

Research papers

Characterization of basin-scale aquifer heterogeneity using transient hydraulic tomography with aquifer responses induced by groundwater exploitation reduction

Fei Liu^{a,c}, Tian-Chyi Jim Yeh^{b,c,*}, Yu-Li Wang^c, Yonghong Hao^b, Jet-Chau Wen^{d,e}, Wenke Wang^f^a School of Water Conservancy and Hydropower, Hebei University of Engineering, 19 Taiji Road, Handan, Hebei 056000, China^b Tianjin Key Laboratory of Water Resources and Environment, Tianjin Normal University, Tianjin 300387, China^c Department of Hydrology and Atmospheric Sciences, University of Arizona, Tucson, AZ 85721, USA^d Department of Safety, Health, and Environmental Engineering, National Yunlin University of Science and Technology, Douliu, Taiwan 64002, China^e Research Center for Soil and Water Resources and Natural Disaster Prevention, National Yunlin University of Science and Technology, Douliu, Taiwan 64002, China^f Key Laboratory of Subsurface Hydrology and Ecological Effects in Arid Region, Chang'an University, Xi'an 710054, China

ARTICLE INFO

This manuscript was handled by Corrado Corradini, Editor-in-Chief, with the assistance of Patrick Lachassagne, Associate Editor

Keywords:

Hydraulic tomography
Basin-scale aquifer heterogeneity
Groundwater exploitation reduction
North China Plain

ABSTRACT

This study exploits aquifer responses to the reduction of pumping rates at different locations in a synthetic groundwater basin as basin-scale hydraulic tomography (HT) surveys to estimate transmissivity (T) and storage coefficient (S) fields. This experiment mimics the situation of groundwater exploitation reduction in a pilot area of groundwater-overexploitation control in the North China Plain. The results of the study show that taking advantage of the groundwater exploitation reduction as HT surveys is a viable approach for basin-scale parameter estimations. Results also suggest that HT analysis should use accurate mean values of T and S for geological zones as initial guesses for the inversion of parameters. Further, we show that the T and S fields estimated from HT yield accurate predictions of the groundwater flow velocities and breakthrough curves (BTCs). However, the BTCs based on kriged and zonal mean fields are inaccurate. The predicted BTCs using homogeneous fields fail to capture the true trend of solute concentration over time. We advocate that utilizing aquifer responses induced by groundwater exploitation reduction could be a new paradigm for basin-scale aquifer characterization.

1. Introduction

Identifying the spatial distribution of hydraulic property in large-scale groundwater basins is crucial to the sustainable management of groundwater resources (Fakhreddine et al., 2016; Straface et al., 2007). Subsurface hydraulic heterogeneity dictates the behaviors of solute transport in groundwater, and hence accurate identification of hydraulic properties in heterogeneous aquifers is of pivotal importance for reliable prediction of solute transport (Jiménez et al., 2015).

Hydraulic properties of large-scale groundwater basin (over hundreds of kilometers) are generally acquired by model calibrations, which yield aquifer characterization with considerable uncertainty given that sources of excitations in aquifers are seldom fully characterized and only sparse temporal and spatial responses of aquifers are available (Kitanidis and Vomvoris, 1983; Yeh et al., 2009). Due to the significant uncertainties in the estimated hydraulic properties by traditional methods, the accuracy of predicting flow and solute transport

in groundwater basins has been seriously questioned (Bredehoeft, 2003).

In the past few decades, numerous studies have focused on identifying the fields of hydraulic conductivity and specific storage using traditional pumping tests and geophysical methods such as electrical resistivity tomography and ground-penetrating radar (Fakhreddine et al., 2016). However, for groundwater basins, these methods are costly and time-consuming that they are impractical (Yeh and Liu, 2000).

As an efficient geophysical tool, the airborne electromagnetic (EM) method has been widely used in geologic mapping, groundwater exploration, and environmental investigations (Paine and Collins, 2003; Palacky, 1993; Yin et al., 2016). Although airborne EM is cost-effective (Guillemeau et al., 2011), its penetration depth is limited and it detects anomalies in electric resistivity rather than hydraulic properties (Fitterman, 1990).

