Wang Yu-Li (Orcid ID: 0000-0003-4324-0955)
Hao Yonghong (Orcid ID: 0000-0003-0678-0389)
Liu Fei (Orcid ID: 0000-0003-0967-9112)

Temporal Sampling and Role of Flux Measurements for Subsurface Heterogeneous Characterization in Groundwater Basins Using Hydraulic Tomography

Fei Liu [1,2], Tian-Chyi Jim Yeh [3,4,*], Xianfang Song [5], Yu-Li Wang [4], Jet-Chau Wen [6,7], Yonghong Hao [3], Wenke Wang [8]

[1] School of Water Conservancy and Hydropower, Hebei University of Engineering, 19 Taiji Road, Handan, Hebei, 056000, China

[2] Hebei Key Laboratory of Intelligent Water Conservancy, Handan, Hebei, 056000, China

[3] Tianjin Key Laboratory of Water Resources and Environment, Tianjin Normal University, Tianjin, 300387, China

[4] Department of Hydrology and Atmospheric Sciences, University of Arizona, Tucson, Arizona, 85721, USA

[5] Key Laboratory of Water Cycle and Related Land Surface Processes, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, 11 A, Datun Road, Chaoyang District, Beijing, 100101, China

[6] Department of Safety, Health, and Environmental Engineering, National Yunlin University of Science and Technology, Douliu, Taiwan, 64002, China

[7] Research Center for Soil and Water Resources and Natural Disaster Prevention, National Yunlin University of Science and Technology, Douliu, Taiwan, 64002, China

[8] Key Laboratory of Subsurface Hydrology and Ecological Effects in Arid Region, Chang'an University, Xi'an, 710054, China

[*] Corresponding author: Tian-Chyi Jim Yeh

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/hyp.14299

E-mail: yeh@hwr.arizona.edu

Running title: Basin-scale subsurface characterization

Abstract

Accurate characterization of heterogeneity in groundwater basins is crucial to the sustainable management of groundwater resources. This study explores the temporal sampling issues and the role of flux measurements in the characterization of heterogeneity in groundwater basins using numerical experiments. The experiments involve a digital basin imitating the groundwater basin of the North China Plain (NCP), where the groundwater exploitation reduction program is ongoing. Using the experiments, we champion that the reduction program could collect groundwater level information induced by operational variations of existing pumping wells at different locations in the basin. Such a dataset could serve as a basin-scale hydraulic tomography (HT) to characterize the basin-scale heterogeneity cost-effectively. Both steady-state and transient-state inversion experiments demonstrate the advantage of HT surveys in characterizing basin-scale heterogeneity over conventional pumping tests at fixed well locations. Additionally, head data at the early, intermediate, and late time from well hydrographs should be selected for the HT analysis to maximize HT's power and save computational costs. When accurate geological zones are incorporated in prior information, flux measurements significantly improve parameter estimates based on conventional pumping tests. However, their effects are less noticeable for long-term HT surveys in such basin-scale aquifers without fissures or fractures. This basin-scale tomographic survey example serves a guide for field data collection and optimization of the analysis of future basin-scale HT.

Keywords: Hydraulic tomography; Basin-scale; Groundwater management; Temporal sampling; Flux measurements; North China Plain

1. Introduction

The effective management of basin-scale groundwater resources depends upon the knowledge of aquifer properties. Behaviors of groundwater flow and solute transport are generally dictated by the spatial patterns of the heterogeneous hydraulic parameters. Thus, the reliability of subsurface heterogeneous characterization in a groundwater basin directly determines the performance of groundwater numerical simulation, prediction, and scenario analysis.

Over the past two decades, hydraulic tomography (HT) has shown a great potential for high-resolution subsurface characterization in numerical studies (Cardiff, Bakhos, Kitanidis, & Barrash, 2013; Illman, Berg, Liu, & Massi, 2010; Illman, Craig, & Liu, 2008), laboratory sandboxes (Illman, Liu, & Craig, 2007; Liu, Illman, Craig, Zhu, & Yeh, 2007; Zhao, Illman, & Berg, 2016), and field-scale studies (Berg & Illman, 2011; Cardiff, Barrash, & Kitanidis, 2012; Fischer et al., 2017; Kuhlman, Hinnell, Mishra, & Yeh, 2008; Liu et al., 2020a; Zha et al., 2016). Many researchers (such as Cardiff et al., 2009; Cho, Zhao, Thomson, & Illman, 2020; Illman et al., 2007; Soueid Ahmed, Zhou, Jardani, Revil, & Dupont, 2015; Yeh & Liu, 2000; Zhao & Illman, 2017) have demonstrated steady-state HT's ability to estimate hydraulic conductivity or transmissivity heterogeneity. Likewise, many have shown the power of transient HT to image hydraulic conductivity and specific storage fields (e.g., Cardiff, Barrash, & Kitanidis, 2013; Hao et al., 2008; Luo, Zhao, Illman, & Berg, 2017; Mohammadi & Illman, 2019; Zhao & Illman, 2018; Zhu & Yeh, 2005).

Conventional pumping tests are time-consuming and costly for the characterization of a groundwater basin. HT has significant advantages over the conventional aquifer tests since HT employs sequential pumping or injection tests in an existing well field to derive spatially varying aquifer properties. It sequentially changes the pumping or injection location and monitors aquifer responses at other wells to yield information about aquifer heterogeneity. HT is analogous to multiple snapshots of subsurface hydraulic heterogeneity at different perspectives and angles (Zhao et al., 2019). It is also equivalent to geophysical tomography surveys (Butler, McElwee, & Bohling, 1999; Jiménez, 2015). As a result, observed data from pumping at different locations throughout the aquifer contain more information about heterogeneity than a single pumping test. Thus, an inverse model with these data can lead to the detailed spatial patterns of high and low permeable zones over a large area.

However, it is challenging to effectively stress the entire groundwater basin using conventional pumping operations in the HT survey. For this reason, Yeh and Lee (2008) promoted changing the way we collect and analyze data for groundwater characterization. Yeh et al. (2008) further proposed exploiting natural stimuli (i.e., atmospheric pressure variation, lightning, solid earth tides, river stage, etc.) as possible energy sources for basin-scale HT surveys. Subsequently, Yeh et al. (2009) demonstrated the possibility of using river stage variations for basin-scale subsurface tomographic surveys. Wang et al. (2017) applied the river stage tomography concept to characterize aquifer heterogeneity in the Zhuoshui river fan, Taiwan.

To facilitate a basin-scale HT survey, Kuhlman et al. (2008) used multiple HT surveys to cover an entire basin. Likewise, Zha et al. (2019) exploited aquifer responses from the change in the flow fields due to variations in pump-and-treat operation as a large-scale HT survey. They mapped the low permeable zones for hydrofracking to enhance remediation effectiveness at Tucson, Arizona, USA. Luo et al. (2020)

demonstrated that the HT analysis using long-term pumping/injection and water-level records could yield reliable hydraulic parameter estimates, significantly improving transport predictions over a large area.

However, many practical issues facing basin-scale HT surveys remain to be investigated: (1) the duration and magnitude of the pumping rate, (2) the frequency of temporal sampling, and (3) the data fusion of head, flux and tracers for enhancement of aquifer characterization. Specifically, should field tests collect steady-state or transient data for HT inversion? How many observation heads at different times in a well hydrograph are required to obtain reasonable estimates of hydraulic parameters while minimizing computational burdens for HT analysis? Whether additional data (flux, tracer, etc.) can contribute to the improvement of the estimates?

The computational efficiency is also a challenging issue for large-scale inverse problems, and high computational costs lie in calculating sensitivities, auto-covariances, and cross-covariances (Zhao & Luo, 2020). In particular, analysis of transient data from a series of pumping tests represents a substantial computational burden (Bohling, Zhan, Butler, & Zheng, 2002). Many previous works have attempted to reduce the computational costs without much loss of accuracy (Kitanidis & Lee, 2014; Lee & Kitanidis, 2014; Lee, Yoon, Kitanidis, Werth, & Valocchi, 2016; Lin, Le, O'Malley, Vesselinov, & Bui-Thanh, 2017; Liu, Zhou, Birkholzer, & Illman, 2013; Sánchez-León, Erdal, Leven, & Cirpka, 2020; Zha et al., 2018). The steady shape approach proposed by Bohling et al. (2002) retains the computational efficiency of a steady-state analysis, in which the head gradient becomes constant with time. However, this condition exists only in the ensemble sense (Zha et al., 2016).

Given that the computational cost is overwhelming when many head data are used for an inverse model, a practical question arises: Are all head data in a well hydrograph used? Zhu and Yeh (2005) showed that the heads over time are highly correlated and suggested using head data at sparse time intervals. Cardiff et al. (2012) interpreted transient HT data only using the selected early-time, intermediate-time, and late-time data points from the drawdown curves. Sun et al. (2013) investigated optimal sampling times of drawdowns for a small-scale HT analysis. However, the temporal sampling issues of subsurface heterogeneous characterization in groundwater basins remain to be investigated.

In addition, head measurements at many fully screened wells only represent depth-averaged head values, and they do not carry information about vertical aquifer heterogeneity. Li et al. (2008) overcame this issue by using estimates of hydraulic conductivity profile from flowmeter measurements along the fully-screened well as additional information for HT. In contrast, Zha et al. (2014) developed an approach to include flux information during HT survey for mapping fracture distributions in a hypothetic geologic medium. Tso et al. (2016) demonstrated that a joint interpretation of head and flux data could enhance the resolution of the HT estimates.

The North China Plain (NCP) is one of the hotspots of groundwater depletion in China, where groundwater contributed to most of the NCP's total water consumption (Zheng et al., 2010). Numerous pumping and monitoring wells are widely distributed in the NCP, and long-term groundwater level monitoring data have been collected for many years (Cao, Zheng, Scanlon, Liu, & Li, 2013). Since 2014, the Chinese government has initiated the groundwater-exploitation reduction program to restore the groundwater resources and alleviate associated environmental problems in the NCP (Xu, 2017; Zhao et al., 2017). As a result, the total amount of groundwater exploitation is expected to be reduced significantly. This reduction plan could involve the changes in operations of pumping wells at different times and locations. Inevitably, the basin-scale flow field could be modified by the alternation, intermittent shutdown, and resumption of pumping wells at different locations of the NCP. Such changes in flow fields are tantamount to a large-scale HT survey in the NCP. Because the reduction plan is a national policy that will be accomplished, it is an opportunity to address the issues of aquifer heterogeneity in the NCP by utilizing long-term datasets of aquifer responses from existing wells. However, such a long-term dataset in the NCP is not available for us at this moment. Consequently, in this study, we use a synthetic two-dimensional, horizontal, confined aquifer to demonstrate the feasibility of the proposed approach.

Previous studies on synthetic HT surveys at the NCP have proven that the feasibility of utilizing head data induced by groundwater exploitation reduction to enhance basin-scale aquifer characterization (Liu et al., 2020a), as well as the potential of HT in identifying the boundary conditions of groundwater basins (Daranond, Yeh, Hao, Wen, & Wang, 2020; Liu et al., 2020b). However, these studies did not consider the temporal sampling issues and the role of flux data for subsurface characterization. For this reason, the main objectives of this paper are as follows. (1) demonstrate the advantages of the hydraulic tomographic survey over the multiple simultaneous pumping tests for steady-state and transient-state inversions; (2) explore the temporal sampling issues and propose a reasonable strategy of data selection in the time domain for HT analysis; (3) evaluate the performance of T and S estimates using head and flux joint inversion. The findings of this study may guide the design of HT surveys and effectively collect appropriate field data in time during the reduction plan for basin-scale aquifer characterization.

