Journal of Hydrology 569 (2019) 117–134

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Reproducibility of hydraulic tomography estimates and their predictions: A two-year case study in Taiwan

Jyun-Lin Chen[a], Jet-Chau Wen[b,c,*], Tian-Chyi Jim Yeh[d], Kun-Yi Andrew Lin[e], Yu-Li Wang[d], Shao-Yang Huang[c], Yang Ma[f], Chia-Yii Yu[g], Cheng-Haw Lee[h]

[a] Doctoral Program, Graduate School of Engineering Science and Technology, National Yunlin University of Science and Technology, Douliou, Yunlin, Taiwan
[b] Department of Safety, Health and Environmental Engineering, National Yunlin University of Science and Technology, Douliou, Yunlin, Taiwan
[c] Research Center for Soil and Water Resources and Natural Disaster Prevention, National Yunlin University of Science and Technology, Douliou, Yunlin, Taiwan
[d] Department of Hydrology and Atmospheric Sciences, University of Arizona, Tucson, AZ, USA
[e] Department of Environmental Engineering, National Chung Hsing University, Taichung, Taiwan
[f] Regional Director Water and Urban Development, China AECOM, Shatin, Hong Kong, China
[g] Safety, Health, and Environmental Center, Group Administration, Formosa Plastics Group, Taipei, Taiwan
[h] Department of Resources Engineering, National Cheng Kung University, Tainan, Taiwan

ARTICLE INFO

This manuscript was handled by P. Kitanidis, Editor-in-Chief, with the assistance of Chunhui Lu, Associate Editor

Keywords:
Reproducibility
Sequential pumping test
Hydraulic tomography (HT)

ABSTRACT

Over the past decades, a new aquifer test technology (sequential pumping tests or hydraulic tomography, HT) has been developed and successfully applied to many field sites to delineate the spatial distributions of hydraulic properties (e.g., transmissivity (T) and storage coefficient (S)). Yet, the reproducibility of its estimated T and S fields and the predictive capabilities of the estimates for different flow scenarios at different time periods remain unexplored. That is to say, if the estimated fields based on sequential pumping tests conducted during different years are the same since the geologic formation and processes may have undergone changes. In order to answer this important question, this study first compares the drawdown-time behaviors from the sequential pumping tests (SPTs) conducted in 2010 with those conducted in 2012 at a field site and then finds they are similar but different in detail. It then uses these data to estimate the T and S fields and checks the reproducibility of the estimates. The estimated heterogeneity patterns are found to be generally reproducible in spite of uncertainties. In addition, the estimates from each year are capable of predicting the observed drawdowns, induced by independent pumping tests during the corresponding year (i.e., self-validation). Moreover, the estimated fields are cross-validated. That is, this study uses the estimates obtained from the 2010 pumping tests to predict the observed drawdowns of the independent pumping tests conducted in 2012. Likewise, it uses the estimates from 2012 pumping tests to forecast the drawdowns of the independent pumping tests of 2010. The results of both self-validation and cross-validation indicate that the estimated T and S fields based on the test in one year can be used to predict bulk flow behavior in the other year. Differences in detailed behaviors may be attributed to changes in the processes, omitted in the depth-averaged flow model.

1. Introduction

Aquifers are inherently heterogeneous at a multiplicity of scales. Such heterogeneity controls hydrologic processes. Comprehensive characterization of aquifer hydraulic properties, thus, is critical for groundwater resources management, and groundwater contamination prevention and remediation. Yet, traditional aquifer test analyses for estimating aquifer properties, like the *Theis* solution (Theis, 1935) and the *Cooper-Jacob* approximation (Cooper and Jacob (1946)), adopt the aquifer homogeneity assumption.

Based on numerical experiments, Meier et al. (1998) and Sánchez-Vila et al. (1999) advocated that the homogeneous hydraulic properties, estimated by the *Cooper-Jacob* approach, represent the averaged properties over the cone of depression in the aquifer. They claimed that the estimates are independent of the location of the observation well. To the contrary, Wu et al. (2005) demonstrated that the averaged properties are heavily weighted towards the heterogeneity in areas close to the pumping and to the observation wells and the dominant or

* Corresponding author at: Department of Safety, Health and Environmental Engineering, National Yunlin University of Science and Technology, Douliou, Yunlin, Taiwan.

E-mail address: wenjc@yuntech.edu.tw (J.-C. Wen).

https://doi.org/10.1016/j.jhydrol.2018.11.064
Received 31 July 2018; Received in revised form 25 October 2018; Accepted 19 November 2018
Available online 01 December 2018

large-scale heterogeneity in other parts of the cone of depression. They further pointed out that the governing flow equation upon which *Cooper-Jacob* solution is based is an ensemble mean equation. The equation describes the averaged behavior of an aquifer with many possible heterogeneous patterns (Yeh et al., 2015a,b). To properly apply the ensemble-mean solution to a field problem, where only one heterogeneity pattern exists, they suggested that *Theis's* or *Cooper-Jacob's* solutions ought to be fitted to drawdown-time data from multiple observation wells over the cone of the depression in order to obtain representative averaged properties. Collectively, field results from Straface et al. (2007) and Wen et al. (2010) corroborated the conclusion by Wu et al. (2005), which stated that applying the *Theis's* or the *Cooper-Jacob's* solutions to one observed drawdown-time curve in a cross-hole pumping test in a heterogeneous aquifer is tantamount to comparing apples and oranges.

Thereafter, Huang et al. (2011) used well hydrographs from 11 observation wells during each of eleven pumping tests at a field site and estimated the transmissivity (*T*) values of an equivalent homogeneous model and a highly parameterized heterogeneous model (Yeh et al., 2015a,b) for each test. They found that these estimated equivalent homogeneous *T* and distributed *T* values vary according to the location of the pumping well. As such, Huang et al. (2011) concluded that using a limited number of pumping tests and observation wells, the estimated hydraulic properties for either an equivalent homogeneous or highly heterogeneous conceptual model likely are scenario dependent and not intrinsic characteristics of an aquifer. They then demonstrated that a joint interpretation of multiple sequential pumping tests (Hydraulic Tomography (HT), survey, and analysis) can minimize the scenario-dependency of the effective homogeneous parameters or distributed parameter fields.

In fact, a tomographic survey is a human instinct. That is, human always views an object from different angles and perspectives to obtain non-redundant information and to synthesize these pieces of information to gain a complete description of the object. Exploiting this human nature, medical scientists developed computed axial tomography (CAT scan) technology and geophysicists created geophysical tomography survey techniques (Yeh et al., 2008). Similarly, many hydrologists had proposed the concept of HT, but Yeh and Liu (2000) are the first who developed a fully 3-D Steady-State Hydraulic Tomography (SSHT) technology and demonstrated its robustness for mapping aquifer heterogeneity. Zhu and Yeh (2005) then expanded the 3-D SSHT to the Transient Hydraulic Tomography (THT), and Xiang et al. (2009) modified the Sequential Successive Linear Estimator (SSLE) to Simultaneous Successive Linear Estimator (SimSLE), which simultaneously includes all the pumping test data sets for THT analysis.

Over the past decades, HT survey with an appropriate inverse model (such as SLE, QLGA, Quasi-Linear Geostatistical Model, Kitanidis, 1995) can identify the patterns of the heterogeneous hydraulic conductivity (*K*) and specific storage (S_s) field in aquifers in a cost-effective manner. These works include (Vesselinov et al., 2001a,b; Liu et al., 2007; Illman et al., 2009; Berg and Illman, 2011a,b; Zhao et al., 2015) in particular.

In field-scale (e.g., tens to hundreds of meters) problems, HT has become a popular method for characterizing the spatial heterogeneity (Straface et al., 2007; Li et al., 2007; Huang et al., 2011; Berg and Illman, 2013; Cardiff et al., 2013; Hochstetler et al., 2016; Sanchez-León et al., 2016; Zhao and Illman, 2017, 2018). More recently, Zha et al. (2014, 2015, 2016, 2017) utilized pumping induced drawdowns and fluxes to map the locations of kilometer-scale faults and fractures as well as their hydraulic properties at a granite site. Likewise, SimSLE has been successfully applied to delineate small-scale fractures (Sharmeen et al., 2012).

Moreover, these studies showed that the hydraulic property fields estimated by HT could lead to a better prediction of flow (Hao et al., 2008; Illman et al., 2008, 2009) and solute transport (Illman et al., 2010, 2012; Ni et al., 2009; Zha et al., 2016) processes than the conventional characterization approaches. Mao et al. (2011) challenged the traditional approaches for analyzing the results of pumping tests in unconfined aquifers, and Mao et al. (2013) extended the HT analysis to variably saturated unconfined aquifers. Meanwhile, Sun et al. (2013) investigated the effects of boundary uncertainty on the HT interpretation and suggested a temporal sampling strategy of the observed head for the HT analysis. Recently, HT was expanded by Zha et al. (2014) to include flux measurements in the characterization of fractured geologic formations. Tso et al. (2016) investigated the usefulness of flux measurements and prior information about the geology in HT analysis. Using the Successive Linear Estimator (SLE), Zha et al. (2017) developed a new HT algorithm that addresses multi-scale heterogeneity and discrete geologic features in geologic formations.

Unequivocally, these HT studies have brought optimism to the development of a cost-effective, high-resolution aquifer characterization technology, in spite of cautionary notes about HT's ability by Bohling and Butler (2010)). Nonetheless, the reproducibility of the estimated parameter fields by HT's for a field site over different years has not been investigated before. That is to say, over a period of time, the regional flow, recharge, geologic heterogeneity, and other unexplored factors may have changed. These changes may lead to different estimates due to the omission of these factors in the model. Consequently, it is logical to ask if the parameter fields estimated, for a given field site, using sequential pumping test data at different years would lead to the same estimates, and if the estimated parameters during a previous campaign can be used to predict the flow field at the site at different times – validation.

Validation is a controversial and philosophical term in groundwater hydrology. For instance, a well-cited paper by Konikow and Bredehoeft (1992), about three decades ago, advocated that groundwater models cannot be validated but tested. They stated "…*Case histories of model applications to the Dakota Aquifer, South Dakota, to bedded salts in New Mexico, and to the upper Coachella Valley, California, illustrate that calibration produces a nonunique solution and that validation, per se, is a futile objective. Although models are definitely valuable tools for analyzing ground-water systems, their predictive accuracy is limited. The terms validation and verification are misleading, and their use in ground-water science should be abandoned in favor of more meaningful model-assessment descriptors.*" We believe this view was built upon knowledge, theories, and models developed more than three decades ago. For example, using pumping test conducted at Wall, South Dakota, they articulated that the *Theis* solution is adequate for predicting the short-term response of the well at Wall (40-hour pumping test); to predict its long-term response, the leakage from the confining layers (i.e., *Hantush* 'modified leaky aquifer solution') must be considered. While they may have recognized effects of large-scale heterogeneity (leakage from the confining layers), they failed to recognize that both *Theis* and *Hantush* (Hantush and Jacob, 1955) solutions are solutions to the ensemble mean equations; they derive the ensemble mean behavior of the aquifer, which is not the same as the observed drawdown-time data at one well. Besides, these equations yield scenario-dependent aquifer properties as discussed previously. Likewise, they used the groundwater modeling study of the upper Coachella Valley, California to illustrate that a calibrated model yielded large uncertainty in predictions. As such, they suggested "model assessment" instead of model validation and verification. Once more, the calibration of the upper Coachella Valley model did not exploit the tomography concept developed recently (i.e., take advantage of non-redundant information). Certainly, we agree with the overall theme of their paper (i.e., model assessment should focus on a complete understanding of the particular hydrologic system or problem of interest). We, however, believe the paper merely presented an argument over semantics.

As articulated in the early part of this section, our understanding of effects of heterogeneity and limitations of traditional models, and advances in inverse modeling using HT have evolved tremendously over the past three decades. A revisit to these "validation", "verification" or "model assessment" issues becomes logical. For this reason, this paper analyzes data from two sequential pumping test experiments conducted in two different years at a field aquifer. Specifically, it first examines the reproducibility of the drawdown data from sequential pumping tests conducted in 2010 and 2012 (i.e., SPT-2010 and SPT-2012, respectively) at the same field site. The reproducibilities of the estimated *T* and storage coefficient (*S*) of an Equivalent Homogeneous Model (EHM) and a Highly Parameterized, Heterogeneous Model (HPHM) are then investigated. In addition to direct comparisons of these *T* and *S* estimates from SPT-2010 and SPT-2012, this study evaluates the ability of these estimates for predicting the flow field due to independent pumping tests conducted in the corresponding year (i.e., self-validation). The independent pumping test means that the drawdown data (in turn, the flow field) from the tests have not been calibrated during the inverse modeling effort for estimating the parameters. The study also cross-validates these estimates. That is, the estimated *T* and *S* fields from SPT-2010 are used to predict the drawdown field induced by the independent pumping tests in SPT-2012 and vice-versa. Finally, the results of these validations are discussed.

2. Field experiments

2.1. Site description and Sequential Pumping Test (SPT)

The SPT experiments were conducted at a field site on the north side of the campus of National Yunlin University of Science and Technology (NYUST) in Taiwan. Eleven fully penetrating wells (screened from 1.5 m below land surface to a depth of 20 m) were installed over an area of 144 m^2; they were distributed as shown in Fig. 1 and named as BH01, BH02, through BH11. Each well was constructed with a schedule 40 high-density polyethylene (HPDE) pipe of 10.16 cm diameter, and the length of the screen section was 18.50 m from 1.50 m below the ground surface and to the depth of 20.00 m and surrounded by a gravel pack. The interval was perforated with 0.05 cm slots. The wells were completed according to the specification of groundwater quality monitoring well (No. 0910091877) by Environmental Protection Administration (EPA), ROC (Taiwan).