Hydraulic tomography (HT) is an emerging technique for imaging

* Corresponding author at: Department of Hydrology and Atmospheric Sciences, University of Arizona, Tucson, AZ 85721, USA.
E-mail address: yeh@hwr.arizona.edu (T.-C.J. Yeh).

detailed aquifer heterogeneity (Hochstetler et al., 2016; Xiang et al., 2009; Yeh et al., 2008; Yeh and Liu, 2000; Zha et al., 2019; Zhao and Illman, 2018; Zhu and Yeh, 2005), which can integrate aquifer responses from numerous pumping tests. The robustness of HT has been confirmed in small-scale synthetic aquifers (Illman et al., 2008, 2010; Illman et al., 2008), laboratory sandboxes (Illman et al., 2007; Liu et al., 2002, 2007; Zhao et al., 2016), and field studies (Berg and Illman, 2011; Cardiff et al., 2012; Doro et al., 2015; Fischer et al., 2017; Kuhlman et al., 2008; Zha et al., 2016). More importantly, the spatial patterns of hydraulic properties estimated by HT can produce better predictions of flow and solute transport than traditional aquifer imaging approaches (Illman et al., 2012; Ni et al., 2009; Tso et al., 2016).

Pumping from a well to create detectable responses throughout large-scale aquifers is unlikely. As a result, how to adequately stress the entire large-scale aquifer becomes a challenge. Kuhlman et al. (2008) proposed and demonstrated the utility of conducting multiple HT surveys in a synthetic groundwater basin. Yeh et al. (2008) promoted exploiting natural stimuli (atmospheric pressure variation, lightning, solid earth tides, river stage, etc.) as energy sources for basin-scale hydrologic and geophysical tomographic surveys. Yeh et al. (2009) demonstrated the feasibility of exploiting the river stage variation for basin-scale aquifer characterization in a synthetic aquifer. Wang et al. (2017) applied the river stage tomography concept to characterize the aquifer in the Zhusui river fan in Taiwan.

There are rooms for improvement in this emerging technology of HT, in particular for imaging basin-scale aquifer heterogeneity (Illman, 2014). The cost-effectiveness of HT derives from the collection of nonredundant information from a limited number of wells (Sun et al., 2013; Yeh et al., 2014). New strategies that collect additional nonredundant information at the existing observation wells used by HT experiments would be the most attractive (Cardiff et al., 2013a,b; Tso et al., 2016). For example, Zha et al. (2019) took advantage of the change in the flow fields due to variations in pump-and-treat operation as a large-scale HT survey to capture the low permeable zones for hydrofracking to enhance remediation effectiveness at the U.S. Air Force Plant 44, Tucson, Arizona, USA.

Analogous to natural stimuli, anthropogenic perturbations can also cause a series of significant groundwater level fluctuations. The North China Plain (NCP), one of the extremely populated and water-stressed areas of the world, has been enduring persistent problems associated with groundwater depletion and deterioration of water quality in the past decades. As a result, the government has initiated the groundwater-exploitation reduction program to restore the groundwater resources and to prevent further deterioration in the NCP since 2014 (Xu, 2017; Zhao et al., 2017). In the NCP, a large number of pumping and monitoring wells have been installed, and groundwater levels have been monitored. More importantly, long-term and high-frequency records of groundwater levels have been collected for many years. Therefore, we proposed to utilize the vast amount of existing wells and datasets to characterize the heterogeneity of aquifers in the NCP. Intermittent shutdown and resumption of different pumping wells in the NCP alters the large-scale flow fields. Such changes in flow fields are equivalent to a large-scale HT survey in the NCP if long-term datasets of aquifer responses are available. 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. A shift of aquifer characterization strategies towards using existing wells and observation data is highly encouraged (Cardiff et al., 2013a).

The main objectives of the study are to (1) explore the feasibility of utilizing aquifer responses induced by anthropogenic excitations to enhance basin-scale aquifer characterization; (2) evaluate the improvement of estimated transmissivity (T) and storage coefficient (S) fields using different prior means (i.e., uniform mean, kriged mean, and zonal mean); (3) assess the accuracy of solute transport prediction with different estimated parameter fields. The results of this study may guide

the future applications of this approach to the aquifer characterization in the NCP when data become available.

2. Methodology

This study focuses on the transient flow and nonreactive solute transport problems in a horizontal, two-dimensional, confined aquifer with spatially varying hydraulic properties. The following sections brief the basic concepts and parameter estimation principles associated with the corresponding governing equations and assumptions for predicting flow and transport in groundwater.