2. Methods

2.1. Groundwater Flow in Two-dimensional Saturated Media

In this study, we assume that the following governing flow equation can describe the groundwater flow induced by pumping in a two-dimensional, depth-averaged, saturated, heterogeneous confined aquifer:

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

subject to the boundary and initial conditions:

$$\mathrm{H}\Big|_{\Gamma_1} = \mathrm{H}_1, [-\mathrm{T}(\mathbf{x})\nabla \mathrm{H}] \cdot \mathbf{n}\Big|_{\Gamma_2} = \mathrm{q}, \text{and } \mathrm{H}\Big|_{\mathrm{t}=0} = \mathrm{H}_0 \tag{2}$$

where, in equation (1), H is the total head [L], **x** is the spatial coordinate (**x**= (x, y), [L]), Q($\mathbf{x}_p$) is the pumping rate per unit area (L/T) at the location $\mathbf{x}_p$, T(**x**) is the transmissivity [L^2/T], and S(**x**) is the storativity [-]. Additionally, the right-hand term of Equation (1) becomes zero when it describes steady-state groundwater flow. In Equation (2), H_1 is the prescribed total head at Dirichlet boundary Γ_1, q is the specified flux at Neumann boundary Γ_2, **n** is a unit vector normal to Γ_2, and H_0 represents the initial head before stressing the aquifer.

In this study, these governing equations under steady and transient conditions are solved by the code of VSAFT2 (Variably Saturated Flow and Transport 2-D) developed by Yeh et al. (1993) available at http://tian.hwr.arizona.edu/downloads. The solution yields spatial-temporal variations of total head and flux, which are used to estimate the spatial distributions of hydraulic properties (*T* and *S*) by an inverse algorithm described below.

2.2. HT Analysis

In this study, steady and transient HT experiments were performed using the Simultaneous Successive Linear Estimator (SimSLE) developed by Xiang et al. (2009). The SimSLE algorithm is an extension of the Sequential Successive Linear Estimator

(SSLE) (Zhu & Yeh, 2005). In the SSLE method, the head information from discrete sources is sequentially assimilated into the inversion process. The SimSLE algorithm has some advantages over the SSLE. (1) SimSLE needs to evaluate the adjoint state equation only once for a given observation location using new parameter estimates from all pumping tests since the adjoint state equation is independent of the pumping rate and pumping location. However, SSLE needs to solve the adjoint state equation for each pumping test because the parameters in the adjoint state equation are modified for each pumping test. (2) SimSLE avoids the loop iteration of SSLE, and thus improves computational efficiency. (3) adding data in different sequences in SSLE may lead to a slightly different final result, but SimSLE does not have such a problem. (4) SimSLE simultaneously incorporates all the aquifer signals from multiple pumping tests during an HT survey in the estimation of hydraulic properties, providing more constraints for the inverse problem and thus converges faster than SSLE. However, the simultaneous inclusion of all data makes SimSLE more computationally intensive than SSLE, as the covariance matrix will become very large. A brief description of the SimSLE algorithm is given below.

In this study, the natural logarithms of hydraulic parameters (T and S) for each geological zone are treated as random variables described by a prior joint probability distribution, characterized by the mean, variance, and correlation scale. Likewise, the hydraulic head (H) and the magnitude of Darcian flux (Q) are also treated as spatial stochastic processes, expressed as the sum of the unconditional mean and the unconditional perturbation (i.e., $\mathrm{H} = \bar{H} + h$, $\mathrm{Q} = \bar{Q} + q$).

The SimSLE starts with the successive linear estimator (Yeh, Jin, & Hanna, 1996) to estimate the conditional expectation of the aquifer parameters (T and S) conditioned on the observed data. The covariance function in the SLE is calculated by using an exponential covariance model, and the adjoint-state method calculates the sensitivity matrix.

Then, the conditional mean estimates f are successively improved by the weighted differences between the observed and the simulated values:

$$f^{(r+1)} = f^{(r)} + w^{(r)T}(h_{obs} - h^{(r)}) \tag{3}$$

where $f^{(r)}$ is the conditional expectation of f at iteration r, h_{obs} is an $m \times 1$ observed data vector composed of m_h head data values and m_q flux data sets (where $m = m_h + m_q$), and $h^{(r)}$ is the simulated data values obtained from the forward model with the estimated hydraulic parameters at iteration r. $w^{(r)}$ is the weight matrix calculated using the conditional auto-covariance of observed data and the cross-covariance between parameter and observed data. The superscript T represents the transpose. The iterative process stops if one of the following two criteria is met. The change in variance of the estimated parameter between the current and previous iterations is smaller than a specified tolerance. The other is that the change of simulated head or flux values between successive iterations is smaller than a user-specified tolerance. Details of SLE are available in Yeh et al. (1996), Xiang et al. (2009), and Zha et al. (2014).

To assess the performance of inversion results from all the cases, we use the coefficient of determination (R^2), the mean square error (L_2), the slope, and the intercept in this study. Generally, the estimates are considered better, if R^2 value and the slope of the linear regression line are closer to 1. Likewise, the smaller L_2 value and the value of intercept approaching 0, signify that the estimated parameter fields are much closer to the reference fields. Additionally, visualization of the posterior uncertainty is also employed as another approach to evaluating the role of flux measurements in parameter estimations.

3. Numerical Experiments

3.1. Site Description

The Heilonggang (HLG) plain in the NCP is a pilot area of groundwater-overexploitation control, covering a total area of approximately 2695 km^2 (Fig. 1). The elevation in the southwest is generally higher than that in the northeast part. The aquifers in the NCP are an aquifer-aquitard system (Cao et al., 2013). The unconfined and the confined aquifers are separated by an intermittent aquitard composed of silt and clay with a thickness of around 10 m (Wei, 2018). The confined aquifer's T and S values vary from 10 to 480 m^2/d, and from 0.001 to 0.008, respectively (Wang, 2011).

The groundwater exploitation reduction program has reduced the total amount of groundwater pumping since 2014, via agricultural water-saving projects and water transfer from the middle route of the South-to-North Water Diversion Project (Zhao et al., 2017). The rebound of the water levels is expected to take place in some regions. Thus, the aquifer signals with operational changes of existing wells may provide a good opportunity for estimating the hydraulic heterogeneity over the whole basin.

3.2. Model Setup

The geometric shape of the synthetic aquifer is identical to the geometry of the real groundwater basin. This basin was discretized into 2720 rectangular elements and 2896 nodes with element dimensions 1 km$\times$1 km. Since the groundwater flows predominantly from the west and southwest parts to the northeast boundary (Fig. 1), specified head boundaries were assigned to the A-B, C-D, and E-F segments with constant head values of 100 m, 100 m, and 60 m, respectively. The remaining boundary segments were prescribed as the impermeable boundaries. The synthetic well-field consists of 70 wells, and the well configuration is similar to those in the real basin. The numbers and locations of pumping and monitoring wells, and steady-state initial head distribution with no pumping are shown in Fig. 1.

Reference heterogeneous *T* and *S* (Fig. 2) were created with the four geological zones of the HLG basin (Fig. 1). Independent random fields of *T* and *S* are generated for each geological zone. In terms of actual geological data, different ranges of *T* are assigned to the zone 1 (120-240 m^2/d), zone 2 (<120 m^2/d), zone 3 (240-360 m^2/d), and zone 4 (360-480 m^2/d). Considering the range of transmissivity for each zone, we can assume that each zone is mildly heterogeneous with small variations. Therefore, the mean ln*T* (m^2/d) values 5.1, 4.38, 5.4, and 5.68 were assigned to zones 1, 2, 3, and 4, respectively. The variance of ln*T* for each zone is 0.1. Likewise, the mean ln*S* for each zone is -5.81, -6.91, -5.12, and -4.89, respectively, and the variance of ln*S* is 0.1. The correlation scales are 20 km in the east-west direction and 5 km in the north-south direction. Different random seeds were used to generate independent random fields of reference *T* and *S* for each geological zone (Fig. 2) using a random field generator (Gutjahr, 1989) embedded in VSAFT2 software.

For inverse modeling of all cases and scenarios in this work, the geometry and locations of the four geological zones in the reference fields are assumed known exactly (Fig. 1). For initial guesses, each zone is prescribed with a mean *T* or *S* value identical to that of the corresponding zone in the reference *T* or *S* fields. The true correlation scales are also assumed known. The details of steady-state and transient-state inversions will be explained in the following sections.

4. Results and Discussion

4.1. Steady-state Experiments

The steady-state experiments examined two cases. As shown in Table 1, Case A involves a simultaneous pumping at 18 wells (see Fig. 1) with a rate of 3000 m^3/d. Case B represents an HT survey with three pumping stresses, and each stress involves pumping six different wells at a rate of 3000 m^3/d. These wells are listed with parentheses in Table 1, and their locations are indicated in Fig. 1.

The simulated groundwater flow field in Case A shows that the simultaneous operation of multiple pumping wells only affected limited portions of the entire aquifer (Fig. 3a). Noticeably, the steady-state flow field is mainly altered by the pumping wells distributed in the southeast areas adjacent to the impermeable boundary (Fig. 3a). However, Case B yields three flow fields with significant differences (Figs. 3b, 3c, and 3d), equivalent to three snapshots of aquifer heterogeneity at different locations. Specifically, each stress produces a different flow (or head) field from the other stress, reflecting the effects of heterogeneity at different parts of the basin. As a result, each observation well collects heads affected by different heterogeneity during each stress. Because of this reason, more information about heterogeneity is recorded at each observation than that from simultaneous hydraulic tests at fixed pumping locations. Thus, interpretation of the heterogeneity based on the head collected from Case B could lead to more detailed heterogeneity.

Compared Fig. 4a to Fig. 4b, we see that the T tomogram from Case B is closer to the reference T field than Case A. For Case A with just one stress, while the main high and low permeable zones are identified, the estimated T pattern is smoother than the reference field, and the localized geometry is not accurately captured (Fig. 4a). For Case B with three stresses, the estimated T tomogram is greatly improved and close to the reference T field (Fig. 4b).

The scatterplots of reference ln$T$ versus estimated ln$T$ are shown in Figs. 4c and 4d. The performance metrics for HT inversion (Case B), with a higher R^2 (0.88), smaller L_2 (0.029), higher slope (0.89), and lower intercept (0.57), are much better than those metrics (R^2=0.82, L_2=0.043, slope=0.84, and intercept=0.84) from conventional simultaneous pumping tests represented by Case A.

In summary, the aquifer's responses to the change in pumping locations, carry additional information about the heterogeneity. This result corroborates the explanation in Wen et al. (2020).

4.2. Transient-state Inversion and Temporal Sampling

4.2.1. Transient HT Experiments

In transient HT experiments, the initial condition was the simulated steady-state responses of the aquifer with no pumping, and two scenarios were considered (Table 2). Scenario A considered one stress, consisting of three simultaneous pumping events. Each event occurred at a group of wells, over three periods (0-1000 days, 1001-2000 days, and 2001-3000 days) with pumping/reduction/shutdown activities. In Scenario B, three stresses and events identical to those in Scenario A were employed, but involved a different group of wells at different locations, operating at different pumping/reduction/shutdown patterns over the three periods.