The stratigraphic cross sections (A-A' and B-B', blue lines in Fig. 1) are depicted in Fig. 2. Generally speaking, the soil profile shows that the aquifer (16.4 m) is mainly comprised of interbedded gravel, sand, silt, and clay layers. It is overlain by 3.4 m thick silty clay, while the bottom is an impermeable clay layer. The aquifer is generally considered as a shallow, confined aquifer although it may not be fully confined over great lateral distances because of the complex interbedded silty clay layers. In addition, the experiment site is located in a forest area and the depth of roots is unknown. The hydraulic heads in 2010 and 2012 fluctuated in the top clay layer (Fig. 2).

The SPT-2010 took place from August 2010 to February 2011. The pumping rates ranged from 9.42 m^3d^{-1} to 12.96 m^3d^{-1}. The SPT in 2012 (SPT-2012) began in October 2012 and ended in January 2013. The pumping rates varied from 8.03 m^3d^{-1} to 15.17 m^3d^{-1} (Table 1). All the pumping tests during two SPT experiments were conducted at days without any precipitation. During each SPT, a pressure transducer with a data logger (precision of 1 mm) was installed in each of the eleven wells to collect the drawdown data measurement. The submersible pumping system (Grundfos Pumps Corporation) was utilized for each pumping test. Each pumping test lasted 72 h to ensure that a steady flow condition was achieved. For each SPT experiment, eleven pumping tests were conducted. During each pumping test, one of the eleven wells was pumped, and the drawdown-time data were recorded at all wells. After the drawdown had reached steady-state, the pumping was stopped. Once the groundwater level fully recovered, a new pumping test began in another well. As a result, 11 pumping tests yielded 121 observed drawdown-time curves (including the pumping wells) for each SPT.

3. Groundwater model

The analysis presented in this paper assumes that the groundwater flow in two-dimensional, depth-averaged, saturated, confined, heterogeneous, or homogeneous aquifers can be described by the following equations:

$$\nabla\cdot[T(\mathbf{x})\nabla H] + Q(\mathbf{x}_p) = S(\mathbf{x})\frac{\partial H}{\partial t} \tag{1}$$

subject to the boundary and initial conditions:

$$H|_{\Gamma 1} = H_1,\ \ [T(\mathbf{x})\nabla H]\cdot n|_{\Gamma 2} = q,\ \ \text{and} H|_{t=0} = H_0 \tag{2}$$

where, *H* is total or hydraulic head [L], **x** is the spatial coordinate ($\mathbf{x} = \{x, y\}$, [L]), $Q(\mathbf{x}_p)$ is the pumping rate ($L^3/T/L^3$) at the location $\mathbf{x}_p$, $T(\mathbf{x})$ is the transmissivity [L^2/T], and $S(\mathbf{x})$ is the storage coefficient [−]. If

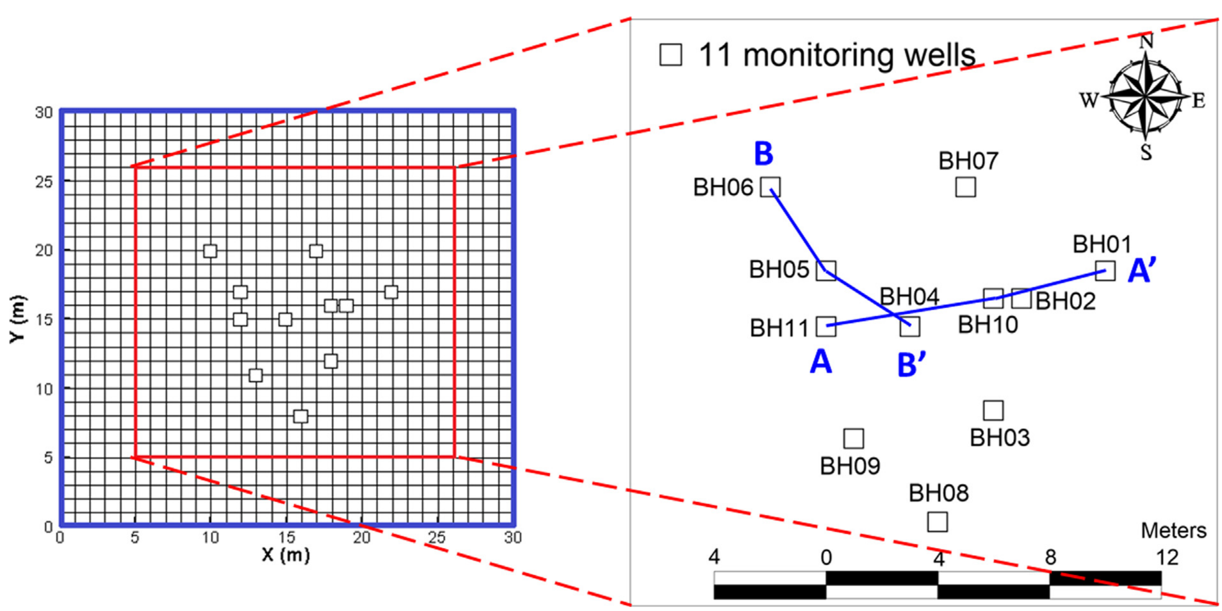

Fig. 1. Well locations and model domain. Each grid is 1 m × 1 m, the total number of grids is 30 × 30, and the boundary conditions are prescribed head. The area indicated by the red square is 21 m × 21 m (441 grids) is the focus of the discussion in the paper. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

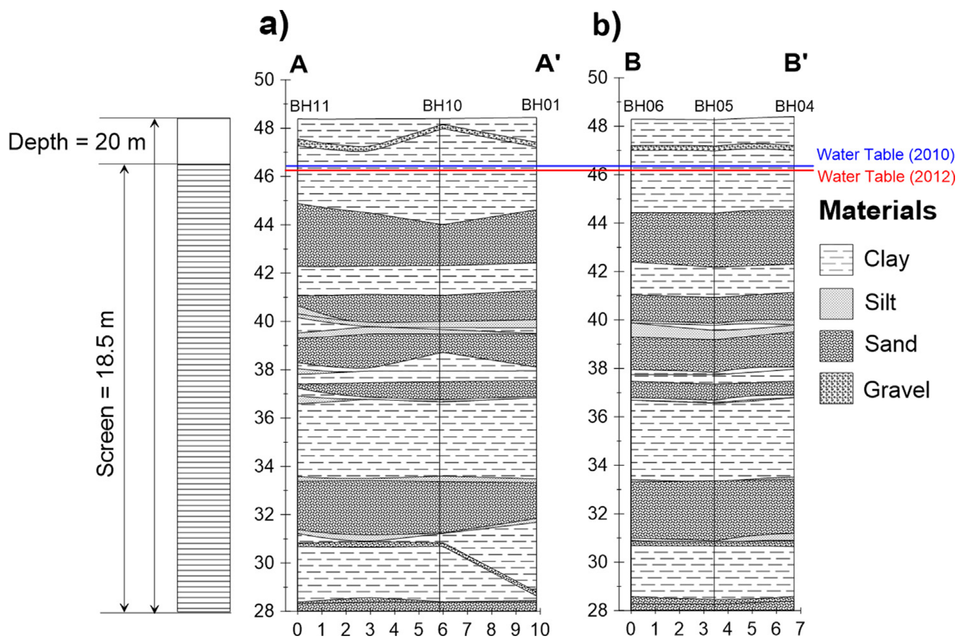

Fig. 2. The stratigraphic profiles and the averaged initial water levels of the aquifer along a) A–A' and b) B–B' lines at the NYUST site (see Fig. 1).

Table 1
Setting conditions of sequential pumping tests between 2010 and 2012.

Cases	SPT-2010	SPT-2012
Date	2010/08-2011/02	2012/10-2013/01
Well No.	Q (m^3/d)	Q (m^3/d)
BH01	11.84	13.61
BH02	12.42	15.17
BH03	11.76	14.30
BH04	12.86	13.64
BH05	12.83	13.29
BH06	12.74	14.10
BH07	12.96	14.58
BH08	12.68	14.45
BH09	12.81	14.05
BH10	9.42	8.03
BH11	10.68	11.90
Average	12.09	13.37

the aquifer is conceptualized as homogenous (EHM), the values of $T(\mathbf{x})$ and $S(\mathbf{x})$ are independent of $\mathbf{x}$. H_1 is the prescribed total head at Dirichlet boundary Γ_1, q is the specific flux at Neumann boundary Γ_2, n is a unit vector normal to the union of Γ_1 and Γ_2, H_0 represents the initial total head.

4. Simultaneous Successive Linear Estimator (SimSLE)

SimSLE (Xiang et al., 2009), a version of the SLE, is used to estimate the parameter fields based on SPT-2010 and SPT-2012 data. The algorithm of the SLE has been discussed by many of Yeh's colleagues (Yeh et al., 1995, 1996; Zhang and Yeh, 1997; Hughson and Yeh, 2000; Yeh and Liu, 2000; Zhu and Yeh, 2005, Zha et al., 2018), and thus only a brief discussion is given below.

SLE adopts a highly parameterized, heterogeneous conceptual model (i.e., every finite element of the model has its own parameters to be estimated). As such, it discretizes the field site into N elements. Each element has a value for each hydraulic parameter (i.e., natural logarithms of K, and S_s, $\ln K$, and $\ln S_s$, respectively). The SLE then considers these parameters as spatial random fields characterized by unconditional means $E[\ln K]$ and $E[\ln S_s]$ (where $E[-]$ denotes the statistical expectation), and the unconditional perturbations $\ln K$-$E[\ln K]$ and $\ln S_s$-$E[\ln S_s]$, respectively. These perturbations represent spatial variability of the parameters, which is characterized by their covariance functions.

SLE estimates the most likely parameter value (i.e., conditional effective parameter value) for each element, given (i.e., conditioned with) the observed drawdown (or head) data from the HT survey. As it is a Bayesian geostatistical method, it starts with some prior knowledge about mean values and covariance functions of the unknown parameter

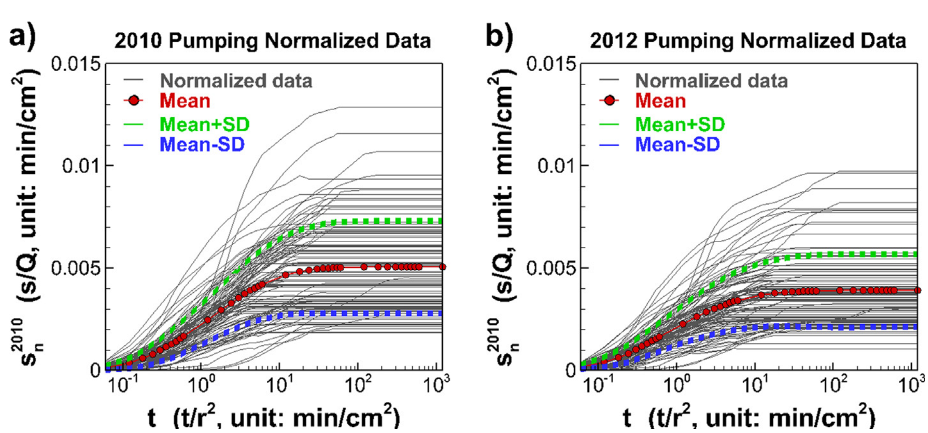

Fig. 3. Normalized drawdown-time plots for a) SPT-2010, and b) SPT-2012. s_n stands for normalized drawdown.