2.1. Groundwater flow and transport model

This study assumes that the following partial differential equation governs groundwater flow in a two-dimensional, depth-averaged, saturated, heterogeneous confined aquifer:

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

subject to the boundary and initial conditions:

$$H|_{\Gamma_1} = H_1, \quad [-T(x) \nabla H] \cdot \mathbf{n} |_{\Gamma_2} = q, \quad \text{and} \quad H|_{t=0} = H_0 \quad (2)$$

where, in Eq. (1), H is the total head [L], x is the spatial coordinate ($x = (x, y)$, [L]), $Q(x_p)$ is the pumping rate per unit area (L/T) at the location x_p , $T(x)$ is the transmissivity [L^2/T], and $S(x)$ is the storativity [–]. In Eq. (2), H_1 is the prescribed total head at Dirichlet boundary Γ_1 , q is the specified flux at Neumann boundary Γ_2 , \mathbf{n} is a unit vector normal to Γ_2 , and H_0 represents the initial total head.

Likewise, the conservative solute transport processes along with groundwater flow in saturated porous media are presumed to be governed by the classical advection-dispersion equations (CADE) as follows:

$$\nabla \cdot [D(x) \nabla C(x, t)] - V(x) \nabla C(x, t) + Q_c(x, t) = \frac{\partial C(x, t)}{\partial t} \quad (3)$$

subject to boundary and initial conditions:

$$C(x, t)|_{\Gamma_1} = C_D, \quad [D(x) \nabla C(x, t)] \cdot \mathbf{n} |_{\Gamma_2} = q_c, \quad \text{and} \quad C(x, t)|_{t=0} = C_0 \quad (4)$$

where, in Eq. (3), $C(x, t)$ is the volumetric solute concentration measured in groundwater [M/L^3], x is the spatial coordinate ($x = (x, y)$, [L]), $Q_c(x, t)$ is the rates where the volumetric solute concentration is pumped from groundwater (M/L^3T) at the location x , $D(x)$ is the dispersion coefficient [L^2/T], and $V(x)$ is the flow velocity [L/T], which is equal to the specific discharge divided by the effective porosity. In Eq. (4), C_D is the prescribed concentration at Dirichlet boundary Γ_1 , q_c is the dispersive flux at Neumann boundary Γ_2 , \mathbf{n} is a unit vector normal to Γ_2 , and C_0 is the initial concentration.

In this study, these governing equations 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>.

2.2. Hydraulic tomography analysis

In this study, transient HT analysis was conducted using the Simultaneous Successive Linear Estimator (SimSLE) developed by Xiang et al. (2009). The SimSLE estimator is inherited from the Successive Linear Estimator (SLE) developed by Yeh et al. (1996). The robustness of the SimSLE algorithm is that the estimation of aquifer properties can be accomplished by simultaneously considering all observed hydraulic head changes due to various pumping events during an HT survey, instead of incorporating the aquifer responses from discrete sources sequentially into the inversion process as is done in the Sequential Successive Linear Estimator (SSLE) (Ni and Yeh, 2008; Zhu and Yeh, 2005). The SimSLE estimator is briefly described below.