For each scenario, we also examine three cases for the temporal sampling investigation (Table 3). Case 1 is the baseline case for the temporal sampling scheme where head data are collected at 52 observation wells with a time interval of 100 days, leading to 30 head data records for each observation well during each stress. In Case 2, ten head data per well at 5, 10, 15, 20, 30, 1000, 1500, 2000, 2500, and 3000 days are selected. To reduce the computational cost to one-third of Case 1, we select these data at early-mid-late times (i.e., 5, 10, 15, 20, 30, 1000, 1500, 2000, 2500, and 3000 days) of the well hydrographs. In Case 3, we further reduce the computational cost by selecting only five observation data per well (i.e., 1000, 1500, 2000, 2500, 3000 days), representing mid-late data of the well hydrographs.

4.2.2. Performance of Scenarios A and B

The tomograms and performance metrics from the baseline case of Scenarios A and B (designed as Scenario A1 and Scenario B1, respectively) are compared. The estimated fields of *T* and *S*, scatterplots as well as performance metrics are shown in Figs. 5 and 6.

The *T* field from Scenario B1 performs slightly better than Scenario A1 in the characterization of hydraulic connections (Figs. 5a and 5b). The scatterplots of ln*T* in both cases also provide further evidence (Figs. 6a and 6b). Scenario A1 captures the dominant low and high *T* zones of this basin-scale aquifer with good evaluation metrics (R^2=0.87, L_2=0.030, slope=0.86, and intercept=0.73), while the scatter points from Scenario B1 are closer to the 1:1 line with a little improvement in R^2 (0.88), L_2 (0.027), slope (0.89), and intercept (0.57). Additionally, the transient-state inversion outperforms the steady-state inversion in *T* estimates by comparing the performance metrics between them (Figs. 4c, 4d, 6a, and 6b).

Regarding *S* estimates, the *S* tomogram from Scenario A1 is comparable to the reference *S* field. However, the geometric shapes of the high and low *S* regions are not accurately captured and present a localized smooth appearance (Fig. 5c). Scenario B1 yields a more detailed aquifer heterogeneity (Fig. 5d) than Scenario A1, relative to the reference *S* field (Fig. 2b). The comparison of estimated ln*S* versus the reference ln*S* also demonstrates that Scenario B1 presents a better performance for *S* estimation than Scenario A1 (Figs. 6c and 6d). The values of R^2 and slope increase from 0.91 and 0.92 in Scenario A1 to 0.93 and 0.93 in Scenario B1, respectively. While the L_2 and the intercept also reduce from 0.043 and -0.46 in Scenario A1 to 0.035 and -0.38 in Scenario B1.

Overall, the estimated *T* and *S* tomograms from the HT survey (Scenario B) are better than those from multiple simultaneous pumping tests (Scenario A). This result confirms the hypothesis that aquifer responses to changes in operations and locations of pumping wells, as an equivalent basin-scale HT survey, carry new information on aquifer heterogeneity. Moreover, it is unnecessary to conduct the pumping test to reach a steady-state, and we recommend that transient data be collected for HT inversion in field hydraulic tests.

4.2.3. Comparison among Temporal Sampling Schemes

The estimated *T* fields from Scenarios B2 and B3 (Figs. 7b and 7d) are closer to the reference fields than those tomograms from Scenarios A2 and A3 (Figs. 7a and 7c), while R^2, L_2, slope and intercept also indicate the improvement of *T* estimates from Scenarios B2 and B3 (Figs. 8b and 8d) in comparison with Scenarios A2 and A3 (Figs. 8a and 8c). Plots of the *S* tomograms and evaluation metrics from Scenarios B2 and B3 (Figs. 7f, 7h, 8f, and 8h) also show better *S* estimates than Scenarios A2 and A3 (Figs. 7e, 7g, 8e, and 8g). Therefore, it is clear that tomographic survey leads to better *T* and *S* estimates than multiple simultaneous pumping tests at the same locations, consistent with the previous results of baseline analysis.

Addtionally, the inversion results from different temporal sampling schemes show that Case 2 yields the best-estimated fields of *T* and *S*, attributing to the optimal selection of head data points controlling the early-mid-late curves of well hydrographs for HT analysis. This finding also indicates that such head data used for HT inversion in Case 2 carry sufficient information to characterize this basin-scale aquifer heterogeneity accurately.

Although the estimated fields (Fig. 7) from Case 3 are comparable to those reference fields, the metrics (R^2, L_2, slope, and intercept) are inferior to those from Case 2 (Fig. 8). The temporal sampling scheme of Case 3, covering only the mid-late curves of hydrographs, can delineate the general patterns of predominant heterogeneity in *T* and *S*. It reduces the computational cost at the expense of some heterogeneous information.

In short, we recommend that HT analysis should use head data at early-mid-late portions of well hydrographs for parameter estimation to maximize the power of HT and to save high computational costs. These results are consistent with findings by Sun et al. (2013).

4.3. Role of Flux Data in Parameter Estimations

The effect of flux data on parameter estimates is examined in this section. Given that Case 2 is the optimal temporal sampling scheme and is efficient in computational cost, head and flux data at ten times (i.e., the sampling strategy in Case 2) in 52 observation wells generated from transient-state forward simulations are included in HT inversion. The estimated *T* and *S* fields from the joint inversion show further local refinement (Fig. 9) compared to those from head inversion (Figs. 7a, 7b, 7e, and 7f). Visually, head and flux inversion in Scenario B2 yields the estimated *T* and *S* tomograms (Figs. 9b and 9d) that resemble the reference fields (Fig. 2) the most, also as reflected in the quantitative metrics in the scatterplots (Fig. 10).

For Scenario A (simultaneous pumping), the addition of flux data contributes to the improvements in *T* and *S* estimates. Specifically, relative to head inversion results (R^2=0.87, L_2=0.03, slope=0.87, intercept=0.67) (Fig. 8a), the performance of *T* estimation (Fig. 10a) from the joint interpretation of both head and flux data is better with higher R^2 (0.90) and slope (0.91) as well as lower L_2 (0.023) and intercept (0.47). While the performance of *S* estimation improves slightly, the R^2 and the slope increase from 0.91 to 0.92, and from 0.91 to 0.92, respectively. The L_2 and the intercept also drop from 0.044 to 0.041 and from -0.51 to -0.42, respectively (Figs. 8e and 10c).

However, for Scenario B (HT), flux data appear to have negligible effects on HT inversion results, using head data only. Specifically, there is no significant improvement in the performance metrics from the joint inversion of both head and flux data (Figs. 10b and 10d), relative to those from the head only inversion (Figs. 8b and 8f). The R^2 value remains unchanged for *T* estimates (0.90) and *S* estimates (0.93). The slope increases from 0.91 to 0.93 and from 0.93 to 0.95 for *T* and *S* estimates. While the intercept also drops from 0.48 to 0.37 and from -0.37 to -0.27 for *T* and *S* estimates, respectively.

Additionally, the uncertainties in *T* and *S* estimates from head and flux joint inversion are compared with those from head inversion (Figs. 11 and 12). We present the uncertainty map of the estimate in addition to the mean estimate, because visualization of uncertainty maps would be informative and advantageous for identifying the role of flux measurements.

According to the uncertainty maps of *T* estimates (Fig. 11), the inter-well areas show low residual variances (uncertainty) of ln*T* while the high variances are mainly in the vicinity of constant head boundaries. Comparing to the residual variances from head inversion in Scenario A2 (Fig. 11a), we observe that the areas with small uncertainty from the joint inversion increase significantly (Fig. 11b), indicating that flux conditioning improves the *T* estimates within the well field for short-term pumping tests in Scenario A. However, the addition of flux data in Scenario B2 seems to have minor changes in the uncertainty of *T* estimates (Figs. 11c and 11d). Only localized improvements in the original low variances zones are observed. So the effect of flux data on *T* estimation appears to be less prominent for Scenario B than Scenario A, supporting the results mentioned above of performance metrics. From another perspective, no matter head inversion or the joint inversion of head and flux, HT surveys

(Scenario B) significantly reduce the uncertainty in *T* estimates (Figs. 11c and 11d) compared to conventional pumping tests (Scenario A) (Figs. 11a and 11b).

By contrast, the effect of additional flux data on *S* estimates is less appreciable than *T* estimates for Scenarios A2 and B2 (Fig. 12). High residual variances of *S* estimates remain along the head boundaries, which may be explained by the non-uniqueness of inversion near the domain boundaries (Yeh et al., 2015). For Scenario A2, flux measurements at the same monitoring locations slightly improve the *S* estimates and reduce the uncertainty in some areas (Figs. 12a and 12b). Additionally, for Scenario B2, a few changes in uncertainty maps between head inversion and joint inversion are visible (Figs. 12c and 12d). Such localized reduction of uncertainty in *S* estimates may contribute to improvements in some metrics (slope and intercept), as illustrated in the scatterplots. On the whole, flux measurements do not significantly enhance mapping of aquifer heterogeneity for long-term HT inversion (Scenario B), which may be ascribed to less room for improvement in *S* estimates considering the excellent performance of head inversion (Fig. 8f).

We emphasize that the geologic zonation model is assumed known, and each zone has low variances of $\ln T$ and $\ln S$ (0.1) in the experiments. Due to low variability in each zone, the benefits of specifying accurate mean values of *T* and *S* outweigh the benefits of flux information. As a result, when the correct zonal *T* and *S* mean values are the initial guesses of inverse models, the improvements due to additional flux data on the estimates are less prominent, substantiating Tso et al. (2016). However, the importance of the inclusion of flux measurements is vivid. Compared with head data, flux measurements carry additional nonredundant information on aquifer heterogeneity, highly pertinent to the connectivity between the pumping wells and the observation wells. This finding corroborates the work by Tso et al. (2016) about the crosscorrelation between flux and hydraulic conductivity. More importantly, when the distributed mean properties are known, the flux data play a more important role in the short-term pumping tests (Scenario A) than the long-term HT survey (Scenario B).

Because flux data contribute more to the improvement of parameter inversion for Scenario A than Scenario B, if reliable geological data are available, we encourage that head and flux data be jointly collected for conventional short-term pumping tests, while the role of flux data is insignificant for the long-term HT surveys in such porous-media aquifers without fissures or fractures. However, the inversion of the zonation model based on inaccurate geological information may yield worse inversion results (Zhao & Illman, 2018; Zhao et al., 2016). Thus, the influence of inaccurate geological data on inversion results remains to be examined in future studies.

Unlike hydraulic head observation, the measurement of groundwater flux at an observation well is far from routine for groundwater monitoring works. Nonetheless, there is a growing interest in flux measurements in a groundwater monitoring network, given that flux measurements can enhance the mapping of fracture distributions in geologic media (Tso et al., 2016; Zha et al., 2014). Furthermore, it plays a vital role in predicting contaminant migration for groundwater remediation (Kuhlman et al., 2008; Liu et al., 2020a; Ni, Yeh, & Chen, 2009). Among numerous borehole methods for groundwater flux measurements, the point dilution method (Drost et al., 1968) is well described in many books, while heat (Melville, Molz, & Güven, 1985) or tracer (Palmer, 1993) is an alternative approach. In addition, electronic borehole flowmeter profiling (Young & Pearson, 1995) is widely used to address flow variations along a borehole.

