate varies with the pumping well location although the estimates are close to the geometric mean of T values of the synthetic aquifer. In comparison with the observed heads of the validation tests, the drawdowns predicted using the estimated T are dispersed and biased in both the synthetic and NYUST aquifers. The bias in the predicted drawdown also depends on the pumping well location employed in the estimation of the T field. This raises a salient question about the utility of the estimates from Theis or Jacob's analysis for a prediction of drawdown, in turn flow, which was defended by *Butler* [2008].

[41] Likewise, approach 2, a heterogeneous approach using data from a single pumping test, produces T spatial distributions capturing the general pattern of the true field of the synthetic aquifer. Although the general pattern of each estimated T field is similar, the estimated T distributions for both the synthetic and NYUST aquifers again vary with the pumping location. Comparisons of the predicted drawdowns of the validation tests show that the estimated T fields, based on approach 2 for both aquifers, consistently yield much better results than those based on approach 1.





Figure 12. Scatterplots of observed drawdowns versus predicted drawdowns of the four validation pumping tests in the synthetic aquifer. The predicted drawdowns in each plot are calculated from the heterogeneous T field, which were estimated using observed heads of a given number of sequential pumping tests (the number of pumping tests used is specified in each figure) (approach 3). The black open circles denote the nonredundant head data if the principle of reciprocity is considered.

That is, even though the estimation problem may be ill-defined, a highly parameterized conceptual model results in more information regarding the pattern of heterogeneity, which leads to a better prediction of the drawdown field under different stresses.

[42] Lastly, approach 3 (HT), which also adopts a highly parameterized conceptual model but jointly interprets data sets from the sequential pumping test, progressively removes the scenario dependence of the T estimates in both the synthetic and NYUST aquifers as more data sets from the sequential pumping tests are included. In addition, the quality of the estimate is continuously improved. The improvement in the T estimate is not limited only to areas near the wells, but the entire aquifer. This improvement is also manifested in the

predicted drawdown of the validation pumping tests, confirming advantages of joint interpretation of the sequential pumping tests (HT) and a highly parameterized conceptual model.

[43] The results of this study are consistent with the cross-correlation analysis of the steady drawdown at a point in the aquifer due to pumping and T values everywhere in the aquifer, and refute the inflated warning of limitations of HT by *Bohling and Butler* [2010], which largely stems from their inappropriate interpretation methodology.

[44] *Bohling et al.* [2002] advocated that drawdown reaches a steady state shape (a condition in which hydraulic gradient between points does not change with time) much earlier than a steady state condition before being affected by boundaries. In addition, drawdown under steady state shape





Figure 13. Scatterplots of observed drawdowns versus predicted drawdowns of the four validation pumping tests in the NYUST field site (approach 3). The predicted drawdowns in each plot are based on the heterogeneous T field, which were estimated using observed heads of a given number of sequential pumping tests (the number of pumping tests used is given in each figure).

conditions can be analyzed as a steady state flow. However, the fact is that the steady shape condition exists only in the sense of approximation, and that early time head data are affected only by heterogeneity close to wells, not as a broad area as the steady state head. Moreover, boundary conditions do not play an import role for steady flow analysis if the pumping rate and many heads are known [*Nelson*, 1960].

[45] Likewise, the cross-correlation analysis and our results expose a crucial weakness of the pilot point approach used by *Bohling and Butler* [2010] and others. Briefly, the pilot point approach employs the maximum likelihood approach to convert some observed head information to obtain hydraulic parameter estimates at some selected pilot points. The number of pilot points is generally much smaller than the number of grid points where parameters are to be filled. The hydraulic parameters at the pilot points are then used to interpolate the parameter at the other grid points based on a given spatial statistic covariance function (or variogram) of the property (note *Bohling and Bulter* [2010] used a spline interpolation function). As a consequence, new head information from the sequential pumping test improves only the estimates at the pilot points and offers little improvement on the estimates at other locations. This limitation is attributed to the fact that the spatial covariance function (or the spline function), which remains the same, does not reflect effects of change in the flow field due to change of the pumping location as indicated by the cross-correlation analysis. Thus, the information about heterogeneity contained in the head data resulting from the sequential pumping tests is not fully utilized. [46] Our results also suggest that interpretation tools based on a ray-tracing technique used in geophysics [e.g., *Vasco et al.*, 2000; *Brauchler et al.*, 2010, 2011], or based on a 2-D radial-vertical flow model [*Bohling et al.*, 2002; *Bohling*, 2009; *Bohling and Butler*, 2010], do not fully consider the cross correlation between the head and *T* heterogeneity over the entire cone of depression. While these tools are useful, again, they do not maximize the information about the heterogeneity contained in the head measurement.