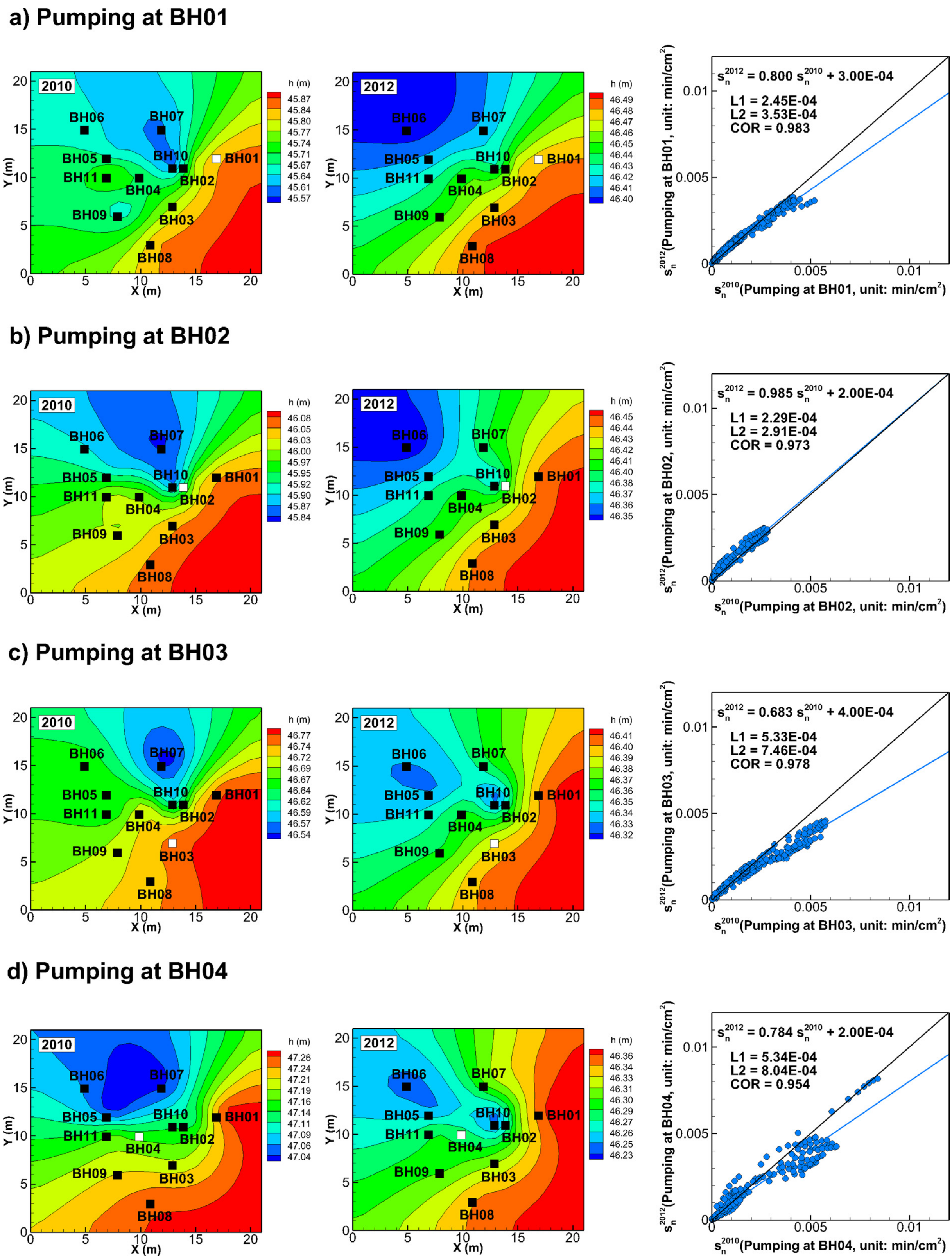

Fig. 4. Contour maps of initial heads (column 1: 2010; column 2: 2012) and the scatter plots of normalized drawdowns of SPT-2010 vs. SPT-2012 (column 3) for the pumping tests at (a) BH01, (b) BH02, (c) BH03, (d) BH04, (e) BH05, (f) BH06, (g) BH07, (h) BH08, (i) BH09, (j) BH10, (k) BH11. s_n denotes the normalized drawdown.

fields (ln$K$ and lnS_s). These prior means are combined as vector $\mathbf{Y}$ (an $n_y \times 1$ parameter vector) in the numerical model. n_y is equal to N (no. of elements of the domain) if ln$K$ is concerned only, and it is doubled ($2N$) if ln$K$ and lnS_s are jointly estimated for transient problems.

Suppose during a pumping test we have collected n_d observed heads in time and space, denoted by the data vector $\mathbf{d}^*$. The estimated parameter vector, given the observation, is $\mathbf{Y}_c$ (subscript c denotes conditional), and is iteratively determined using the following linear

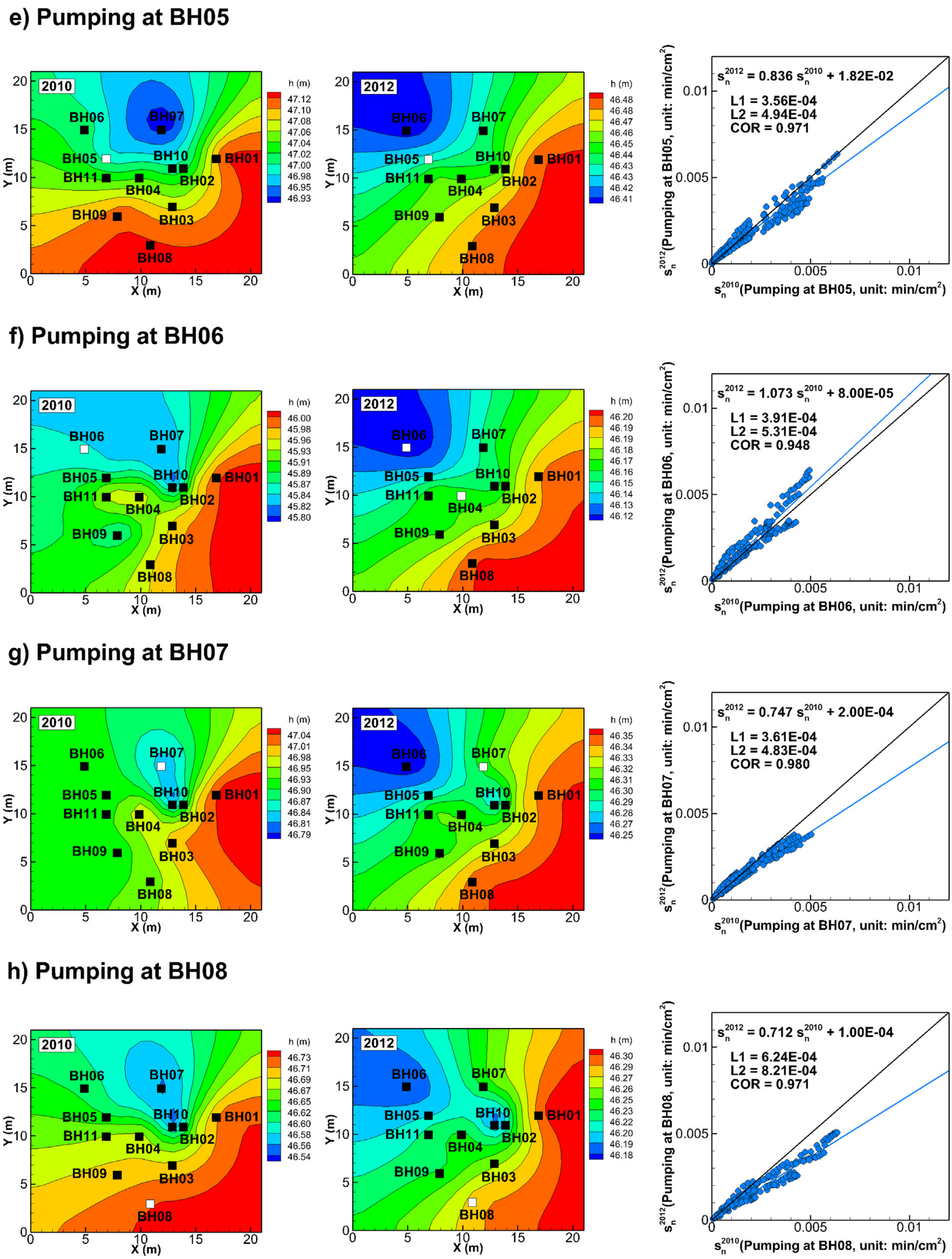

Fig. 4. (*continued*)

estimator:

$$\mathbf{Y}_c^{(r+1)} = \mathbf{Y}_c^{(r)} + \boldsymbol{\omega}^{\mathrm{T}}(\mathbf{d}^* - \mathbf{d}^{(r)}) \tag{3}$$

where r is the iteration index; the vector $\mathbf{d}^{(r)}$ is the simulated heads at the observation locations and times obtained from the forward model (Eqs. (1) and (2)), using the parameters obtained at iteration r. When $r = 0$, $\mathbf{Y}_c = E[\mathbf{Y}]$. The coefficient matrix, $\boldsymbol{\omega}$ $(n_d \times n_y)$, denotes the weights, which assign the contribution of difference between the observed and simulated head at each observation location and time at each iteration to a previously estimated parameter value at each element. The superscript T denotes the transpose. The coefficient matrix $\boldsymbol{\omega}$ is determined by solving the following equations:

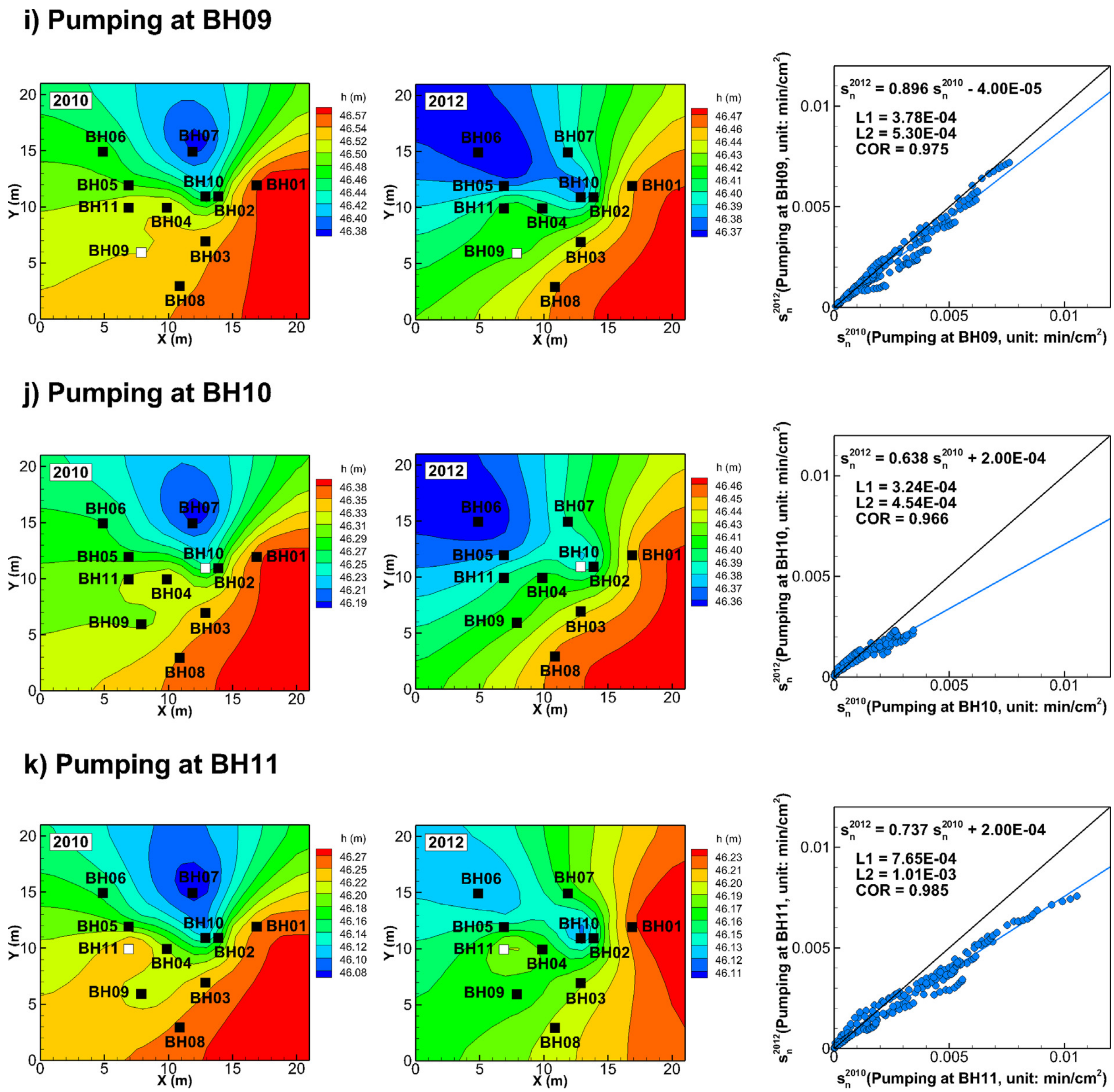

Fig. 4. (*continued*)

Table 2
The input parameters for HT modeling (S is non-dimensional and correlation scale is 10 m).

Cases	Performance Statistics				
	H[a]	T[b]		S[c]	
	I.C[d] and B.C[e] (m)	Mean of ln$T$	Variance of ln$T$	Mean of ln$S$	Variance of ln$S$
2010	46.44	2.633	0.192	−5.307	0.600
2012	46.33				

[a] Pressure head.
[b] Effective transmissivity.
[c] Effective storage coefficient.
[d] Initial condition.
[e] Boundary condition.

Table 3
The estimated hydraulic properties for field data used in self- and cross-validation (S is non-dimensional).