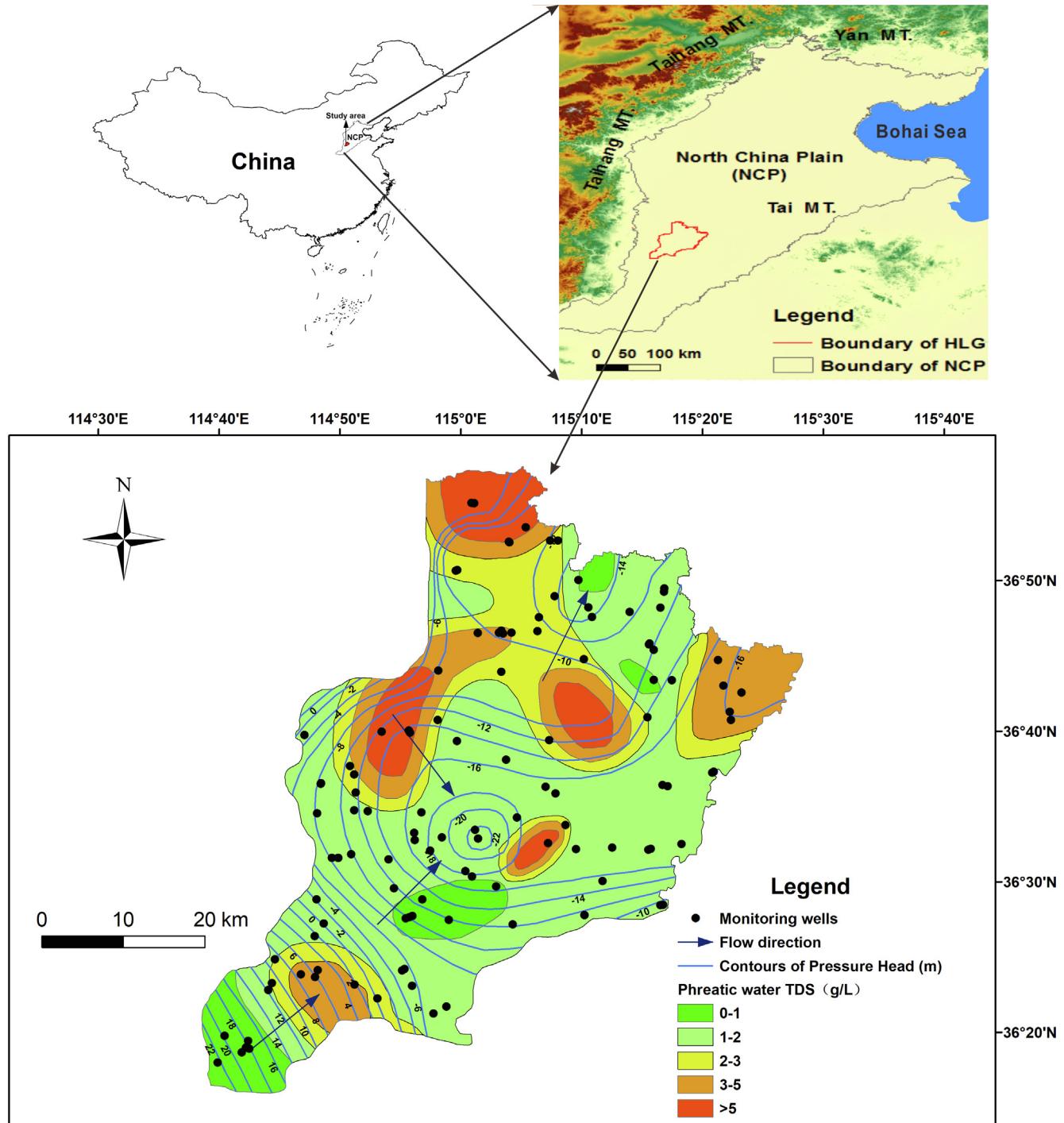


Fig. 1. The geographical location of the Heilonggang plain and the spatial distribution of phreatic water TDS in the NCP.

In the SimSLE method, the aquifer properties (T and S) are conceptualized as highly-parameterized random fields, described by a prior geostatistical model (mean, variance and correlation scales). The SimSLE starts with cokriging to estimate the conditional expectation of the aquifer parameters (T and S) conditioned on $f(x_i)$ and $h(k, x_j, t_d)$. The term $f(x_i)$ denotes the perturbations of either $\ln T$ or $\ln S$ measured at location x_i ($i = 1, \dots, N_f$, where N_f is the total number of measured parameters). Another term $h(k, x_j, t_d)$ represents the observed head perturbation at location x_j and time t_d during the k th pumping event. The linear estimator is

$$\hat{f}^{(1)}(x_0) = \sum_{i=1}^{N_f} \lambda_{0i} f(x_i) + \sum_{k=1}^{N_p} \sum_{j=1}^{N_h(k)} \sum_{d=1}^{N_t(k,j)} \beta_{0kjd} h(k, x_j, t_d) \quad (5)$$

where $\hat{f}^{(1)}(x_0)$ is the cokriged f value at location x_0 ; N_p is the total number of pumping events; $N_h(k)$ is the total number of monitoring wells for the k th pumping event. The total number of transient head measurements at location x_j during the k th pumping event is denoted by $N_t(k, j)$. Meanwhile, λ_{0i} is the cokriging weight, which represents the contribution of parameter measurement at location x_i to the estimate at location x_0 . β_{0kjd} denotes the contribution from the head perturbation $h(k, x_j, t_d)$ to the estimate at location x_0 . These weights are calculated by solving the cokriging system of equations (Xiang et al., 2009).