[47] In conclusion, results of our analysis of sequential pumping tests in both the synthetic and NYUST aquifers promote the use of a highly parameterized conceptual model for inverse modeling with a conjunctive use of the hydraulic tomography interpretation. There is no doubt that hydraulic tomography combined with an appropriate analysis and interpretation can maximize the utility of a well field to improve the characterization of the aquifer. In contrast to arguments from *Butler* [2008], our results support the call for change on the way data are collected and analyzed [*Yeh and Lee*, 2007].

[48] **Acknowledgments.** The authors would like to acknowledge the National Science Council of Taiwan for support under grant NSC 99-2221-E-224-028 and the Water Resources Agency, Ministry of Economic Affairs, of Taiwan for grant MOEAWRA-0990029. Site pumping tests, which were conducted by Huang-Jia Huang and support from the National Yunlin University of Science and Technology, are greatly appreciated. Support for T.-C. J. Yeh and W. Lu from the China Education Department and Jilin University, Jilin, China is also acknowledged. T.-C. J. Yeh also acknowledges the support from US NSF EAR-1014594.

References

- Berg, S. J., and W. A. Illman (2011), Capturing aquifer heterogeneity: Comparison of approaches through controlled sandbox experiments, *Water Resour. Res.*, 47, W09514, doi:10.1029/2011WR010429.
- Bohling, G. C. (2009), Sensitivity and resolution of tomographic pumping tests in an alluvial aquifer, *Water Resour. Res.*, 45, W02420, doi:10.1029/ 2008WR007249.
- Bohling, G. C., and J. J. Butler Jr. (2010), Inherent limitations of hydraulic tomography, *Ground Water*, 48(6), 809–824.
- 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:10.1029/2001WR 001176.
- 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.
- Brauchler, R., R. Hu, T. Vogt, D. Halbouni, T. Heinrichs, T. Ptak, and M. Sauter (2010), Cross-well interference slug tests: An efficient tool for high resolution characterization of hydraulic heterogeneity, *J. Hydrol.*, 384, 33–45, doi:10.1016/j.jhydrol.2010.01.004.
- Brauchler, R., R. Hu, P. Dietrich, and M. Sauter (2011), A field assessment of high resolution aquifer characterization based on hydraulic travel time and hydraulic attenuation tomography, *Water Resour. Res.*, 47, W03503, doi:10.1029/2010WR009635.
- Bruggeman, G. A. (1972), The reciprocity principle in flow through heterogeneous porous media, in *Development in Soil Science, Fundamental of Transport Phenomena in Porous Media*, vol. 71, pp. 136–149, IAHR, Elsevier, Amsterdam.
- Butler, J. J., Jr. (2008), Pumping tests for aquifer evaluation time for a change? Technical Commentary, *Groundwater*, 47(5), doi:10.1111/ j.1745-6584.2008.00488.x.
- Cardiff, M., W. Barrash, P. K. Kitanidis, B. Malama, A. Revil, S. Straface, and E. Rizzo (2009), A potential-based inversion of unconfined steady-

state hydraulic tomography, *Ground Water*, 47(2), 259–270, doi:10.1111/j.1745-6584.2008.00541.x.