Methods	Years	Estimates	
		ln$T$	ln$S$
Unconditional Direct Average Approach	2010	2.430	−5.075
	2012	2.836	−5.539
Unconditional HT Approach	2010	1.605	−4.491
	2012	1.841	−4.639

$$[\varepsilon_{dd}^{(r)} + \theta^{(r)}\mathrm{diag}(\varepsilon_{dd}^{(r)})]\omega^{(r)} = \varepsilon_{dy}^{(r)} \tag{4}$$

where ε_{dd} is the unconditional covariance of observed heads and ε_{dy} is the cross-covariance between parameter and data where $r = 0$. They become conditional (or residual) covariance and cross-covariance when $r > 0$ (when observed heads are used to condition the estimates). The parameter θ is a dynamic stability multiplier, and diag (ε_{dd}) is a stability matrix, which is a diagonal matrix consisting of the diagonal elements of ε_{dd}. The solution of Eq. (4) requires knowledge of covariance ε_{dd} and cross-covariance ε_{dy}, which are derived from the first-order numerical approximation (Yeh and Liu, 2000):

$$\varepsilon_{dd}^{(r)} = \mathbf{J}_d^{(r)}\varepsilon_{yy}^{(r)}\mathbf{J}_d^{(r)\mathrm{T}}, \quad \varepsilon_{dy}^{(r)} = \mathbf{J}_d^{(r)}\varepsilon_{yy}^{(r)} \tag{5}$$

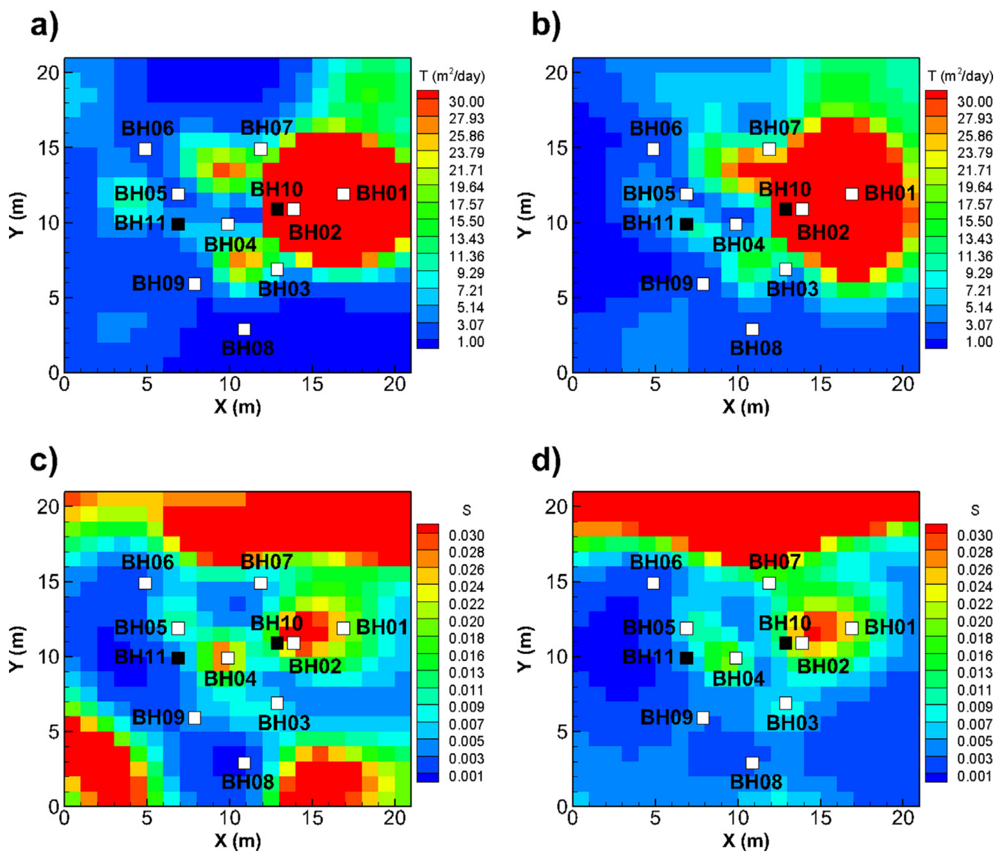

Fig. 5. Conditional HT estimates: The estimated T field (a) 2010, and (b) 2012 from drawdown datasets of 9 pumping tests. The estimated S fields of (c) 2010, and (d) 2012.

where $\mathbf{J}_d$ ($n_d \times n_y$) is the sensitivity matrix of head data with respect to the element-wise parameter, using the parameters estimated at current iteration. The sensitivity is solved by an adjoint approach. At iteration $r = 0$, the ε_{yy} ($n_y \times n_y$) is the unconditional covariance of parameters. For $r \geq 1$ the unconditional covariance becomes the residual or conditional (residual) covariance, which is evaluated according to

$$\varepsilon_{yy}^{(r+1)} = \varepsilon_{yy}^{(r)} - \omega^{\mathrm{T}}\varepsilon_{dy} \tag{6}$$

The iteration between Eqs. (3) and (6) continues until some criteria (Xiang et al., 2009) are met. The final estimate is the conditional effective parameters, and residual variances are the uncertainty associated with the estimate.

Note that SimSLE can also adopt an EHM to estimate the effective parameters for the entire aquifer.

5. Analysis and results

5.1. Evaluation criteria

The mean absolute error (L1 norm), mean square error (L2 norm), linear regression analysis, and standard correlation coefficient (COR) ($0 \leq |r| \leq 1$) were the performance statistics used in the evaluations of results of the following analyses.

5.2. Reproducibility of Drawdown-time curves

In order to investigate the reproducibility of the drawdown-time curves between SPT-2010 and SPT-2012, the drawdown-time curves were normalized according to the following equations:

$$s_n = \bar{s}_{i,j}^{k}(t) = s_{i,j}^{k}(t)/Q_i^k \tag{7}$$

$$t_n = \bar{t}_{i,j} = t_{i,j}/r_{i,j}^2 \tag{8}$$

where the $\bar{s}_{i,j}^{k}(t)$ represents the normalized drawdown, which is the drawdown, $s_{i,j}^{k}(t)$, observed at time t at the observation well j during pumping at well, i, divided by the pumping rate Q_i in the kth year. This is necessary since different pumping rates were used at different pumping wells and years. The term $\bar{t}_{i,j}$ is the normalized time, which is the observation time $t_{i,j}$ divided by the square of the distance between the observation well and pumping well ($r_{i,j}^2$). This normalization is based on *Theis* solution for homogeneous, confined aquifers (Yeh et al., 2015a); it aims to eliminate the distance effect such that effects of heterogeneity or other unknown or unaccounted factors (e.g., recharge or leakage, and others) on the shapes of drawdowns can be illuminated.

The 110 normalized drawdown-time curves from SPT-2010 and SPT-2012 and their means, mean plus one standard deviation, and mean minus one standard deviation are presented in Fig. 3a and b. Overall, the normalized observed drawdowns in SPT-2010 are greater than those in SPT-2012. Since the drawdowns have been normalized with pumping rates, differences in pumping rates are not the cause but other factors. Nevertheless, as indicated in Fig. 3a and b, the general shapes of drawdown-time curves of the same observation wells in response to pumping at different wells (effects of heterogeneity) appear similar (i.e., reproducible) over the two experiments.

The first two columns of Fig. 4a–k show the contour map of the initial head distribution before each of the 11 pumping tests of SPT-2010 and SPT-2012. Name of the pumping well and the year are labeled in figures. These maps were constructed using the kriging tool built in Tecplot (i.e., unknown correlation scales and variograms) and the observed heads at the 11 wells before pumping started. These head values were obtained using measurements of the water levels in the wells before each pumping test started. The third column illustrates the scatter plots of the normalized drawdowns observed in SPT-2010 versus

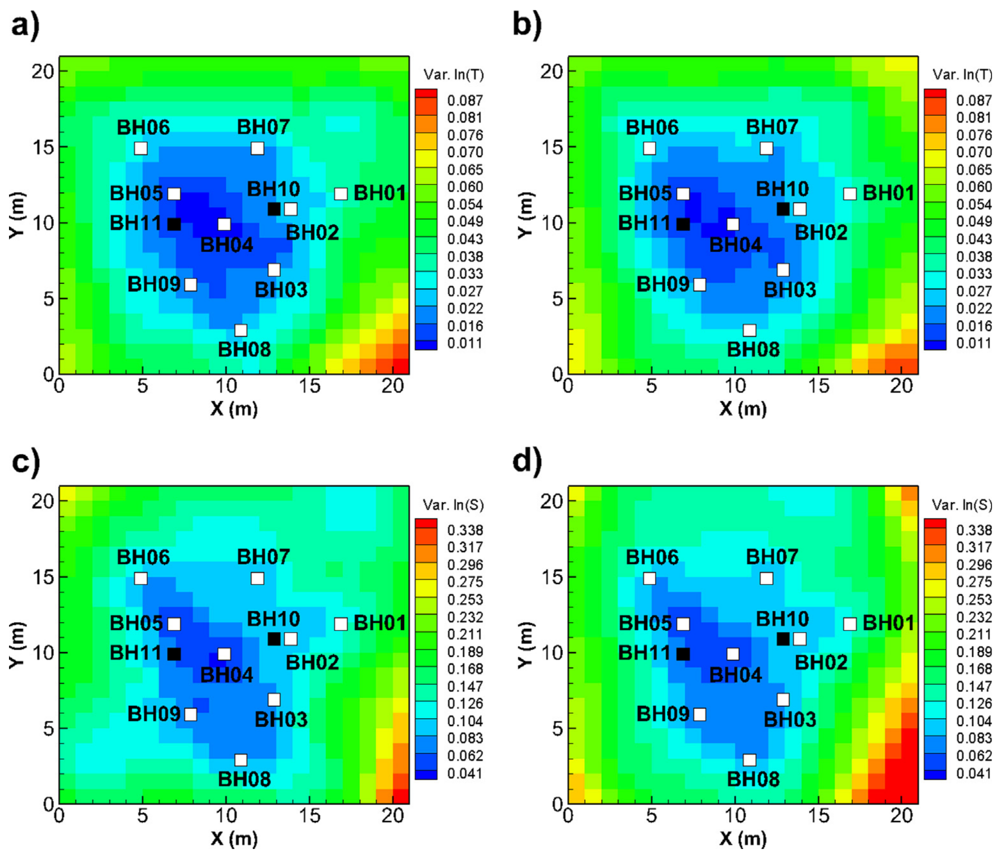

Fig. 6. The residual variance of estimated ln*T* of (a) 2010, and (b) 2012 from drawdown datasets of 9 pumping tests. The residual variance of estimated ln*S* of (c) 2010, and (d) 2012.

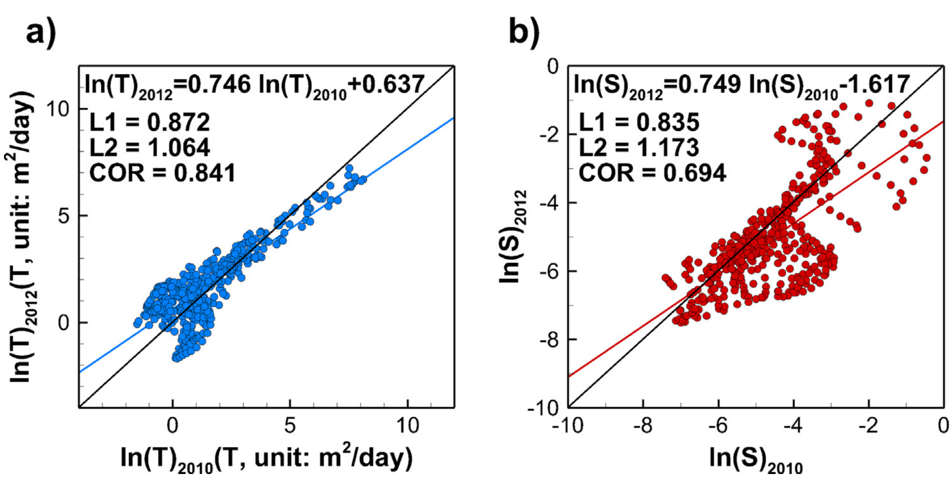

Fig. 7. Comparison of the estimated (a) ln*T* and (b) ln*S* between 2010 and 2012 based on 9 pumping data sets and conditional HT analysis (note: each plot has 441 datasets).

those in SPT-2012 for the corresponding pumping test. Generally, the spatial trends of the initial head distributions before each pumping test between the two years are similar, indicating a regional flow pattern from southeast to northwest direction, but they are different locally. In particular, notice the persistently higher initial heads near the southwest corners of Fig. 4b–e, and h–k of SPT-2010 than those of SPT-2012. These are reflections of higher head values at BH03, BH09, BH08, BH11, BH05, and BH04, rather than effects of interpolations. These differences also are manifested in the scatter plots of drawdowns as displayed in the third column of the figures. That is, most pairs of 2010 and 2012 drawdowns cluster along a straight line, but some pairs at large drawdowns deviate from the line, indicative of larger drawdowns at some wells during SPT-2010 than SPT-2012 experiments. Otherwise, the drawdowns are generally reproducible.

5.3. Estimation of T and S using field data

Using the field pumping test data, we first estimated the effective *T* and *S* based on the EHM (they are the unconditional effective parameters) first, and we then estimated *T* and *S* fields of an HPHM (conditional effective parameters). The word “unconditional” hereafter implies that the parameters do not attempt to reproduce the observed drawdown at each observation location but the overall drawdown trend of all the wells. On the other hand, the term “conditional” means that

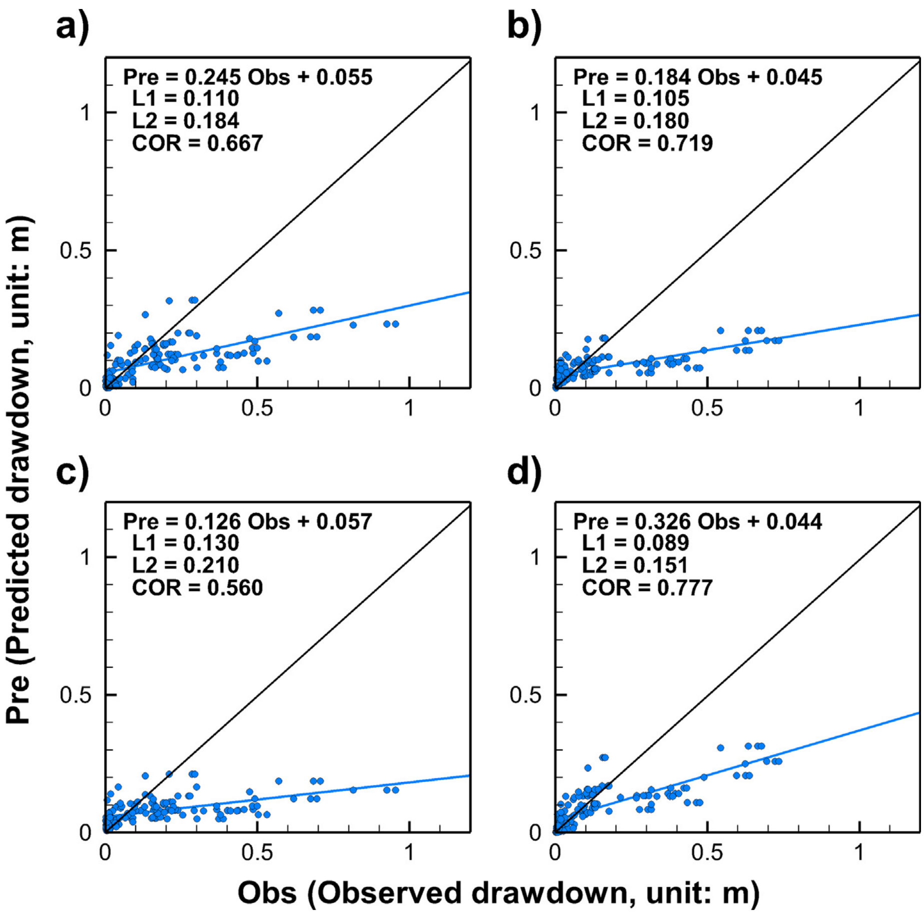

Fig. 8. Validation of effective T and S estimates based on unconditional direct average approach: Self-validation (a) using the estimates from SPT-2010 to predict drawdowns of two independent pumping test in SPT-2010; (b) using the estimates from SPT-2012 to predict drawdowns of the independent tests in SPT-2012. Cross-validation (c) using the estimates from SPT-2012 to predict drawdowns in two independent tests in SPT- 2010; (d) using the estimates from SPT-2010 to predict drawdowns in SPT-2012 (note: each plot has 140 datasets).

the parameters aim to reproduce the observed heads at each location and time as close as possible.