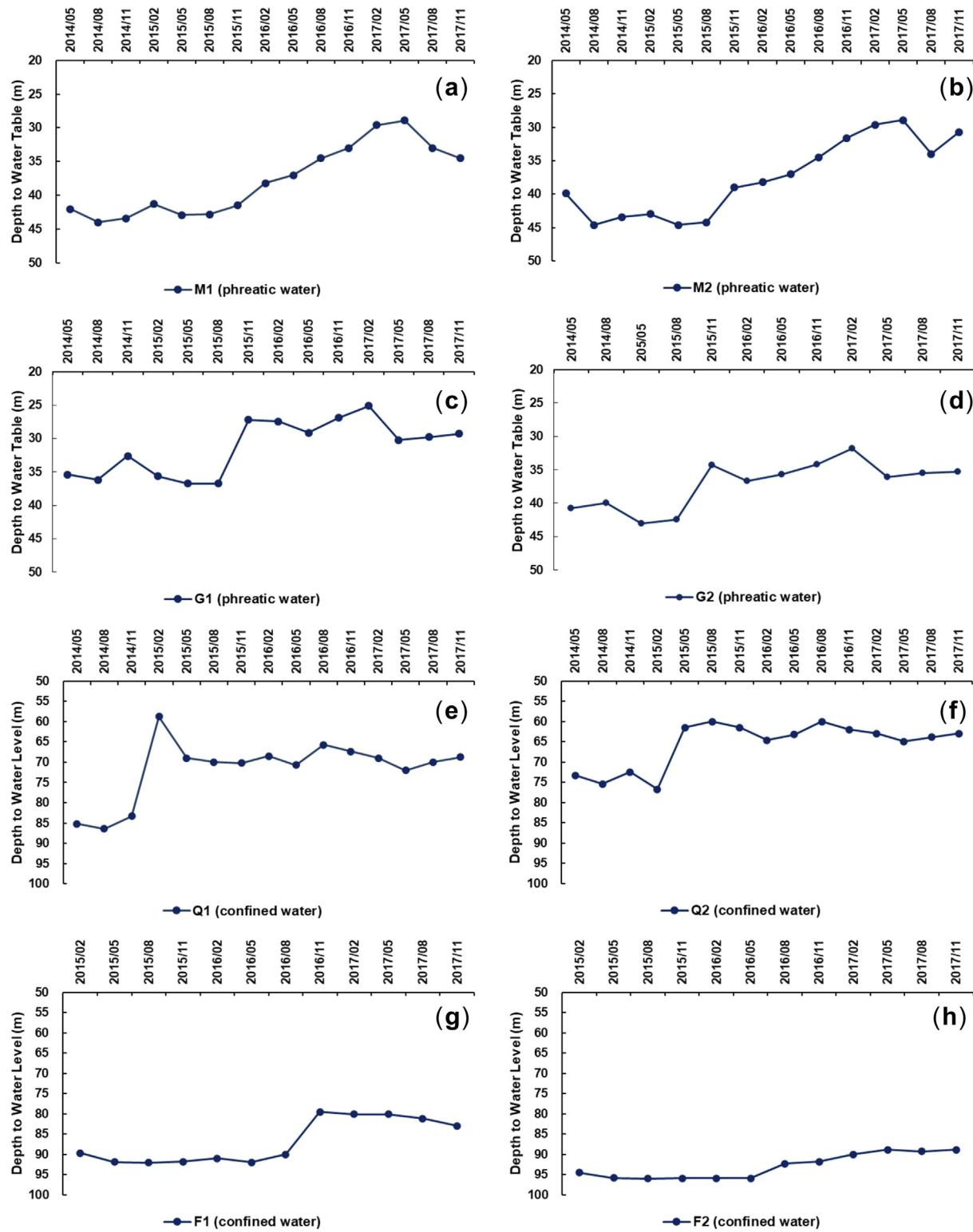


Fig. 2. Hydrographs in representative monitoring wells screened at phreatic aquifer (a-d) and confined aquifer (e-h) after the groundwater exploitation reduction program.

The new estimate for all the elements can be generated by using cokriging, and then the conditional covariance, $\varepsilon_{ff}^{(1)}$, is computed as follows:

$$\begin{aligned}
 \varepsilon_{ff}^{(1)}(x_m, x_n) &= R_{ff}(x_m, x_n) - \sum_{i=1}^{N_f} \lambda_{mi} R_{ff}(x_i, x_n) - \sum_{k=1}^{N_p} \sum_{j=1}^{N_h(k)} \sum_{d=1}^{N_l(k,j)} \beta_{0kjd} \\
 &= R_{hf}((k, x_j, t_d), x_n)
 \end{aligned} \tag{6}$$