- Castagna, M., and A. Bellin (2009), A Bayesian approach for inversion of hydraulic tomographic data, *Water Resour. Res.*, 45, W04410, doi:10.1029/2008WR007078.
- Cooper, H. H., Jr., and C. E. Jacob (1946), A generalized graphical method for evaluating formation constants and summarizing well-field history, *Eos Trans. AGU*, 27(4), 526–534.
- Delay, F., P. Ackerer, and A. Guadagnini (2011), Theoretical analysis and field evidence of reciprocity gaps during interference pumping tests, *Adv. Water Resour.*, 34(5), 592–606.
- 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:10.1088/02665611/11/2/005.
- Gutjahr, A. L. (1989), Fast Fourier transforms for random field generation, *Project Rep. for Los Alamos Grant, Contract 4-R58-2690R*, Dep. of Math., N. M. Tech., Socorro.
- Hanna, S., and T.-C. J. Yeh (1998), Estimation of co-conditional moments of transmissivity, hydraulic drawdown, and velocity fields, *Adv. Water Resour.*, 22(1), 87–95, doi:10.1016/S0309-1708(97)00033-X.
- Hao, Y., T.-C. J. Yeh, J. Xiang, W. A. Illman, K. Kenichi, K. C. Hsu, and C. H. Lee (2008), Hydraulic tomography for detecting fracture zone, *Ground Water*, 46(2), 183–192, doi:10.1111/j.1745-6584.2007.00388.x.
- 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., X. Liu, S. Takeuchi, T.-C. J. Yeh, K. Ando, and H. Saegusa (2009), Hydraulic tomography in fractured granite: Mizunami underground research site, Japan, *Water Resour. Res.*, 45, W01406, doi:10.1029/ 2007WR006715.
- Illman, W. A., J. Zhu, A. J. Craig, and D. Yin (2010), Comparison of aquifer characterization approaches through steady state groundwater model validation: A controlled laboratory sandbox study, *Water Resour. Res.*, 46, W04502, doi:10.1029/2009WR007745.
- Li, W., A. Englert, O. A. Cirpka, J. Vanderborght, and H. Vereechen (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), Threedimensional geostatistical inversion of flowmeter and pumping test data, *Ground Water*, 46(2), 193–201.
- Liu, S., T.-C. J. Yeh, and R. Gardiner (2002), Effectiveness of hydraulic tomography: Sandbox experiment, *Water Resour. Res.*, 38(4), 1034, doi:10.1029/2001WR000338.
- Liu, X., and P. K. Kitanidis (2011), Large-scale inverse modeling with an application in hydraulic tomography, *Water Resour. Res.*, 47, W02501, doi:10.1029/2010WR009144.
- 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.
- Meier, P. M., J. Carrera, and X. Sanchez-Vila (1998), An evaluation of Jacob's method for interpretation of pumping tests in heterogeneous formations, *Water Resour. Res.*, 34(5), 1011–1025, doi:10.1029/98WR00008.
- Nelson, R. W. (1960), In-place measurement of permeability in heterogeneous media, 1, Theory of a proposed method, J. Geophys. Res., 65(6), 1753–1758, doi:10.1029/JZ065i006p01753.
- Sanchez-Vila, X., P. M. Meier, and J. Carrera (1999), Pumping test 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, doi:10.1029/1999WR900007.
- 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.
- Theis, C. V. (1935), The relation between lowering the piezometric surface and the rate and duration of discharge of a well using ground water storage, *Eos Trans. AGU*, *16*, 519–524.
- 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.
- Wen, J. C., C. M. Wu, T.-C. J. Yeh, and C. M. Tseng (2010), Estimation of effective aquifer hydraulic properties from an aquifer test with muti-well

observations (Taiwan), *Hydrogeol. J.*, 18, 1143–1155, doi:10.1007/s10040-010-0577-1.

- 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.
- 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.
- Yeh, T.-C. J., and C. H. Lee (2007), Time to change the way we collect and analyze data for aquifer characterization, *Ground Water*, 45(2), 116–118, doi:10.1111/j.17456584.2006.00292.x.
- 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., D. Mao, L. Wan, K.-C. Hsu, J.-C. Wen, and C.-H. Lee (2011), Uniqueness, scale, and resolution issues in groundwater model parameter identification, Tech. Rep., Dept. of Hydrology and Water Resources, Univ. Arizona.

- Yin, D., and W. A. Illman (2009), Hydraulic tomography using temporal moments of drawdown recovery data: A laboratory sandbox study, *Water Resour. Res.*, 45, W01502, doi:10.1029/2007WR006623.
- 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.

K.-C. Chang, H.-L. Juan, J.-H. Lee, and C.-M. Tseng, Water Resource Agency, Ministry of Economic Affairs, 41-3, Hsin-Yi Road, Section 3, Taipei 106, Taiwan.

S.-Y. Huang, Graduate School of Engineering Science and Technology, National Yunlin University of Science and Technology, 123, Section 3, University Road, Douliou, Yunlin 64002, Taiwan.

W. Lu, Department of Groundwater Science and Engineering, College of Environment and Resources, Jilin University, 6 Ximinzhu St., Changchun, Jilin, 130026, China.

J.-C. Wen, Department and Graduate School of Safety Health and Environment Engineering, Research Center for Soil and Water Resources and Natural Disaster Prevention, National Yunlin University of Science and Technology, 123, Section 3, University Road, Douliou, Yunlin 64002, Taiwan. (wenjc@yuntech.edu.tw)

T.-C. J. Yeh, Department of Hydrology and Water Resources, University of Arizona, John Harshbarger Building, 1133 E. North Campus Drive, Tucson, AZ 85721, USA.