Two approaches were employed for estimating the unconditional effective T and S parameters (i.e., uniform effective T and S of an EHM). The first approach adopted the traditional *Theis* analysis based on the data from 11 pairs of cross-hole tests and averaged these estimated T and S values (we will call this as the unconditional direct average approach). The other approach (the unconditional HT approach) employed SimSLE to estimate the effective T and S for the EHM associated with this aquifer, using the drawdown-time data sets (110 sets) from the 10 pumping tests, simultaneously. While this approach also estimates the effective T and S of the EHM, the estimates should be less scenario dependent because of the effects of scanning. That is, they consider different flow fields under various stresses.

Afterward, a conditional HT approach, which employs the SimSLE with HPHM, was used to estimate the T and S values at each element of the numerical model for the field site.

For these estimations, a study domain of 30 m × 30 m in size was selected, covering all the eleven wells of the field site and avoiding boundary effects on the simulations. The domain was discretized into 900 elements of 1 m × 1 m. Eleven wells were located within the red square (441 elements, 21 m × 21 m) (Fig. 1). Boundary and initial conditions for the HT analysis of SPT-2010 and SPT-2012 data sets were assumed the arithmetic average of the initial water levels of all the tests in SPT-2010 and all of those in SPT-2012. They are 46.44 m and 46.33 m, respectively. All the observed drawdown-time data were converted to water level-time data, and therefore, the differences in initial and boundary conditions were avoided.

For the parameter estimation, drawdown values at 7 different times were selected from each drawdown-time data to expedite the analysis. This selection was based on the temporal sampling strategy suggested by Sun et al. (2013). That is, at the early rising limb of the drawdown-log time curve, the drawdown is most sensitive to change in S. Thus, data at three closely spaced time intervals were selected. The other four points were distributed over the rest of drawdown-time curve.

5.3.1. Unconditional T and S parameters for EHM

Direct Average Approach. The means of the effective ln$T$ and ln$S$ values of the EHM for each SPT are listed in Table 3. Notice that the T estimate from SPT-2010 are smaller than that from SPT-2012, and the S estimates from SPT-2010 is larger thant that from SPT-2012.

Unconditional HT Approach. This approach used the observed 90 drawdown data due to pumping at the 9 wells simultaneously. The simulation domain, grids, initial and boundary conditions, and input the mean and variance of ln$T$ and ln$S$ were listed in Table 2. Notice, the approach estimates only two parameters T and S, which are uniform in space since it adopts the EHM. As shown in Table 3, the estimated value for ln$T$ is 1.605, and that for ln$S$ is -4.491 based on data of SPT-2010, and that is 1.841 for ln$T$ and is -4.639 for ln$S$ based on SPT-2012 data. Similar to the direct approach, this approach also yields a smaller T estimate from SPT-2010 than that from SPT-2012, and greater S estimate from SPT-2010 than that from SPT-2012 (Table 3). Overall, this approach yields smaller T and larger S values than the direct average approach from the two SPTs.

5.3.2. Conditional T and S parameter fields for HPHM

Next, the 90 observed drawdown-time data of SPT-2010 were used to estimate the T and S values of each element in the domain of the HPHM. The simulation domain, grids, initial and boundary conditions, and input mean and variance were identical to those used in the unconditional HT approach, and the correlation scales were 10 m in both directions (Huang et al., 2011). The same procedure was applied to the

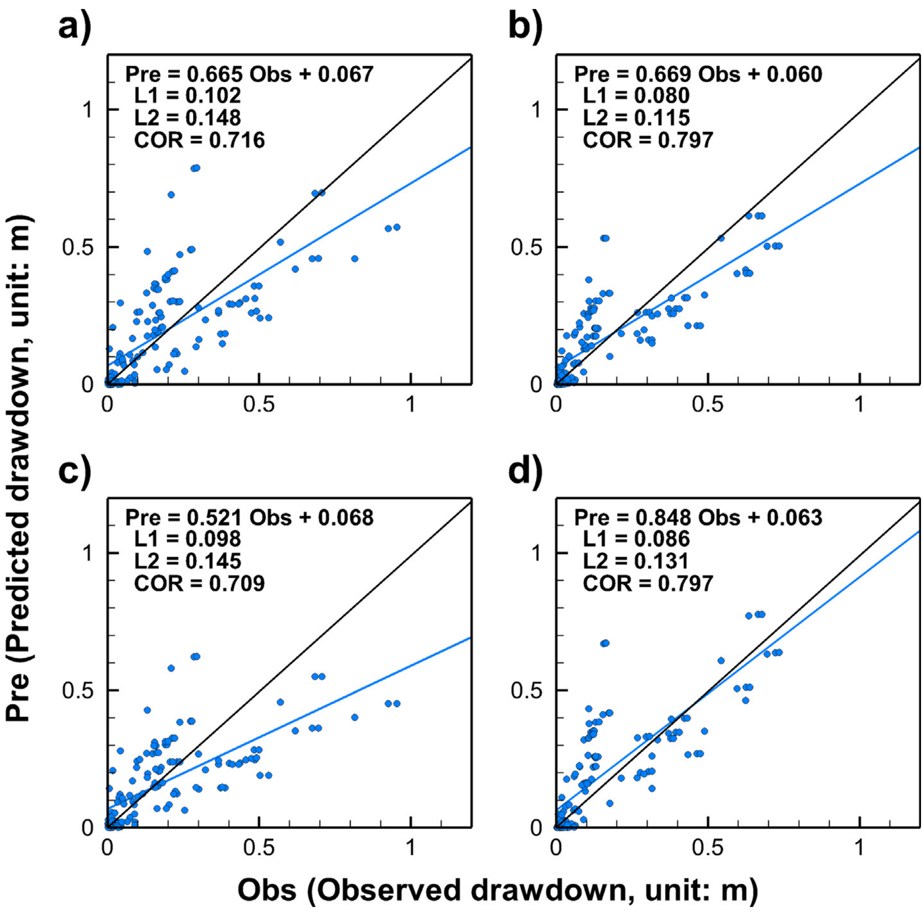

Fig. 9. Validation of effective T and S estimates based on unconditional HT approach: Self-validation (a) using the estimates from SPT-2010 to predict drawdowns of two independent pumping test in SPT-2010; (b) using the estimates from SPT-2012 to predict drawdowns of the independent tests in SPT-2012. Cross-validation (c) using the estimates from SPT-2012 to predict drawdowns in two independent tests in SPT-2010; (d) using the estimates from SPT-2010 to predict drawdowns in SPT-2012 (note: each plot has 140 data sets).

data from SPT-2012. Notice that data recorded from 10 observation wells from the two pumping tests at BH10 and BH11 were excluded from this estimation; they were treated as independent pumping tests for validation of the estimates (Section 5.4).

The estimated T fields of SPT-2010 and SPT-2012 are depicted in Fig. 5a and b, respectively. They show a similar pattern of ln$T$ distributions: a high T zone (red color) appears around BH01, BH02, and BH10 and low T zones around the boundaries. Similarly, the estimated S fields for SPT-2010 and SPT-2012 (Fig. 5c and d, respectively) are low around the lower, left and right boundaries and high S values along the upper (north) boundaries with two high-value zones at the low corners in SPT-2010 estimated field (Fig. 5c). As reported by Sun et al. (2013), misrepresenting no-flux boundaries as constant head boundaries during HT analysis could lead to low T zones around the boundaries and vice versa. Accordingly, we postulate that the low T and high S zones near the upper boundaries may be indicative of the existence of low permeable boundaries.

Notice that the SPT-2010 S estimates suggest two high S regions near the lower boundary where SPT-2012 estimates infer low S values. These two anomalies appear to agree the greater drawdowns observed in SPT-2010 than those in SPT-2012 since more water was stored at these locations and less water was available for the pumping. The differences in the initial head distribution between SPT-2010 and SPT-2012 (Fig. 4) seem to collaborate the differences in the estimates.

The residual variance maps of the estimated T fields for SPT-2010 and SPT-2012 are illustrated in Fig. 6a and b, respectively, while those of the estimated S fields are in Fig. 6c and d. The patterns of the residual variance maps for T estimates of the two SPTs are similar. The same is also true for the S estimates. The similarities are attributed to the fact that the residual variances solely depend on the locations of the wells and cross-correlations between head and T or S. A comparison of the residual variance maps of T and S, the area of low variances of T estimates is broader than that of S estimates. This is consistent with the results of cross-correlation analysis by Wu et al. (2005) and Sun et al. (2013): the observed head at a well is highly correlated with the T anomalies at upstream regions of the pumping well and the observation well. Meanwhile, the head is correlated highly with the S anomalies in a narrow area between the pumping and the observation well. Therefore, the estimates of S over the domain involve greater uncertainty than the T estimates. Notice the slight differences between the residual variance maps of S between Fig. 6c and d, likely due to numerical accuracy issues (i.e., truncation errors and convergence criteria).

Reproducibility of the T estimates is shown in Fig. 7a, where the estimated Ts from SPT-2010 are plotted against those from SPT-2012, whereas that of the S estimates is in Fig. 7b, in which the estimated Ss from SPT-2010 are plotted against those from SPT-2012. Overall, these estimates are reproducible in terms of the general trends. However, T estimates are generally more reproducible than S estimates, owing to the narrow area of the cross-correlation between head measurements and the S anomalies. For this reason, densely distributed observation wells are needed to estimate S distribution correctly (Sun et al., 2013).

5.4. Self- and Cross-Validations

The estimated T and S fields from the field experiments cannot be directly compared with the unknown "true" fields. In addition, the ultimate goal of an aquifer characterization is to predict flow or solute transport processes in the aquifer accurately under any stresses. Therefore, validation of the estimates by predicting drawdown fields (flow fields), which are not used during the inverse modeling efforts, is a desirable approach. As mentioned previously, the pumping wells (BH10 and BH11) were not used as either a pumping or an observation

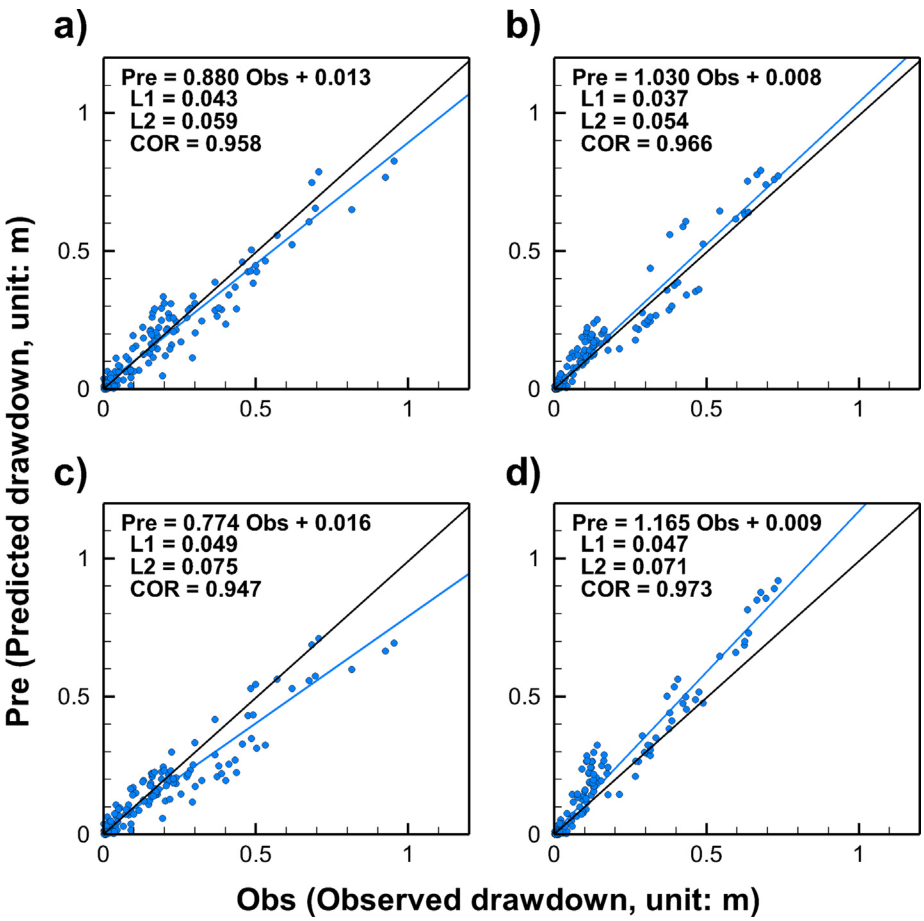

Fig. 10. Validation of effective T and S estimates based on conditional HT approach: Self-validation (a) using the estimates from SPT-2010 to predict drawdowns of two independent pumping test in SPT-2010; (b) using the estimates from SPT-2012 to predict drawdowns of the independent tests in SPT-2012. Cross-validation (c) using the estimates from SPT-2012 to predict drawdowns in two independent tests in SPT-2010; (d) using the estimates from SPT-2010 to predict drawdowns in SPT-2012 (note: each plot has 140 datasets).

well in the inverse analysis. These pumping tests thus created new flow fields, and do not have the reciprocity issue in SPT experiments (i.e., they do not carry repeated information about the heterogeneity). Further, the cross-correlation analysis suggests that heterogeneity around BH10 and BH11 likely remains unknown even after the calibration. That is to say, if the estimates from the 9 pumping tests do not capture the heterogeneity at the site, they will not predict satisfactorily the drawdown fields induced by the two new tests, and vice versa.