In unconsolidated geologic media, much attention is paid to field in-situ methods, because they are free from borehole effects and are superior to borehole methods in determining both magnitude and direction of groundwater flux (Cremeans, Devlin, McKnight, & Bjerg, 2018; Essouayed, Annable, Momtbrun, & Atteia, 2019; Osorno, Devlin, & Firdous, 2018; Thomle, Strickland, Johnson, Zhu, & Stegen, 2020).

In practice, head measurements at many fully screened wells only represent depth-averaged head values, and they do not carry important information about vertical aquifer heterogeneity (Li et al., 2008). Flux measurements along the well screen during HT surveys may overcome the limitation of the depth-averaged head measurements at full-screen observation wells by incorporating information about vertical distribution of hydraulic conductivity (Tso et al., 2016; Zha et al., 2014). Although this work discussed the role of flux measurements for subsurface characterization in a groundwater basin where geological information is well known, further investigations remain on the inclusion of flux data into inverse models for regions that lack geological data.

4.4. Implications on NCP Groundwater Management Practices

Implementing the groundwater exploitation reduction program in the NCP provides an excellent opportunity for characterizing the subsurface hydraulic heterogeneity over the basin. It is the time to collect the data of regional groundwater exploitation intelligently, requiring multi-agencies collaboration and data sharing. This study suggests that data collection should include the spatial locations and screen intervals of existing pumping and observation wells, the time-variant pumping rates, the observed well hydrographs, geological data, etc. Moreover, we must take advantage of the well hydrographs and pumping operations data in the NCP, accumulated over many years. So we recognize that it is a realistic and feasible way to resolve the issues of aquifer heterogeneity in the NCP.

5. Conclusions

A cost-effective, large-scale HT survey for groundwater basins is feasible by changing the operations of existing pumping wells. Specifically, the operational variations of pumping wells at different locations cause the overlap between the stressed areas of different pumping events. The head fluctuations in the monitoring wells respond differently to the variations of well operation, which results in the time-variant groundwater flow fields. The disparity of regional head distributions suggests that well hydrographs likely carry new information about the hydraulic connections among pumping and monitoring wells at various locations. This basin-scale HT survey saves a significant amount of investment by taking full advantage of aquifer responses to the changes in operations of existing pumping wells.

The results of both steady-state and transient-state inversions demonstrate the advantage of HT surveys in characterizing basin-scale heterogeneity over conventional pumping tests at fixed locations. Information during HT surveys carries more non-redundant information on the hydraulic heterogeneity than those from multiple simultaneous pumping tests. Additionally, the transient-state inversion outperforms the steady-state inversion in T estimates. For this reason, we recommend that transient data in field tests should be collected for HT inversion.

We propose an optimal temporal sampling strategy for basin-scale HT analysis. Head data at the early, intermediate and late-time from well hydrographs should be selected for the HT analysis to maximize the power of HT and save high computational costs. Such an optimal temporal sampling scheme can yield the best-estimated fields of T and S in this basin-scale aquifer. The use of the mid-late data of hydrographs captures the general patterns of predominant heterogeneity and reduces the computational expenses, but at the cost of some heterogeneous information.

For conventional pumping tests, the addition of flux data contributes to significant improvements in T and S estimates. However, for HT surveys, flux data appear to have little effect on HT inversion results. There is no significant improvement in the performance metrics from the head and flux joint inversion. Due to low variability in each geological zone, the use of distributed zonal mean values as initial guesses generates good *T* and *S* estimates from head inversion, which leaves less room for improvement in estimates from additional flux information. Therefore, if geological knowledge is available in the real world, we encourage that head and flux data be jointly collected for conventional short-term pumping tests. In contrast, flux measurements are insignificant for long-term HT surveys in such basin-scale porous-media aquifers without fissures or fractures.

Admittedly, there are limitations inherent in this study. Firstly, a two-dimensional synthetic numerical experiment was examined without considering other unknown sources (such as leakage, etc.). Secondly, this study uses the known geologic zonations as prior information of inverse models for all scenarios and cases. This study also assumes that boundary conditions for inverse modeling are perfectly known, reducing the uncertainty of inverse solutions. Besides, only one possible hypothetic basin was investigated. The inverse problem is not well-defined. Such as, it could have many possible solutions. Our results should be treated as one possibility. Unless a Monte Carlo simulation (Wang et al., 2021) with many realizations is conducted, our results may vary.

Despite these limitations, the results of this study demonstrate the superiority of HT survey over the multiple simultaneous pumping tests and advance the knowledge of the effective selection of informative monitoring data in the time domain and the role of flux measurements in parameter estimations. The basin-scale tomographic survey is a step forward in aquifer characterization, which could optimize the design of HT surveys and guide field data collection for basin-scale subsurface characterization as championed by Yeh and Lee (2008) and Yeh et al. (2008).

Acknowledgements

This research was supported by the National Natural Science Foundation of China (Grant No. 41901039); the Natural Science Foundation of Hebei Province (Grant No. D2019402045); the Department of Education in Hebei Province (Grant No. BJ2021014). T.-C. Jim Yeh also acknowledges the partial support by the U.S. NSF grant EAR1931756. Special thanks go to the editor and the reviewers for their constructive comments.

Conflict of Interest：None.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

Berg, S. J., & Illman, W. A. (2011). Three-dimensional transient hydraulic tomography in a highly heterogeneous glaciofluvial aquifer-aquitard system. *Water Resources Research*, *47*(10), W10507. https://doi.org/10.1029/2011wr010616

Bohling, G. C., Zhan, X., Butler, J. J., & Zheng, L. (2002). Steady shape analysis of tomographic pumping tests for characterization of aquifer heterogeneities. *Water Resources Research*, *38*(12), 60-61-60-15. https://doi.org/10.1029/2001wr001176

Butler, J. J., McElwee, C. D., & Bohling, G. C. (1999). Pumping tests in networks of multilevel sampling wells: Motivation and methodology. *Water Resources Research*, *35*(11), 3553-3560. https://doi.org/10.1029/1999wr900231

Cao, G., Zheng, C., Scanlon, B. R., Liu, J., & Li, W. (2013). Use of flow modeling to assess sustainability of groundwater resources in the North China Plain. *Water*

Resources Research, *49*(1), 159-175. https://doi.org/10.1029/2012wr011899

Cardiff, M., Bakhos, T., Kitanidis, P. K., & Barrash, W. (2013). Aquifer heterogeneity characterization with oscillatory pumping: Sensitivity analysis and imaging potential. *Water Resources Research*, *49*(9), 5395-5410. https://doi.org/10.1002/wrcr.20356

Cardiff, M., Barrash, W., & Kitanidis, P. K. (2012). A field proof-of-concept of aquifer imaging using 3-D transient hydraulic tomography with modular, temporarily-emplaced equipment. *Water Resources Research*, *48*(5), W05531. https://doi.org/10.1029/2011wr011704

Cardiff, M., Barrash, W., & Kitanidis, P. K. (2013). Hydraulic conductivity imaging from 3-D transient hydraulic tomography at several pumping/observation densities. *Water Resources Research*, *49*(11), 7311-7326. https://doi.org/10.1002/wrcr.20519

Cardiff, M., Barrash, W., Kitanidis, P. K., Malama, B., Revil, A., Straface, S., & Rizzo, E. (2009). A potential-based inversion of unconfined steady-state hydraulic tomography. *Ground Water*, *47*(2), 259-270. https://doi.org/10.1111/j.1745-6584.2008.00541.x

Cho, M. S., Zhao, Z., Thomson, N. R., & Illman, W. A. (2020). Use of steady-state hydraulic tomography to inform the selection of a chaotic advection system. *Journal of Contaminant Hydrology*, *229*, 103559. https://doi.org/10.1016/j.jconhyd.2019.103559

Cremeans, M. M., Devlin, J. F., McKnight, U. S., & Bjerg, P. L. (2018). Application of new point measurement device to quantify groundwater-surface water interactions. *Journal of Contaminant Hydrology*, *211*, 85-93. https://doi.org/10.1016/j.jconhyd.2018.03.010

Daranond, K., Yeh, T.-C. J., Hao, Y., Wen, J.-C., & Wang, W. (2020). Identification of groundwater basin shape and boundary using hydraulic tomography. *Journal of Hydrology, 588*, 125099. https://doi.org/10.1016/j.jhydrol.2020.125099

Drost, W., Klotz, D., Koch, A., Moser, H., Neumaier, F., & Rauert, W. (1968). Point dilution methods of investigating ground water flow by means of radioisotopes. *Water Resources Research*, *4*(1), 125-146. https://doi.org/10.1029/WR004i001p00125

Essouayed, E., Annable, M. D., Momtbrun, M., & Atteia, O. (2019). An innovative tool for groundwater velocity measurement compared with other tools in laboratory and field tests. *Journal of Hydrology X*, *2*, 100008. https://doi.org/10.1016/j.hydroa.2018.100008

Fischer, P., Jardani, A., Soueid Ahmed, A., Abbas, M., Wang, X., Jourde, H., & Lecoq, N. (2017). Application of Large-Scale Inversion Algorithms to Hydraulic Tomography in an Alluvial Aquifer. *Ground Water*, *55*(2), 208-218. https://doi.org/10.1111/gwat.12457

Gutjahr, A. L. (1989). *Fast Fourier transforms for random field generation: project report for Los Alamos Grant to New Mexico Tech* New Mexico Institute of Mining and Technology]. Socorro.

Hao, Y., Yeh, T. C., Xiang, J., Illman, W. A., Ando, K., Hsu, K. C., & Lee, C. H. (2008). Hydraulic tomography for detecting fracture zone connectivity. *Ground Water*, *46*(2), 183-192. https://doi.org/10.1111/j.1745-6584.2007.00388.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. *Environmental Science & Technology*, *44*(22), 8609-8614.

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

Illman, W. A., Liu, X., & Craig, A. (2007). Steady-state hydraulic tomography in a laboratory aquifer with deterministic heterogeneity: Multi-method and multiscale validation of hydraulic conductivity tomograms. *Journal of Hydrology*, *341*(3-4), 222-234. https://doi.org/10.1016/j.jhydrol.2007.05.011

Jiménez, S. (2015). *High resolution aquifer characterization using hydraulic tomography and tracer tomography* ETH Zurich].