where m and n represent the element number. R_{ff} denotes the

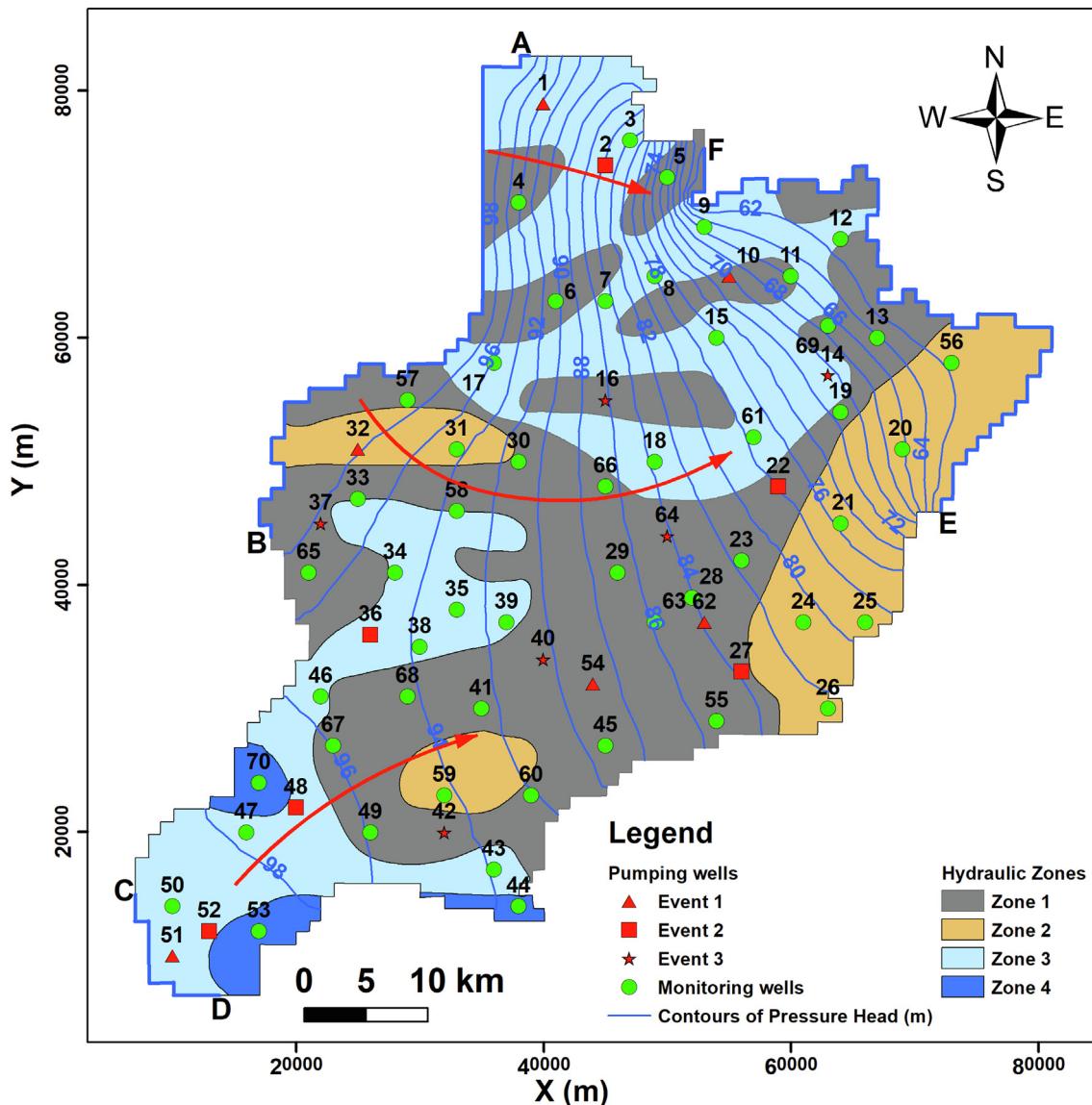


Fig. 3. Map of the synthetic groundwater basin showing the steady-state flow field with no pumping, configurations of pumping wells and monitoring wells, boundary conditions, as well as the zones of hydraulic properties.

unconditional covariance of the parameter, while R_{hf} is the cross-covariance between h and f . The conditional covariance can reflect the reduction of uncertainty in the parameter estimates along with the data assimilation of hydraulic head observations. Subsequently, the estimated fields of parameters are used to generate the conditional effective head fields.