For these reasons, we use the estimated fields from the SPT-2010 and SPT-2012 experiments to forecast the temporal evolution of drawdowns at 10 wells of the pumping events at BH10 and BH11 during 2010 and 2012 (i.e., "self-validation"). In addition, "cross-validations" were conducted, in which the estimated T and S fields from SPT-2010 were used to predict the observed drawdown fields induced by the two independent pumping tests in SPT-2012. Likewise, estimates from SPT-2012 are used to simulate the observed drawdown fields from SPT-2010. This cross-validation also serves as a means to check the reproducibility of the estimates with respect to their forecast abilities.

Results of self-validation of the estimates from the unconditional direct average approach, based on SPT-2010 and SPT-2012 data are presented as scatter plots of the predicted vs. the observed heads in Fig. 8a and b, respectively. The cross-validations of the two estimates are illustrated in Fig. 8c and d. According to these figures, both estimated fields yield significantly biased head fields of the two tests. Specifically, the predicted drawdowns are significantly lower than the observed ones, indicative of over-estimate of T and under-estimate of S values (Table 3).

Self-validation results of the estimates based on unconditional HT approach with the data of SPT-2010 and SPT-2012 are displayed in Fig. 9a and b. Meanwhile, results of the cross-validations of the two sets of estimates are shown in Fig. 9c and d. As expected, the T and S estimates from joint inversion (HT) produce less biased predictions in comparison with those in Fig. 8a and b. That is, the estimates using SimSLE are more realistic than those based on the direct average approach. Again, the use of the 9 pumping test datasets allows the estimates to consider the 9 different flow fields simultaneously such that the estimates are representative (i.e., ergodic) for different flow fields. The predicted head values are dispersed albeit unbiased. That is, they do not capture detailed behaviors of the aquifer because of unresolved heterogeneity and scale disparity between the head measurements and model (see Yeh et al., 2015a,b).

The self- and cross-validation results of conditional estimates for the HPHM based on SPT-2010 and SPT-2012 are displayed in Fig. 10a–d, respectively. These results indicate that the estimates yield less biased predicted heads of the two independent pumping tests with smaller scatterings (variances) than the unconditional effective T and S of EHM. Similarly, cross-validation results suggest that the T and S estimates based on SPT-2010 and SPT-2012 data are satisfactory.

In addition to the scatter plots in Fig. 10, observed, predicted well hydrographs and their upper and lower bounds of the SPT-2010 and SPT-2012 self-validation are presented in Figs. 11 and 12, respectively, and cross-validation in Figs. 13 and 14. The upper and lower bounds are the predicted head based on the conditional HT estimates plus and minus one residual standard deviation (i.e., the square root of the residual variance, Fig. 6) of the estimates. Notice that the predicted heads within the bounds are located at the locations where the residual variance are small and vice versa. We emphasize the fact that the magnitude of the residual variance depends on the value of the input guess variance for SLE. Its real magnitude does not affect the estimates and the pattern but the magnitude of the residual variance. As mentioned before, the input variance for the SLE was derived from the estimates based on *Theis* approach. These estimates are spatially smooth and have

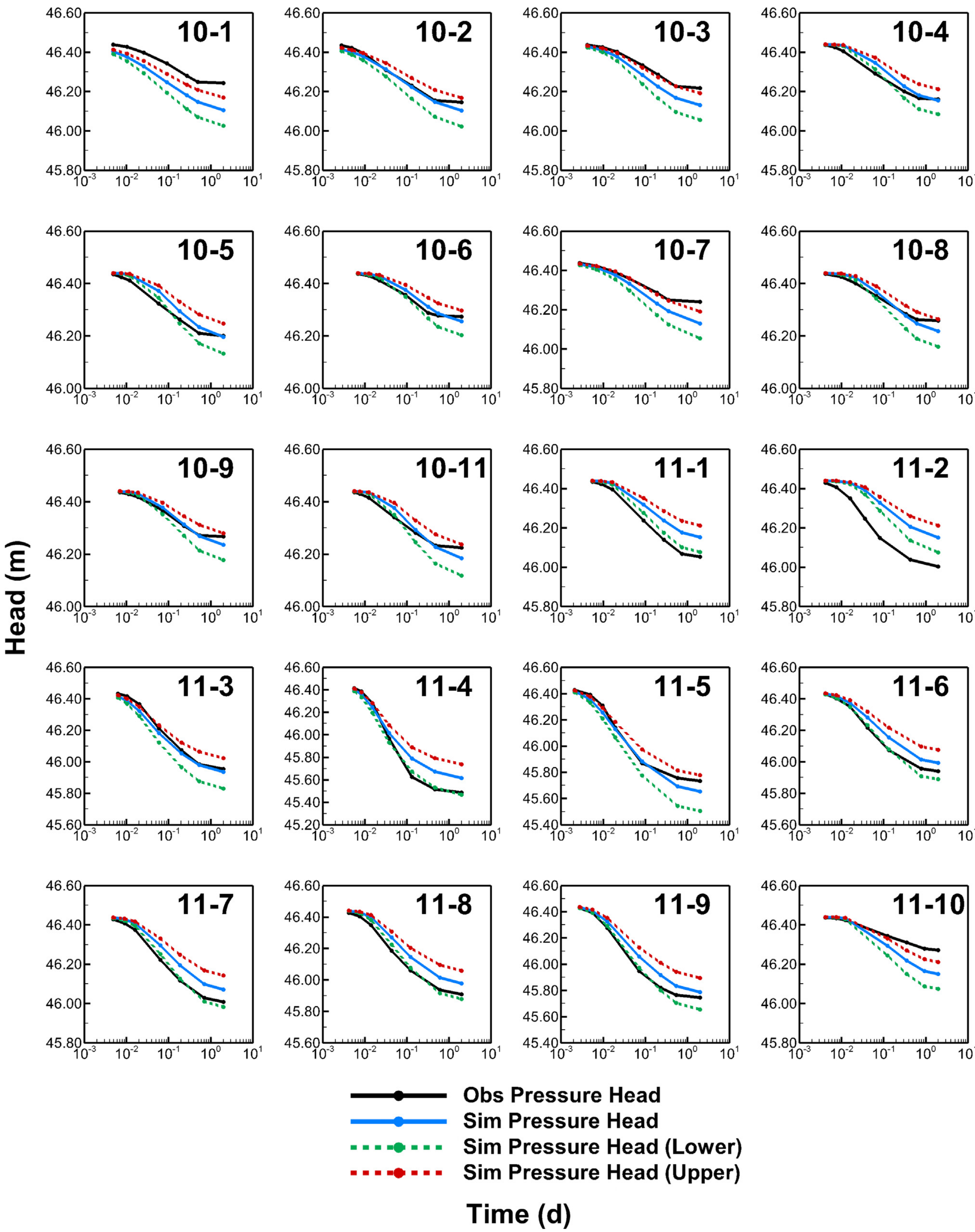

Fig. 11. Predicted well hydrographs with associated upper and lower bounds and observed hydrographs of the SPT-2010 self-validation, using pump wells, BH10 and BH11 individually. Nomenclature: 10-1 indicates that BH10 is the pumping well and BH01 is the observation well. The rest follows.

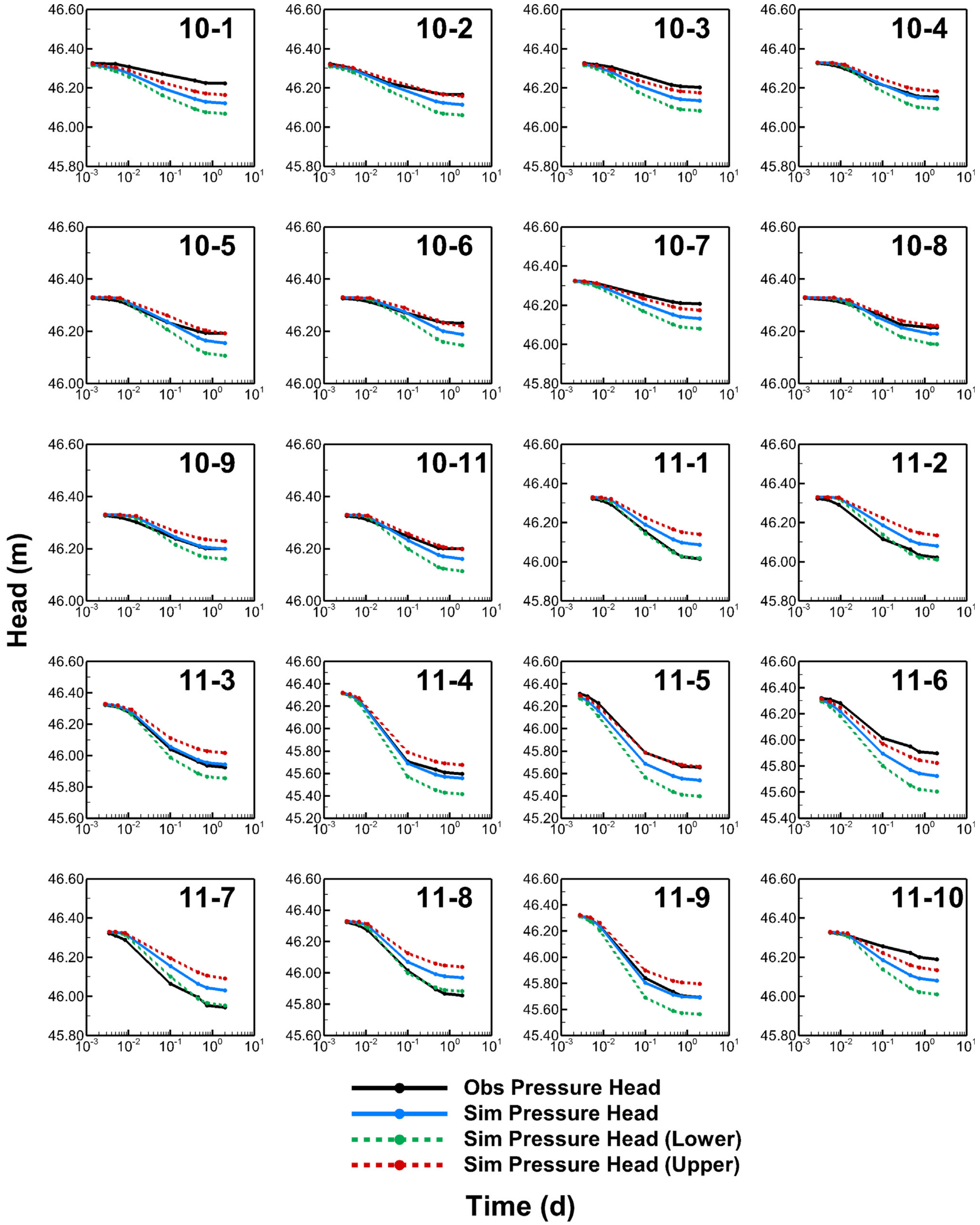

Fig. 12. Predicted well hydrographs with associated upper and lower bounds and observed hydrograph of the SPT-2012 self-validation, using pump wells, BH10 and BH11 individually. Nomenclature is the same as in Fig. 11.