Kitanidis, P. K., & Lee, J. (2014). Principal Component Geostatistical Approach for large-dimensional inverse problems. *Water Resources Research*, *50*(7), 5428-5443. https://doi.org/10.1002/2013wr014630

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

Lee, J., & Kitanidis, P. K. (2014). Large-scale hydraulic tomography and joint inversion of head and tracer data using the Principal Component Geostatistical Approach (PCGA). *Water Resources Research*, *50*(7), 5410-5427. https://doi.org/10.1002/2014wr015483

Lee, J., Yoon, H., Kitanidis, P. K., Werth, C. J., & Valocchi, A. J. (2016). Scalable subsurface inverse modeling of huge data sets with an application to tracer concentration breakthrough data from magnetic resonance imaging. *Water Resources Research*, *52*(7), 5213-5231. https://doi.org/10.1002/2015wr018483

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

Lin, Y., Le, E. B., O'Malley, D., Vesselinov, V. V., & Bui-Thanh, T. (2017). Large-scale inverse model analyses employing fast randomized data reduction. *Water Resources Research*, *53*(8), 6784-6801. https://doi.org/10.1002/2016wr020299

Liu, F., Yeh, T.-C. J., Wang, Y.-L., Hao, Y., Wen, J.-C., & Wang, W. (2020a). Characterization of basin-scale aquifer heterogeneity using transient hydraulic tomography with aquifer responses induced by groundwater exploitation reduction. *Journal of Hydrology*, *588*, 125137. https://doi.org/10.1016/j.jhydrol.2020.125137

Liu, F., Yeh, T. C. J., Wang, Y. L., Song, X., Lei, X., Wen, J. C., Wang, W., & Hao, Y. (2020b). Potential of Hydraulic Tomography in Identifying Boundary Conditions of Groundwater Basins. *Water Resources Research*, *56*(12), e2020WR028331. https://doi.org/10.1029/2020wr028331

Liu, X., Illman, W. A., Craig, A. J., Zhu, J., & Yeh, T. C. J. (2007). Laboratory sandbox validation of transient hydraulic tomography. *Water Resources Research*, *43*(5). https://doi.org/10.1029/2006wr005144

Liu, X., Zhou, Q., Birkholzer, J., & Illman, W. A. (2013). Geostatistical reduced-order models in underdetermined inverse problems. *Water Resources Research*, *49*(10), 6587-6600. https://doi.org/10.1002/wrcr.20489

Luo, N., Illman, W. A., Zha, Y., Park, Y.-J., & Berg, S. J. (2020). Three-dimensional hydraulic tomography analysis of long-term municipal wellfield operations: Validation with synthetic flow and solute transport data. *Journal of Hydrology*, *590*, 125438. https://doi.org/10.1016/j.jhydrol.2020.125438

Luo, N., Zhao, Z., Illman, W. A., & Berg, S. J. (2017). Comparative study of transient hydraulic tomography with varying parameterizations and zonations: Laboratory sandbox investigation. *Journal of Hydrology*, *554*, 758-779.

https://doi.org/10.1016/j.jhydrol.2017.09.045

Melville, J. G., Molz, F. J., & Güven, O. (1985). Laboratory Investigation and Analysis of a Ground-Water Flowmeter. *Groundwater*, *23*(4), 486-495. https://doi.org/10.1111/j.1745-6584.1985.tb01498.x

Mohammadi, Z., & Illman, W. A. (2019). Detection of karst conduit patterns via hydraulic tomography: A synthetic inverse modeling study. *Journal of Hydrology*, *572*, 131-147. https://doi.org/10.1016/j.jhydrol.2019.02.044

Ni, C.-F., Yeh, T.-C. J., & Chen, J.-S. (2009). Cost-effective hydraulic tomography surveys for predicting flow and transport in heterogeneous aquifers. *Environmental Science & Technology*, *43*(10), 3720-3727.

Osorno, T. C., Devlin, J. F., & Firdous, R. (2018). An In-Well Point Velocity Probe for the rapid determination of groundwater velocity at the centimeter-scale. *Journal of Hydrology*, *557*, 539-546. https://doi.org/10.1016/j.jhydrol.2017.12.033

Palmer, C. D. (1993). Borehole dilution tests in the vicinity of an extraction well. *Journal of Hydrology*, *146*, 245-266. https://doi.org/10.1016/0022-1694(93)90279-I

Sánchez-León, E., Erdal, D., Leven, C., & Cirpka, O. A. (2020). Comparison of Two Ensemble Kalman-Based Methods for Estimating Aquifer Parameters from Virtual 2-D Hydraulic and Tracer Tomographic Tests. *Geosciences*, *10*(7), 276. https://doi.org/10.3390/geosciences10070276

Soueid Ahmed, A., Zhou, J., Jardani, A., Revil, A., & Dupont, J. P. (2015). Image-guided inversion in steady-state hydraulic tomography. *Advances in Water Resources*, *82*, 83-97. https://doi.org/10.1016/j.advwatres.2015.04.001

Sun, R., Yeh, T.-C. J., Mao, D., Jin, M., Lu, W., & Hao, Y. (2013). A temporal sampling strategy for hydraulic tomography analysis. *Water Resources Research*, *49*(7), 3881-3896. https://doi.org/10.1002/wrcr.20337

Thomle, J., Strickland, C., Johnson, T. C., Zhu, Y., & Stegen, J. (2020). A Flux Detection Probe to Quantify Dynamic Groundwater-Surface Water Exchange in the Hyporheic Zone. *Ground Water*. https://doi.org/10.1111/gwat.13001

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 Resources Research*, *52*(1), 3-20. https://doi.org/10.1002/2015wr017191

Wang, B. (2011). *Study on Comprehensive Quality Assessment and Coupling-Coordination Relationship of Water-Land Resources in Heilonggang District* Chinese Academy of Geological Sciences. (in Chinese)].

Wang, Y.-L., Yeh, T.-C. J., Wen, J.-C., Huang, S.-Y., Zha, Y., Tsai, J.-P., Hao, Y., & Liang, Y. (2017). Characterizing subsurface hydraulic heterogeneity of alluvial fan using riverstage fluctuations. *Journal of Hydrology*, *547*, 650-663. https://doi.org/10.1016/j.jhydrol.2017.02.032

Wang, Y.-L., Yeh, T.-C. J., Xu, D., Li, K., Wen, J.-C., Huang, S.-Y., Wang, W., & Hao, Y. (2021). Stochastic analysis of oscillatory hydraulic tomography. *Journal of Hydrology*, *596*, 126105. https://doi.org/10.1016/j.jhydrol.2021.126105

Wei, R. (2018). *Evaluation of Groundwater Resources in Plain Area of Zhangwei River Basin and Relationship between Surface Water and Groundwater Conversion in Typical Section* China University of Geosciences (Beijing).]. Beijing. (in Chinese).

Wen, J.-C., Chen, J.-L., Yeh, T.-C. J., Wang, Y.-L., Huang, S.-Y., Tian, Z., & Yu, C.-Y. (2020). Redundant and Nonredundant Information for Model Calibration or Hydraulic Tomography. *Groundwater*, *58*(1), 79-92. https://doi.org/10.1111/gwat.12879

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 Resources Research*, *45*(2), W02432. https://doi.org/10.1029/2008wr007180

Xu, C. (2017). *Study on the dynamic variation of groundwater level and its influence factors in Guangping County* China University of Geosciences (Beijing).]. Beijing. (in Chinese).

Yeh, T.-C. J., Jin, M., & Hanna, S. (1996). An Iterative Stochastic Inverse Method: Conditional Effective Transmissivity and Hydraulic Head Fields. *Water Resources Research*, *32*(1), 85–92.

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 Resources Research*, *44*(3). https://doi.org/10.1029/2007wr006375

Yeh, T.-C. J., Lee, C.-H., Hsu, K.-C., & Wen, J.-C. (2008). Fusion of hydrologic and geophysical tomographic surveys. *Geosciences Journal*, *12*(2), 159-167. https://doi.org/10.1007/s12303-008-0017-6

Yeh, T.-C. J., & Liu, S. (2000). Hydraulic tomography: Development of a new aquifer test method. *Water Resources Research*, *36*(8), 2095-2105. https://doi.org/10.1029/2000wr900114

Yeh, T.-C. J., Mao, D.-q., Zha, Y.-y., Wen, J.-c., Wan, L., Hsu, K.-c., & Lee, C.-h. (2015). Uniqueness, scale, and resolution issues in groundwater model parameter identification. *Water Science and Engineering*, *8*(3), 175-194. https://doi.org/10.1016/j.wse.2015.08.002

Yeh, T.-C. J., Srivastava, R., Guzman, A., & Harter, T. (1993). A Numerical Model for

Water Flow and Chemical Transport in Variably Saturated Porous Media. *Groundwater*, *31*(4), 634-644. https://doi.org/10.1111/j.1745-6584.1993.tb00597.x

Yeh, T.-C. J., Xiang, J., Suribhatla, R. M., Hsu, K.-C., Lee, C.-H., & Wen, J.-C. (2009). River stage tomography: A new approach for characterizing groundwater basins. *Water Resources Research*, *45*(5), W05409. https://doi.org/10.1029/2008wr007233

Young, S. C., & Pearson, H. S. (1995). The Electromagnetic Borehole Flowmeter: Description and Application. *Groundwater Monitoring & Remediation*, *15*(4), 138-147. https://doi.org/10.1111/j.1745-6592.1995.tb00561.x

Zha, Y., Yeh, T.-C. J., Illman, W. A., Mok, C. M. W., Tso, C.-H. M., Carrera, B. A., & Wang, Y.-L. (2019). Exploitation of pump-and-treat remediation systems for characterization of hydraulic heterogeneity. *Journal of Hydrology*, *573*, 324-340. https://doi.org/10.1016/j.jhydrol.2019.03.089

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. *Advances in Water Resources*, *71*, 162-176. https://doi.org/10.1016/j.advwatres.2014.06.008

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 Resources Research*, *54*(3), 1616-1632. https://doi.org/10.1002/2017wr021884

Zha, Y., Yeh, T. 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. *Ground Water*, *54*(6), 793-804. https://doi.org/10.1111/gwat.12421

Zhao, Y., & Luo, J. (2020). Reformulation of Bayesian Geostatistical Approach on Principal Components. *Water Resources Research*, *56*(4), e2019WR026732. https://doi.org/10.1029/2019wr026732

Zhao, Y., Zhu, Y., Lin, Z., Wang, J., He, G., Li, H., Li, L., Wang, H., Jiang, S., He, F., Zhai, J., Wang, L., & Wang, Q. (2017). Energy Reduction Effect of the South-to-North Water Diversion Project in China. *Scientific Reports*, *7*(1), 15956. https://doi.org/10.1038/s41598-017-16157-z

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. *Journal of Hydrology*, *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. *Journal of Hydrology*, *559*, 392-410. https://doi.org/10.1016/j.jhydrol.2018.02.024

Zhao, Z., Illman, W. A., & Berg, S. J. (2016). On the importance of geological data for hydraulic tomography analysis: laboratory sandbox study. *Journal of Hydrology*, *542*, 156-171. https://doi.org/10.1016/j.jhydrol.2016.08.061

Zhao, Z., Illman, W. A., Zha, Y., Yeh, T.-C. J., Mok, C. M. B., Berg, S. J., & Han, D. (2019). Transient Hydraulic Tomography Analysis of Fourteen Pumping Tests at a Highly Heterogeneous Multiple Aquifer–Aquitard System. *Water*, *11*(9), 1864. https://doi.org/10.3390/w11091864

Zheng, C., Liu, J., Cao, G., Kendy, E., Wang, H., & Jia, Y. (2010). Can China Cope with Its Water Crisis?—Perspectives from the North China Plain. *Groundwater*, *48*(3), 350-354. https://doi.org/10.1111/j.1745-6584.2010.00695_3.x

Zhu, J., & Yeh, T.-C. J. (2005). Characterization of aquifer heterogeneity using transient hydraulic tomography. *Water Resources Research*, *41*(7), W07028. https://doi.org/10.1029/2004wr003790

Table 1. Design of Case A and Case B in the steady-state experiments.