A stochastic linear estimator is then used to improve the estimate for iteration $r > 1$, in which the nonlinear relationship between f and h is handled by a successive estimation scheme:

$$\hat{f}^{(r+1)}(x_0) = \hat{f}^{(r)}(x_0) + \sum_{k=1}^{N_p} \sum_{j=1}^{N_h(k)} \sum_{d=1}^{N_t(k,j)} \omega_{0kjd}^{(r)} [h(k, x_j, t_d) - h^{(r)}(k, x_j, t_d)] \quad (7)$$

where $\omega_{0kjd}^{(r)}$ is the weights, representing the contribution of the difference between the observed head $h(k, x_j, t_d)$ and simulated $h^{(r)}$ (k, x_j, t_d) at iteration r at location x_j and time t_d during the k th pumping test to the estimation at location x_0 . The weighting coefficients are determined by

$$\begin{aligned} & \sum_{k=1}^{N_p} \sum_{j=1}^{N_h(k)} \sum_{d=1}^{N_t(k,j)} \omega_{0kjd}^{(r)} [\varepsilon_{hh}^{(r)}((p, x_m, t_e), (k, x_j, t_d)) + \Theta^{(r)} \delta_{kjd}] \\ &= \varepsilon_{hf}^{(r)}((p, x_m, t_e), x_0) \end{aligned} \quad (8)$$

where $p = 1, \dots, N_p$, $m = 1, \dots, N_h$ and $e = 1, \dots, N_t$, $\varepsilon_{hh}^{(r)}$ is the conditional auto-covariance of observed head data at iteration r , and the term $\varepsilon_{hf}^{(r)}$ is the conditional cross-covariance between parameter f and head data at iteration r . The covariances involved in Eq. (8) are calculated by the first-order numerical approximation (Yeh and Liu, 2000). SimSLE stabilizes the solution by adding a dynamic stabilizer $\Theta^{(r)}$ (δ_{kjd} is the Dirac delta, equal to one if $k = j = d$ and zero otherwise) to the diagonal elements of $\varepsilon_{hh}^{(r)}$. The diagonal elements vary with iterations, reflecting the degree of improvement of the estimates. Thus, this approach is different from that of Levenberg-Marquardt algorithm (Pujol, 2007).

ε_{ff} is the conditional covariance of parameter f , which is constructed by a geostatistical prior model with a description of mean, variance, and correlation scales at $r = 0$. After achieving the new parameter estimates using Eq. (7), ε_{ff} is updated subsequently according to the following equation:

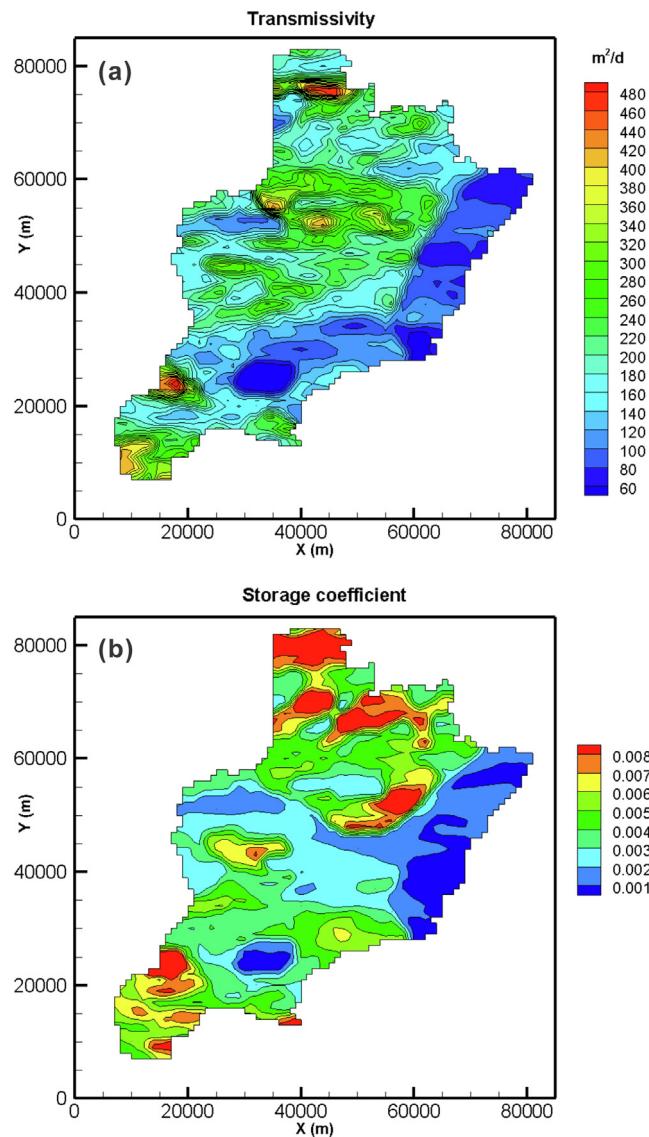


Fig. 4. Spatial distribution of true transmissivity (a) and storage coefficient (b) in the synthetic domain.

Table 1
Designation of transient hydraulic tomography in the synthetic experiment.