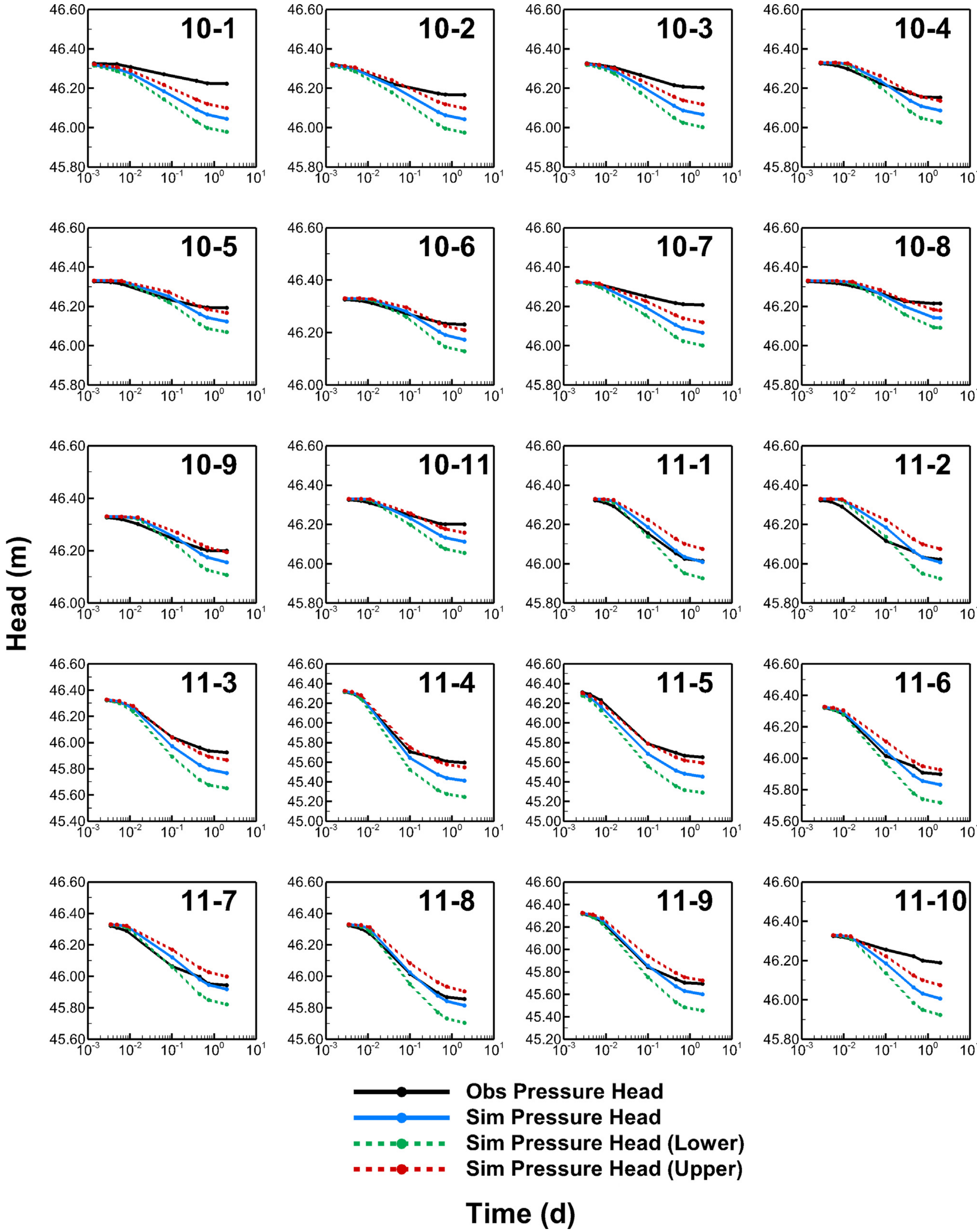

Fig. 13. Predicted well hydrographs with associated upper and lower bounds and the observed hydrographs of the SPT-2010 cross-validation: using the conditional T and S estimates from SPT-2010 to predict drawdowns in two independent tests (BH10 and BH11) in SPT-2012. Nomenclature is the same as in Fig. 11.

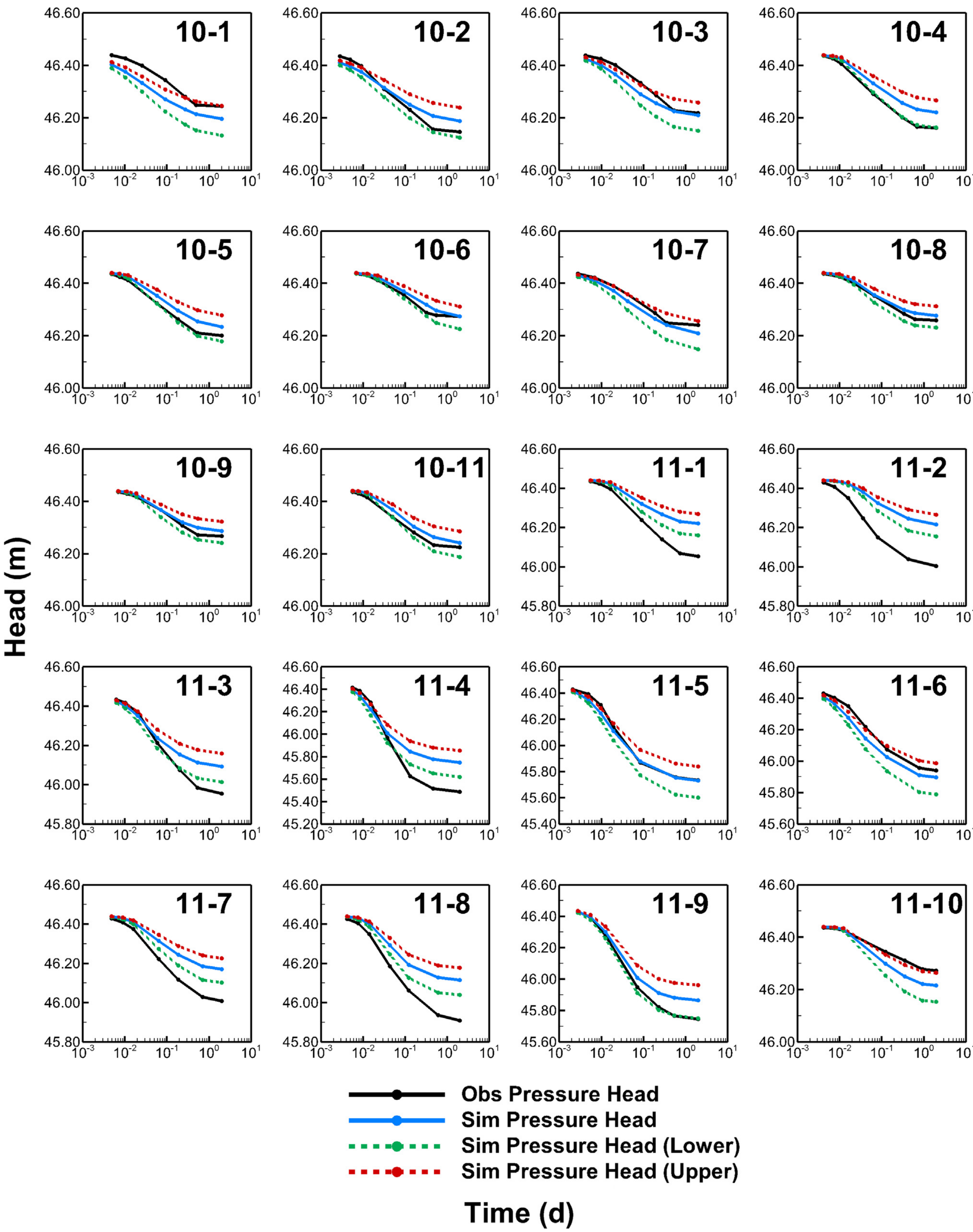

Fig. 14. Predicted well hydrographs with associated upper and lower bounds and the observed hydrographs of the SPT-2012 cross-validation: using the estimates from SPT- 2012 to predict drawdowns in two independent tests (BH10 and BH11) in SPT-2010. Nomenclature is the same as in Fig. 11.

small spatial variability and in turn, the variance. In order words, the actual variance (or spatial variability) is likely larger than this value, and the gap between the upper and lower bound should be larger. Therefore, all predicted well hydrographs would be enclosed within these bounds.

6. Discussion

This study shows that the shapes of the drawdown hydrographs of the SPT-2010 and SPT-2012 were generally the same, but the magnitudes of the drawdowns in SPT-2010 were greater than those in SPT-2012. Since the drawdowns were normalized by the pumping rates, the difference in magnitudes may be attributed to other factors, such as regional flow, recharge, or alteration of the hydraulic properties due to earthquakes or other processes. On the other hand, the time axis of the drawdown curves was normalized by the square of the distance between the pumping well and each observation, and the shapes of the drawdown time curves of the two tests are similar. Accordingly, we may exclude the possibility of changes in hydraulic properties of the site over the two years.

Yet, contoured maps of the initial head fields before each of the experiments show that the regional flow trends, as well as the boundary conditions, are similar but different locally, particularly at the southern boundary. Despite the difference, we believe that the spatial trends of the initial head should not affect the estimates of unconditional T and S for this field site because drawdowns were used in the above analysis. Nevertheless, both unconditional direct average and HT EHM approach yielded smaller T estimates from SPT-2010 than those from SPT-2012, and larger S estimates from SPT-2010 than those from SPT-2012. Interestingly, such results seem to explain the large drawdowns observed in SPT-2010. If this is true, the hydraulic properties then have been altered over the two-year time span – contradicting the previous conclusion based on the shape of the drawdown-time curves.

We also recognized that the conditional HT HPHM estimates for the two experiments, closely resemble each other, with some noticeable differences in S estimates near one of the boundaries of the aquifer. In effect, these results are consistent with the estimates based on unconditional approaches. That is, the large S values along the south side of the boundary contribute to the large estimated unconditional S value. The question now arises: If these changes in S estimate at the southern boundary are real or artifact.

As discussed above, it is unlikely the aquifer properties could change significantly over the two years period and the differences in initial and boundary conditions unlikely could affect the estimates. Therefore, we speculate that the discrepancies may be attributed to unexplored processes omitted in this depth-average 2-D model, such as recharge or root extractions. Perhaps, these discrepancies may also support the claim that the T and S parameters in a depth-average, two-dimensional flow model are phenomenological parameters. That is, the effects of the vertical variations of hydraulic conductivities (see Fig. 2) and hydraulic gradients of the three-dimensional, heterogeneous aquifer are lumped into the effective T and S parameters. As the regional flow or other processes change, the vertical gradient varies, and in turn, the parameter values (p168, Yeh et al., 2015a).

On the other hand, since the wells are 20 m long, one might speculate upon the wellbore effect at the pumping wells. Albeit the heads of the pumping wells were not used in the analysis and early time drawdowns, which may be affected by the wellbore effect, were avoided, the wellbore effect at the pumping wells might slightly delay the arrival of drawdown at the observation wells and in turn, affect the estimates of S. However, as indicated by the residual variance, accurate S estimates are inherently difficult to obtain in spite of the absence of the borehole effect. Note that the T estimates are not affected by the borehole effects since they are not influenced by the early time data.

The self-validation and cross-validation of the estimated T and S fields using three different approaches yield different results. The unconditional T and S values using the direct average approach consistently yield biased predicted drawdowns. On the other hand, the unconditional T and S values based on the unconditional HT approach yield less biased prediction with similar degrees of scatterings. This corroborates with the result that unconditional T and S estimates based on the direct average approach are much larger than those based on the unconditional HT approach are. Simultaneous inclusion of all the test data into the analysis for the effective T and S of EHM corrects the inconsistencies between the estimates from using *Theis* solution for each cross-hole test. Further, the unconditional HT approach (even it is unconditional) using a large number of data reduces the effects of noises on the estimates.

As expected, the estimates from the conditional HT approach result in drawdown field predictions with the least bias and variance. The reproducibility of the productivity of the estimated T and S is generally satisfactory although it is imperfect. This imperfection is likely due to unknown processes, which are neglected in the depth-average model. These issues are the subject of our next research. While HT is a logical approach for collecting analyzing data (Yeh and Lee, 2007), a fully 3-D dimensional modeling and monitoring of flow in variably saturated media may be the only way to unravel the unknown processes and to reduce phenomenological effects on parameters.

In conclusion, as articulated by Yeh et al. (2015a,b), reproducibility and predictability in groundwater hydrology would depend on the model resolution (i.e., model scale), and scales of our observation and interest. At coarse resolution, the traditional homogeneous approaches (coarse model scale) may work – they may yield unbiased parameter estimates and flow predictions with large uncertainty at fine observation scales. On the other hand, HT with highly parameterized models (fine model scale) can lead to unbiased and less uncertain parameter estimates and flow predictions at the fine observation scale. Nonetheless, reproducibility and prediction of groundwater flow always involve some uncertainty due to multi-scale temporal and spatial variabilities and the limited resolutions of our models (model scale) as well as noise and model structure errors.

The results of this study corroborate the theme of Konikow and Bredehoeft (1992): model assessment or validation aims to improve our understanding of the particular hydrologic system or problem of interest. Nonetheless, results of this study offer a more optimistic view than that those expressed by them. That is, if the aquifer is sufficiently characterized in terms of initial, boundary, hydraulic parameters, and other processes, the groundwater model is not just a tool for better understanding of the system, but also a tool can be used to provide reasonable forecasts, so long as scales of the model and observation are consistent (Yeh et al., 2015a,b). Of course, our statement rests upon the results from a small experimental site and the two-year periods of our investigation.

Acknowledgments

Jet-Chau Wen would like to acknowledge the research support from NSC 101-2221-E-224-050, NSC 102-2221-E-224-050, MOST 103-2221-E-224-054, MOST 104-2221-E-224-039, MOST 105-2625-M-224-002, MOST 106-2625-M-224-002, and MOST 107-2625-M-224-002, by the Minister of Science and Technology, Taiwan. Tian-Chyi Jim Yeh acknowledges the Global Expert award through Tianjin Normal University from the Thousand Talents Plan of Tianjin City, and the support from CRDF under the award number (DAA2-15-61224-1): Hydraulic tomography in shallow alluvial sediments: Nile River Valley, Egypt. All data from this work is available upon request through the corresponding author.