	Stress	Pumping Wells	Pumping Rate (m^3/d)	Annotation
Case A	Stress 1	18 pumping wells (combined below)	-3000	Multiple Simultaneous Pumping tests
Case B	Stress 1	6 pumping wells (1/10/32/51/54/62)	-3000	HT Survey
	Stress 2	6 pumping wells (2/22/27/36/48/52)	-3000	
	Stress 3	6 pumping wells (14/16/37/40/42/64)	-3000	

Table 2. Design of Scenario A and Scenario B in transient-state experiments.

Scenario	Stress	Pumping Rate (m^3/d)	Time Period (days)	Status	Pumping Well ID	Annotation
Scenario A	Stress 1	-8000	0-1000	normal rate	1/10/32/51/54/62 (pumping event 1)	Multiple Simultaneous Pumping Tests
		-4000	1001-2000	reduction		
		0	2001-3000	shut down		
		-8000	0-1000	normal rate	2/22/27/36/48/52 (pumping event 2)	
		0	1001-2000	shut down		
		-4000	2001-3000	reduction		
		0	0-1000	shut down	14/16/37/40/42/64 (pumping event 3)	
		-8000	1001-2000	normal rate		
		-4000	2001-3000	reduction		
Scenario B	Stress 1	-8000	0-1000	normal rate	1/10/32/51/54/62 (pumping event 1)	HT Survey
		-4000	1001-2000	reduction		
		0	2001-3000	shut down		
		-8000	0-1000	normal rate	2/22/27/36/48/52 (pumping event 2)	
		0	1001-2000	shut down		
		-4000	2001-3000	reduction		
		0	0-1000	shut down	14/16/37/40/42/64 (pumping event 3)	
		-8000	1001-2000	normal rate		
		-4000	2001-3000	reduction		
	Stress 2	-8000	0-1000	normal rate	2/22/27/36/48/52 (pumping event 1)	
		-4000	1001-2000	reduction		
		0	2001-3000	shut down		
		-8000	0-1000	normal rate	14/16/37/40/42/64 (pumping event 2)	
		0	1001-2000	shut down		
		-4000	2001-3000	reduction		
		0	0-1000	shut down	1/10/32/51/54/62 (pumping event 3)	
		-8000	1001-2000	normal rate		
		-4000	2001-3000	reduction		
	Stress 3	-8000	0-1000	normal rate	14/16/37/40/42/64 (pumping event 1)	
		-4000	1001-2000	reduction		
		0	2001-3000	shut down		
		-8000	0-1000	normal rate	1/10/32/51/54/62 (pumping event 2)	
		0	1001-2000	shut down		
		-4000	2001-3000	reduction		
		0	0-1000	shut down	2/22/27/36/48/52 (pumping event 3)	
		-8000	1001-2000	normal rate		
		-4000	2001-3000	reduction		

Table 3. Design of three cases for temporal sampling.

	Selected data for inversion	Annotation
Case 1	all head data with a time step of 100 days	Baseline analysis
Case 2	Ten head data for each observation well (5, 10, 15, 20, 30, 1000, 1500, 2000, 2500, 3000 day)	Early-Mid-Late data
Case 3	Five head data for each observation well (1000, 1500, 2000, 2500, 3000 day)	Mid-Late data

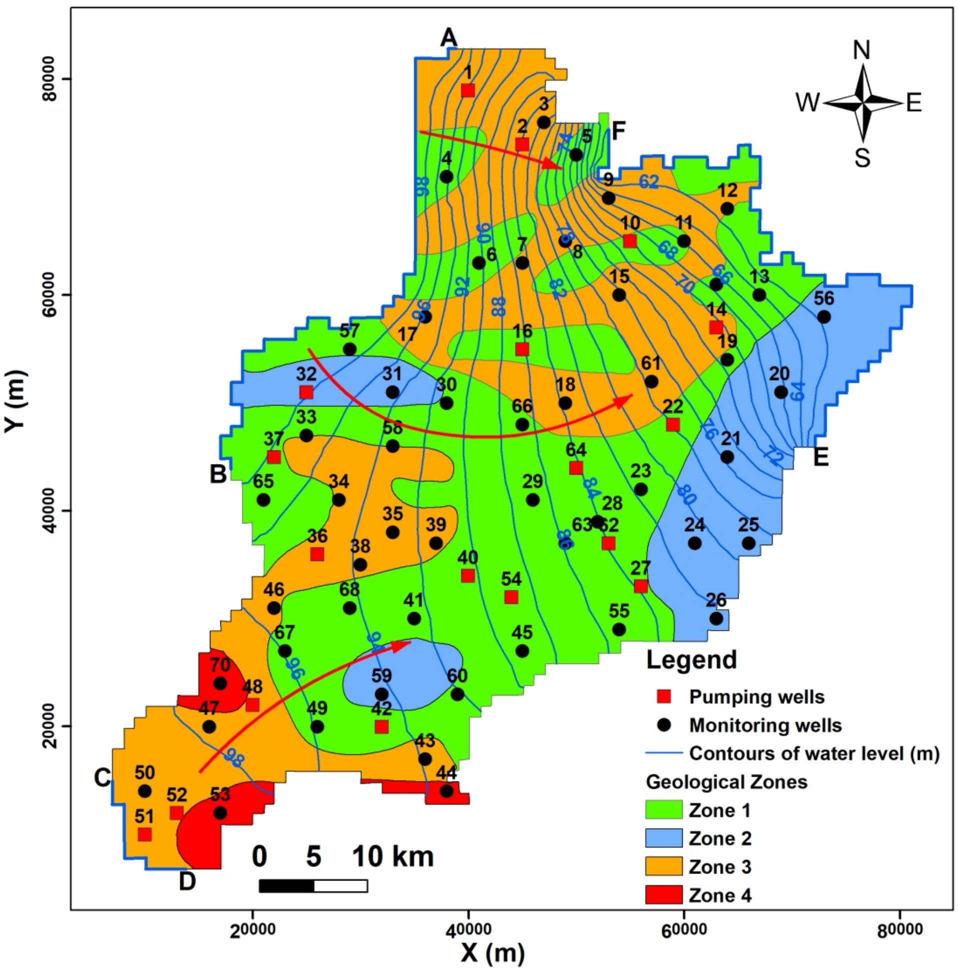

Fig. 1. Map of the synthetic groundwater basin showing the steady-state flow field with no pumping, configurations of pumping wells and monitoring wells, boundary conditions, as well as geological zones.

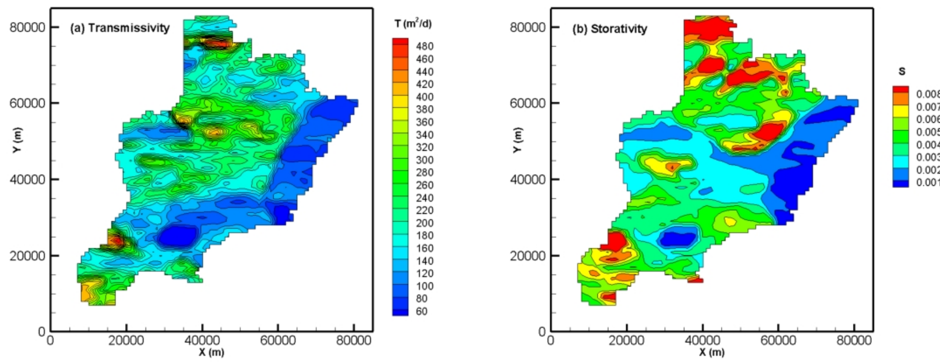

Fig. 2. Reference transmissivity (a) and storativity (b) fields in the synthetic domain.

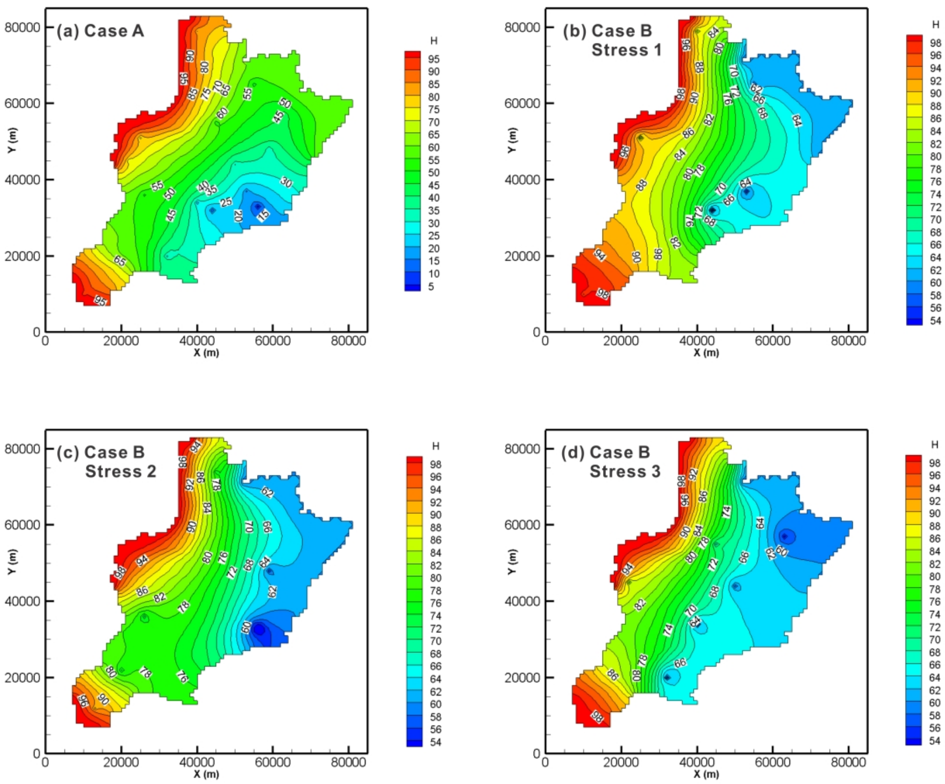

Fig. 3. Groundwater flow fields generated from (a) Case A and (b-d) Case B.

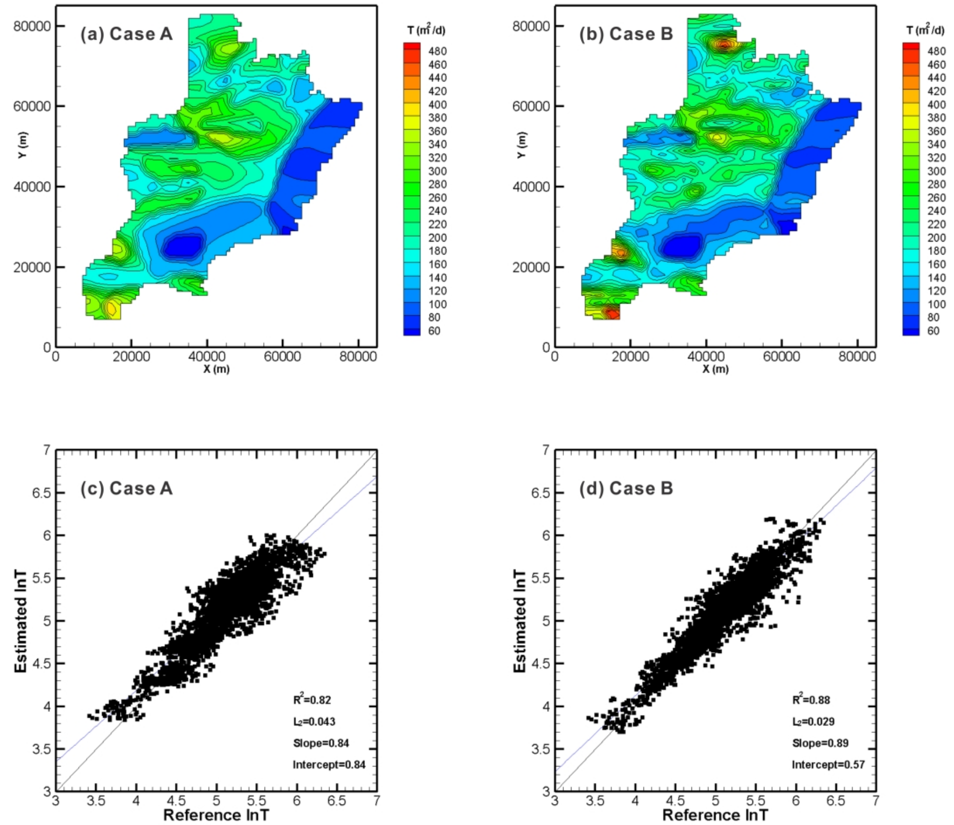

Fig. 4. T tomograms for (a) Case A and (b) Case B, as well as the corresponding scatterplots of reference lnT versus estimated lnT for (c) Case A and (d) Case B.

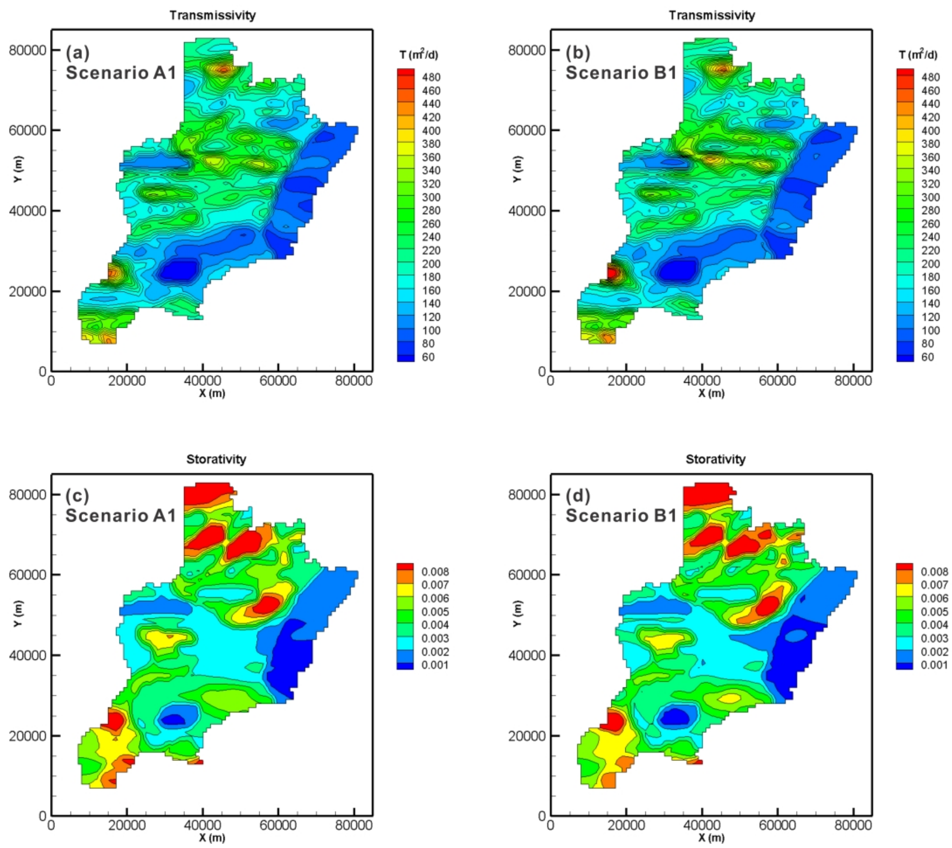

Fig. 5. The estimated transmissivity tomograms (a, b) and the corresponding storativity tomograms (c, d) for Scenarios A1 and B1.

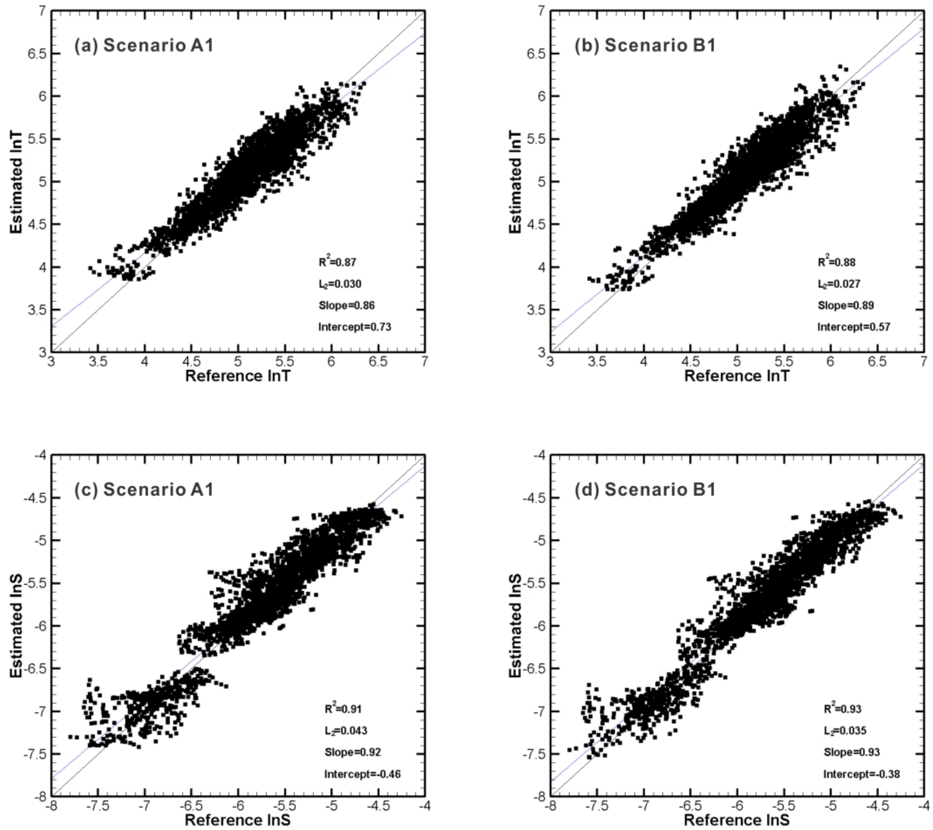

Fig. 6. Scatterplots of reference versus estimated values of the transmissivity (a, b) and the storativity (c, d) for Scenarios A1 and B1.

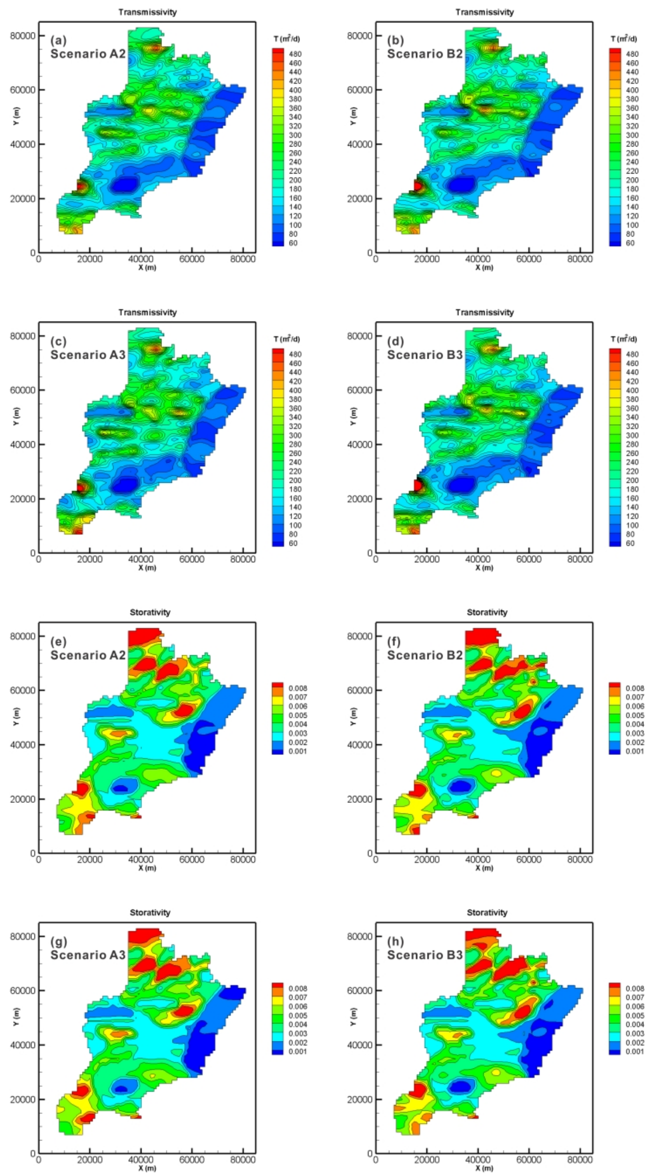

Fig. 7. The estimated transmissivity tomograms (a-d) and the corresponding storativity tomograms (e-h) for Scenarios A2, B2, A3, and B3.

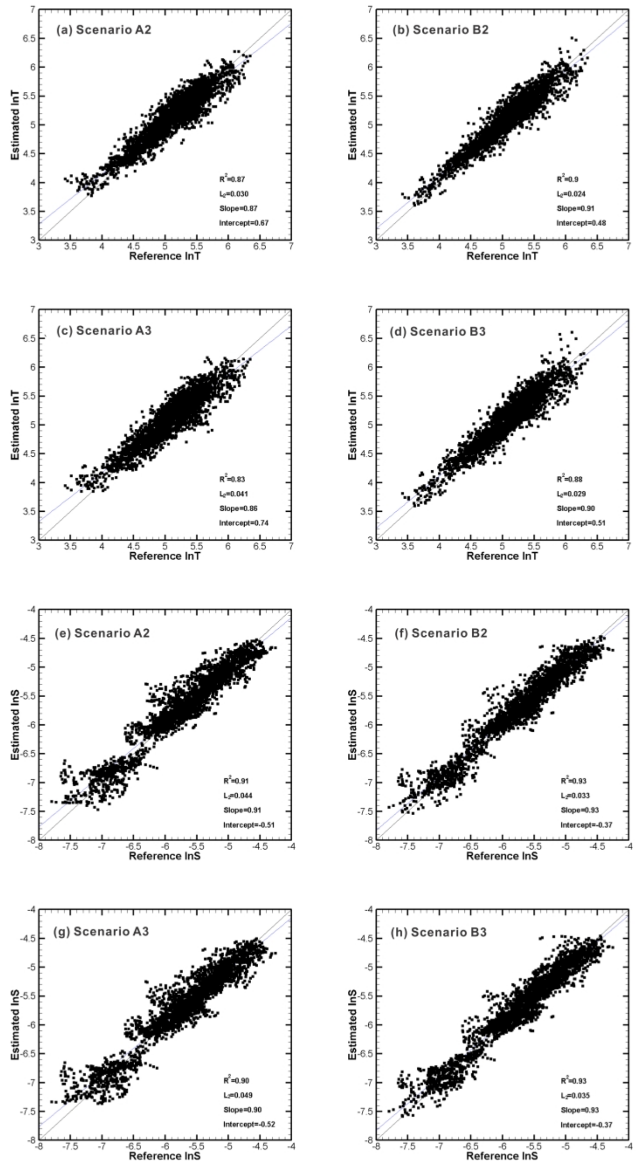

Fig. 8. Scatterplots of reference versus estimated values of the transmissivity (a-d) and the storativity (e-h) for Scenarios A2, B2, A3, and B3.