References

Berg, S.J., Illman, W.A., 2013. Field study of subsurface heterogeneity with steady-state hydraulic tomography. Groundwater 51, 29–40. https://doi.org/10.1111/j.1745-

6584.2012.00914.x.
Berg, S.J., Illman, W.A., 2011a. Capturing aquifer heterogeneity: comparison of approaches through controlled sandbox experiments. Water Resour. Res. 47, W09514. https://doi.org/10.1029/2011WR010429.
Berg, S.J., Illman, W.A., 2011b. Three-dimensional transient hydraulic tomography in a highly heterogeneous glaciofluvial aquifer-aquitard system. Water Resour. Res. 47, W10507. https://doi.org/10.1029/2011WR010616.
Bohling, G.C., Butler, J.J., 2010. Inherent limitations of hydraulic tomography. Ground Water 48, 809–824. https://doi.org/10.1111/j.1745-6584.2010.00757.x.
Cardiff, M., Bakhos, T., Kitanidis, P.K., Barrash, W., 2013. Aquifer heterogeneity characterization with oscillatory pumping: sensitivity analysis and imaging potential. Water Resour. Res. 49, 5395–5410. https://doi.org/10.1002/wrcr.20356.
Cooper, H.H., Jacob, C.E., 1946. A generalized graphical method for evaluating formation constants and summarizing well-field history. Eos Trans. AGU 27, 526–534.
Hantush, M.S., Jacob, C.E., 1955. Non-steady radial flow in an infinite leaky aquifer. Eos Trans. AGU 36, 95–100. https://doi.org/10.1029/TR036i001p00095.
Hao, Y., Yeh, T.-C.J., Xiang, J., Illman, W.A., Ando, K., Hsu, K.-C., Lee, C.-H., 2008. Hydraulic tomography for detecting fracture zone connectivity. Ground Water 46, 183–192. https://doi.org/10.1111/j.1745-6584.2007.00388.x.
Hochstetler, D.L., Barrash, W., Leven, C., Cardiff, M., Chidichimo, F., Kitanidis, P.K., 2016. Hydraulic tomography: continuity and discontinuity of high-k and low-k zones. Groundwater 54, 171–185. https://doi.org/10.1111/gwat.12344.
Huang, S.-Y., Wen, J.-C., Yeh, T.-C.J., Lu, W., Juan, H.-L., Tseng, C.-M., Lee, J.-H., Chang, K.-C., 2011. Robustness of joint interpretation of sequential pumping tests: numerical and field experiments. Water Resour. Res. 47, W10530. https://doi.org/10.1029/2011WR010698.
Hughson, D.L., Yeh, T.-C.J., 2000. An inverse model for three-dimensional flow in variably saturated porous media. Water Resour. Res. 36, 829–839. https://doi.org/10.1029/2000WR900001.
Illman, W.A., Berg, S.J., Yeh, T.-C.J., 2012. Comparison of approaches for predicting solute transport: sandbox experiments. Ground Water 50, 421–431. https://doi.org/10.1111/j.1745-6584.2011.00859.x.
Illman, W.A., Berg, S.J., Liu, X., Massi, A., 2010. Hydraulic/partitioning tracer tomography for DNAPL source zone characterization: small-scale sandbox experiments. Environ. Sci. Technol. 44, 8609–8614. https://doi.org/10.1021/es101654j.
Illman, W.A., Liu, X., Craig, A.J., 2008. Evaluation of transient hydraulic tomography and common hydraulic characterization approaches through laboratory sandbox experiments. J. Environ. Eng. Manage 18, 249–256.
Illman, W.A., Liu, X., Takeuchi, S., Yeh, T.-C.J., Ando, K., Saegusa, H., 2009. Hydraulic tomography in fractured granite: Mizunami underground research site. Japan. Water Resour. Res. 45, W01406. https://doi.org/10.1029/2007WR006715.
Kitanidis, P.K., 1995. Quasi-inear geostatistical theory for inversing. Water Resour. Res. 31, 2411–2419. https://doi.org/10.1029/95WR01945.
Konikow, L.F., Bredehoeft, J.D., 1992. Ground-water models cannot be validated. Adv. Water Resour. 15, 75–83. https://doi.org/10.1016/0309-1708(92)90033-X.
Li, W., Englert, A., Cirpka, O.A., Vanderborght, J., Vereecken, H., 2007. Two-dimensional characterization of hydraulic heterogeneity by multiple pumping tests. Water Resour. Res. 43, W04433. https://doi.org/10.1029/2006WR005333.
Liu, X., Illman, W.A., Craig, A.J., Zhu, J., Yeh, T.-C.J., 2007. Laboratory sandbox validation of transient hydraulic tomography. Water Resour. Res. 43, W05404. https://doi.org/10.1029/2006WR005144.
Mao, D., Wan, L., Yeh, T.-C.J., Lee, C.-H., Hsu, K.-C., Wen, J.-C., Lu, W., 2011. A revisit of drawdown behavior during pumping in unconfined aquifers. Water Resour. Res. 47, W05502. https://doi.org/10.1029/2010WR009326.
Mao, D., Yeh, T.-C.J., Wan, L., Wen, J.-C., Lu, W., Lee, C.-H., Hsu, K.-C., 2013. Joint interpretation of sequential pumping tests in unconfined aquifers. Water Resour. Res. 49, 1782–1796. https://doi.org/10.1002/wrcr.20129.
Meier, P.M., Carrera, J., Sánchez-Vila, X., 1998. An evaluation of Jacob's method for interpretation of pumping tests in heterogeneous formations. Water Resour. Res. 34, 1011–1025. https://doi.org/10.1029/98WR00008.
Ni, C.-F., Yeh, T.-C.J., Chen, J.-S., 2009. Cost-effective hydraulic tomography surveys for predicting flow and transport in heterogeneous aquifers. Environ. Sci. Technol. 43, 3720–3727. https://doi.org/10.1021/es8024098.
Sanchez-León, E., Leven, C., Haslauer, C.P., Cirpka, O.A., 2016. Combining 3D hydraulic tomography with tracer tests for improved transport characterization. Groundwater 54, 498–507. https://doi.org/10.1111/gwat.12381.
Sánchez-Vila, X., Meier, P.M., Carrera, J., 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, 943–952. https://doi.org/10.1029/1999WR900007.
Sharmeen, R., Illman, W.A., Berg, S.J., Yeh, T.-C.J., Park, Y.-J., Sudicky, E.A., Ando, K., 2012. Transient hydraulic tomography in a fractured dolostone: laboratory rock block experiments. Water Resour. Res. 48, W10532. https://doi.org/10.1029/2012WR012216.
Straface, S., Yeh, T.-C.J., Zhu, J., Troisi, S., Lee, C.-H., 2007. Sequential aquifer tests at a well field, Montalto Uffugo Scalo, Italy. Water Resour. Res. 43, W07432. https://doi.org/10.1029/2006WR005287.
Sun, R., Yeh, T.-C.J., Mao, D., Jin, M., Lu, W., Hao, Y., 2013. A temporal sampling strategy for hydraulic tomography analysis. Water Resour. Res. 49, 3881–3896. https://doi.org/10.1002/wrcr.20337.
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.
Tso, C.-H.M., Zha, Y., Yeh, T.-C.J., Wen, J.-C., 2016. The relative importance of head, flux, and prior information in hydraulic tomography analysis. Water Resour. Res. 52, 3–20. https://doi.org/10.1002/2015WR017191.
Vesselinov, V.V., Neuman, S.P., Illman, W.A., 2001a. Three-dimensional numerical inversion of pneumatic cross-hole tests in unsaturated fractured tuff: 1. Methodology and borehole effects. Water Resour. Res. 37, 3001–3017. https://doi.org/10.1029/2000WR000133.
Vesselinov, V.V., Neuman, S.P., Illman, W.A., 2001b. Three-dimensional numerical inversion of pneumatic cross-hole tests in unsaturated fractured tuff: 2. Equivalent parameters, high-resolution stochastic imaging and scale effects. Water Resour. Res. 37, 3019–3041. https://doi.org/10.1029/2000WR000135.
Wen, J.-C., Wu, C.-M., Yeh, T.-C.J., Tseng, C.-M., 2010. Estimation of effective aquifer hydraulic properties from an aquifer test with multi-well observations (Taiwan). Hydrogeol. J. 18, 1143–1155. https://doi.org/10.1007/s10040-010-0577-1.
Wu, C.-M., Yeh, T.-C.J., Zhu, J., Lee, T.H., Hsu, N.-S., Chen, C.-H., Sancho, A.F., 2005. Traditional analysis of aquifer tests: comparing apples to oranges? Water Resour. Res. 41, W09402. https://doi.org/10.1029/2004WR003717.
Xiang, J., Yeh, T.-C.J., Lee, C.-H., Hsu, K.-C., Wen, J.-C., 2009. A simultaneous successive linear estimator and a guide for hydraulic tomography analysis. Water Resour. Res. 45, W02432. https://doi.org/10.1029/2008WR007180.
Yeh, T.-C.J., Gutjahr, A.L., Jin, M., 1995. An iterative cokriging-like technique for groundwater flow modeling. Ground Water 33, 33–41. https://doi.org/10.1111/j.1745-6584.1995.tb00260.x.
Yeh, T.-C.J., Lee, C.-H., Hsu, K.-C., Illman, W.A., Barrash, W., Cai, X., Daniels, J., Sudicky, E., Wan, L., Li, G., Winter, C.L., 2008. A view toward the future of subsurface characterization: CAT scanning groundwater basins. Water Resour. Res. 44, W03301. https://doi.org/10.1029/2007WR006375.
Yeh, T.-C.J., Jin, M., Hanna, S., 1996. An iterative stochastic inverse method: conditional effective transmissivity and hydraulic head fields. Water Resour. Res. 32, 85–92. https://doi.org/10.1029/95WR02869.
Yeh, T.-C.J., Liu, S., 2000. Hydraulic tomography: development of a new aquifer test method. Water Resour. Res. 36, 2095–2105. https://doi.org/10.1029/2000WR900114.
Yeh, T.-C.J., Lee, C.-H., 2007. Time to change the way we collect and analyze data for aquifer characterization. Ground Water 45, 116–118. https://doi.org/10.1111/j.1745-6584.2006.00292.x.
Yeh, T.-C.J., Khaleel, R., Carroll, K.C., 2015a. Flow Through Heterogeneous Geologic Media. Cambridge, New York.
Yeh, T.-C.J., Mao, D., Zha, Y., Wen, J.-C., Wan, L., Hsu, K.-C., Lee, C.-H., 2015b. Uniqueness, scale, and resolution issues in groundwater model parameter identification. Water Sci. Eng. 8, 175–194. https://doi.org/10.1016/j.wse.2015.08.002.
Zha, Y., Yeh, T.-C.J., Mao, D., Yang, J., Lu, W., 2014. Usefulness of flux measurements during hydraulic tomographic survey for mapping hydraulic conductivity distribution in a fractured medium. Adv. Water Resour. 71, 162–176. https://doi.org/10.1016/j.advwatres.2014.06.008.
Zha, Y., Yeh, T.-C.J., Illman, W.A., Onoe, H., Mok, C.-M.W., Wen, J.-C., Huang, S.-Y., Wang, W., 2017. Incorporating geologic information into hydraulic tomography: a general framework based on geostatistical approach. Water Resour. Res. 53, 2850–2876. https://doi.org/10.1002/2016WR019185.
Zha, Y., Yeh, T.-C.J., Illman, W.A., Tanaka, T., Bruines, P., Onoe, H., Saegusa, H., 2015. What does hydraulic tomography tell us about fractured geological media? A field study and synthetic experiments. J. Hydrol. 531, 17–30. https://doi.org/10.1016/j.jhydrol.2015.06.013.
Zha, Y., Yeh, T.-C.J., Illman, W.A., Tanaka, T., Bruines, P., Onoe, H., Saegusa, H., Mao, D., Takeuchi, S., Wen, J.-C., 2016. An application of hydraulic tomography to a large-scale fractured granite site, Mizunami, Japan. Groundwater 54, 793–804. https://doi.org/10.1111/gwat.12421.
Zha, Y., Yeh, T.-C.J., Illman, W.A., Zeng, W., Zhang, Y., Sun, F., Shi, L., 2018. A reduced-order successive linear estimator for geostatistical inversion and its application in hydraulic tomography. Water Resour. Res. 54, 1616–1632. https://doi.org/10.1002/2017WR021884.
Zhang, J., Yeh, T.-C.J., 1997. An iterative geostatistical inverse method for steady flow in the vadose zone. Water Resour. Res. 33, 63–71. https://doi.org/10.1029/96WR02589.
Zhao, Z., Illman, W.A., 2017. On the importance of geological data for three-dimensional steady-state hydraulic tomography analysis at a highly heterogeneous aquifer-aquitard system. J. Hydrol. 544, 640–657. https://doi.org/10.1016/j.jhydrol.2016.12.004.
Zhao, Z., Illman, W.A., 2018. Three-dimensional imaging of aquifer and aquitard heterogeneity via transient hydraulic tomography at a highly heterogeneous field site. J. Hydrol. 559, 392–410. https://doi.org/10.1016/j.jhydrol.2018.02.024.
Zhao, Z., Illman, W.A., Yeh, T.-C.J., Berg, S.J., Mao, D., 2015. Validation of hydraulic tomography in an unconfined aquifer: a controlled sandbox study. Water Resour. Res. 51, 4137–4155. https://doi.org/10.1002/2015wr016910.
Zhu, J., Yeh, T.-C.J., 2005. Characterization of aquifer heterogeneity using transient hydraulic tomography. Water Resour. Res. 41, W07028. https://doi.org/10.1029/2004WR003790.