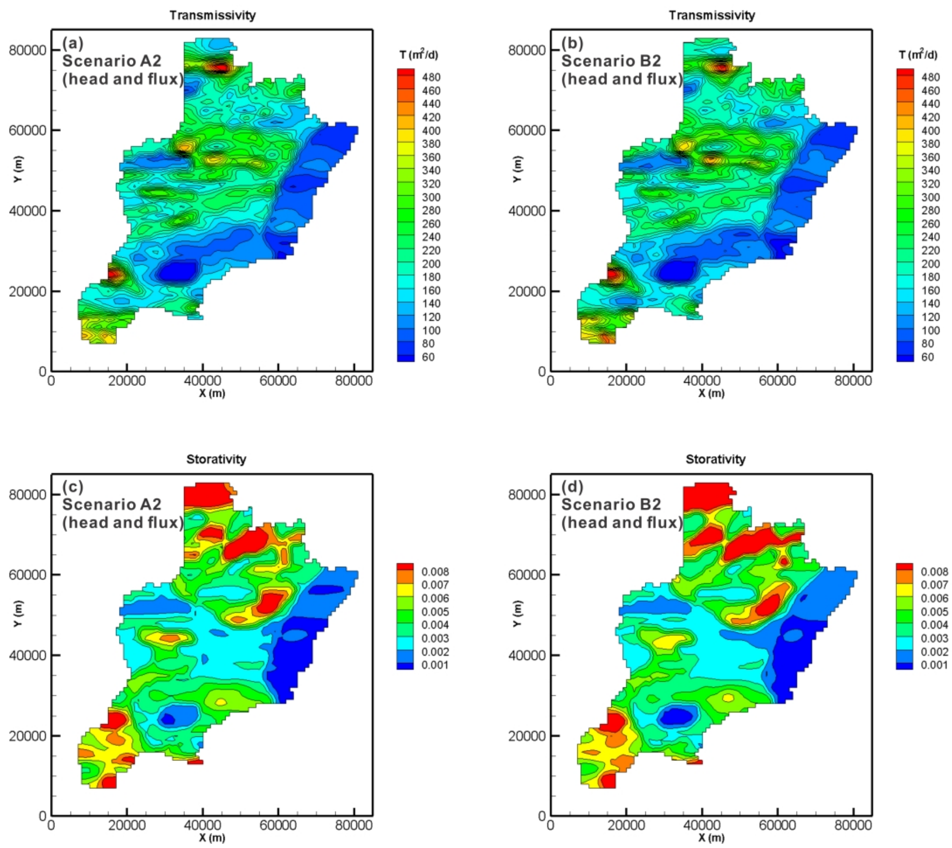

Fig. 9. The estimated transmissivity (a, b) and storativity tomograms (c, d) from head and flux joint inversions for Scenarios A2 and B2.

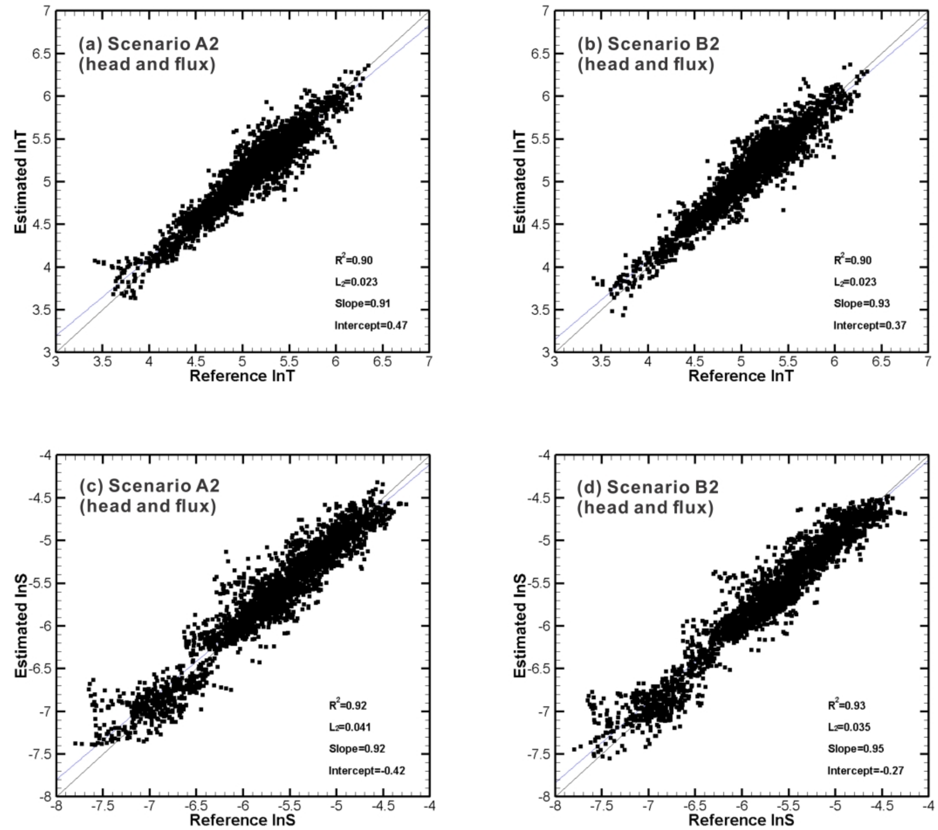

Fig. 10. Scatterplots of reference versus estimated values of transmissivity (a, b) and storativity (c, d) from head and flux joint inversions for Scenarios A2 and B2.

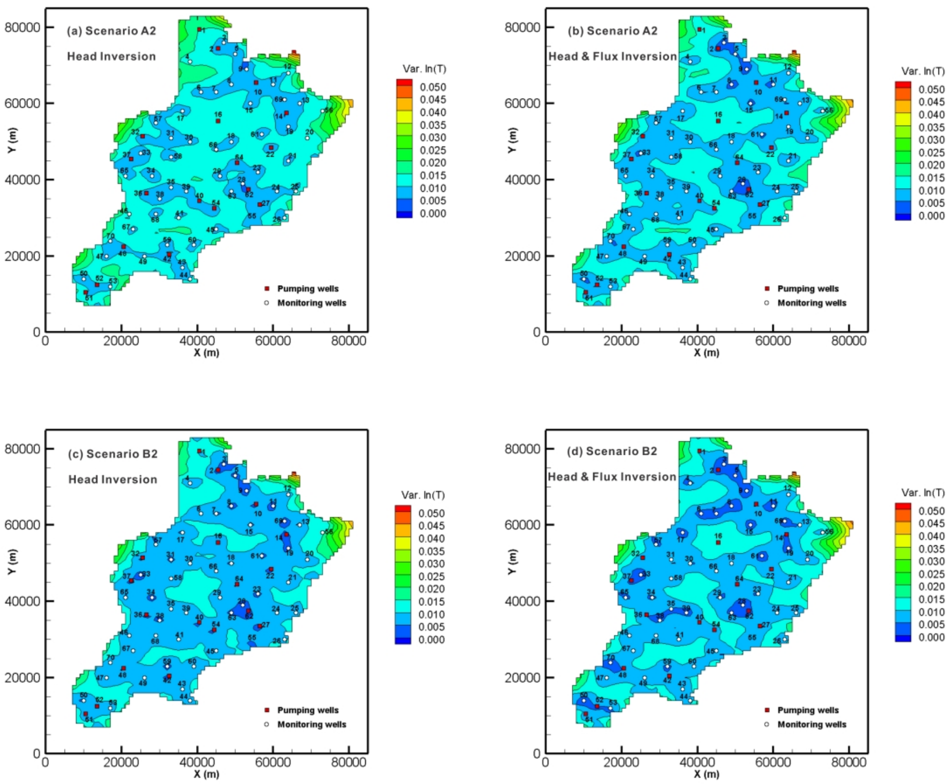

Fig. 11. Variances (uncertainties) of the T estimates from (a) head inversion in Scenario A2, (b) head and flux inversion in Scenario A2, (c) head inversion in Scenario B2, and (d) head and flux inversion in Scenario B2.

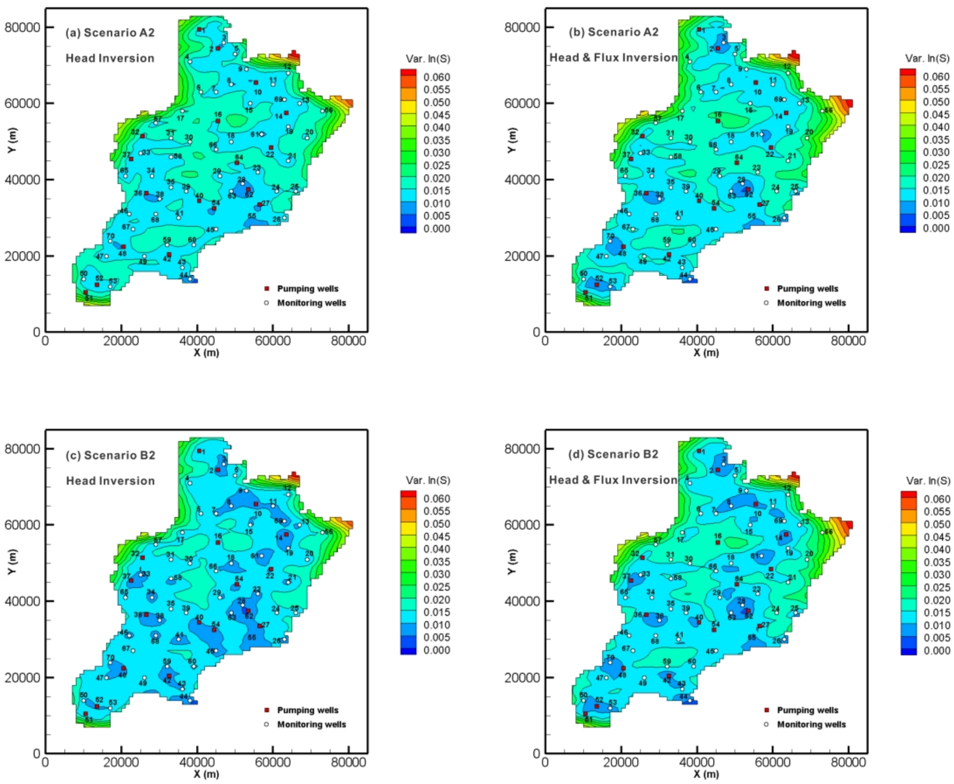

Fig. 12. Variances (uncertainties) of the S estimates from (a) head inversion in Scenario A2, (b) head and flux inversion in Scenario A2, (c) head inversion in Scenario B2, and (d) head and flux inversion in Scenario B2.