Scenario	Well ID	Pumping rate (m ³ /d)	Time period (day)	Status
Pumping event 1	1/10/32/ 51/54/62	-8000	0–1000	Normal rate
		-4000	1001–2000	Reduction
		0	2001–3000	Shut down
Pumping event 2	2/22/27/ 36/48/52	-8000	0–1000	Normal rate
		0	1001–2000	Shut down
		-4000	2001–3000	Reduction
Pumping event 3	14/16/37/ 40/42/64	0	0–1000	Shut down
		-8000	1001–2000	Normal rate
		-4000	2001–3000	Reduction

$$\varepsilon_{ff}^{(r+1)}(x_m, x_n) = \varepsilon_{ff}^{(r)}(x_m, x_n) - \sum_{k=1}^{N_p} \sum_{j=1}^{N_h(k)} \sum_{d=1}^{N_t(k,j)} \omega_{mkjd}^{(r)} \varepsilon_{hf}^{(r)}((k, x_j, t_d), x_n) \quad (9)$$

The iterative process among Eqs. (7)–(9) continues until one of the following two criteria is met. One is that change in variances between

the current estimated hydraulic properties and prior iterations is smaller than a specified tolerance, while the other is that the change of simulated heads between successive iterations is smaller than the user-specified head tolerance (Xiang et al., 2009).

2.3. Performance metrics for HT inversion

Scatterplots of the true T and S values versus their estimated counterparts are used to assess the HT outcomes with various pieces of prior information. Moreover, the criteria of the coefficient of determination (R^2), the mean absolute error (L_1), the mean square error (L_2), the slope and intercept are also used to evaluate the performance of HT inverse model. L_1 and L_2 are described as the following equations:

$$L_1 = \frac{1}{N} \sum_{i=1}^N |x_i - \hat{x}_i| \quad L_2 = \frac{1}{N} \sum_{i=1}^N (x_i - \hat{x}_i)^2 \quad (10)$$

where N is the total number of elements, i is the element number, x_i is the true T or S value at the i th element, and \hat{x}_i is the estimated T or S value at the i th element. In general, the higher R^2 indicates a better result; the smaller values for L_1 and L_2 are the better estimates. The slope is an indicator of bias. The closer the slope of the linear regression line to 1 and the intercept to 0, the better the estimates are.

3. Numerical experiments

3.1. Site description

The Heilonggang (HLG) plain in the NCP is a pilot area of groundwater-overexploitation control, which stretches 9–57 km wide from west to east and 76 km long from south to north, covering a total area of approximately 2695 km² (Fig. 1). The elevation in the study area ranges from 96 m above sea level (m.a.s.l.) in the southwest, to 34 m.a.s.l. in the northeast part. The overall terrain slope ranges from 0.1‰ to 0.2‰. The main landforms in the study area are alluvial-proluvial plain, and alluvial-lacustrine plain, respectively. A temperate semi-humid climate prevails in the study area, characterized by an obvious seasonal pattern. The mean annual precipitation was in the range of 500–600 mm, and 70–80% of annual precipitation fell during the monsoon from June to September. The average annual pan evaporation ranged from 1100 to 1800 mm (Wang, 2011).

The HLG is located in the southern section of the North China subsidence zone, where the unconsolidated Quaternary sediments are well developed with a maximum thickness of 560 m (Zhou, 2008). The Quaternary deposits consist of gravel, sand, silt and clay. The aquifer system in the study area can be divided into two types: phreatic aquifer and confined aquifer. The unconfined and the confined aquifers are separated by an discontinuous aquitard composed of silt and clay with thickness around 10 m.

The phreatic aquifer is mainly composed of medium to fine-grained sands with depth to the bottom of the aquifer varying from 100 to 150 m, and the transmissivity of the aquifer ranges from 100 to 300 m²/d. Precipitation infiltration and irrigation return flow via vadose zone as well as lateral flow are the major sources of phreatic water. The phreatic water mainly discharges as lateral outflow, artificial exploitation and leakage. Given that the depth to the water table is generally large in the study area, direct evaporation from the phreatic aquifer is small. It is noteworthy that the saline water with total dissolved solids (TDS) exceeding 2 g/L is widespread in the phreatic aquifer (Fig. 1), which is a potential contaminant source for groundwater quality in the underlying confined aquifer.

This study focuses on the confined aquifer, which is the most promising water-supply aquifer for agricultural, domestic, and industrial usages. Along with the groundwater exploitation reduction program, the pressure head is anticipated to vary significantly. The confined aquifer is mainly composed of alluvial sands and gravels during the

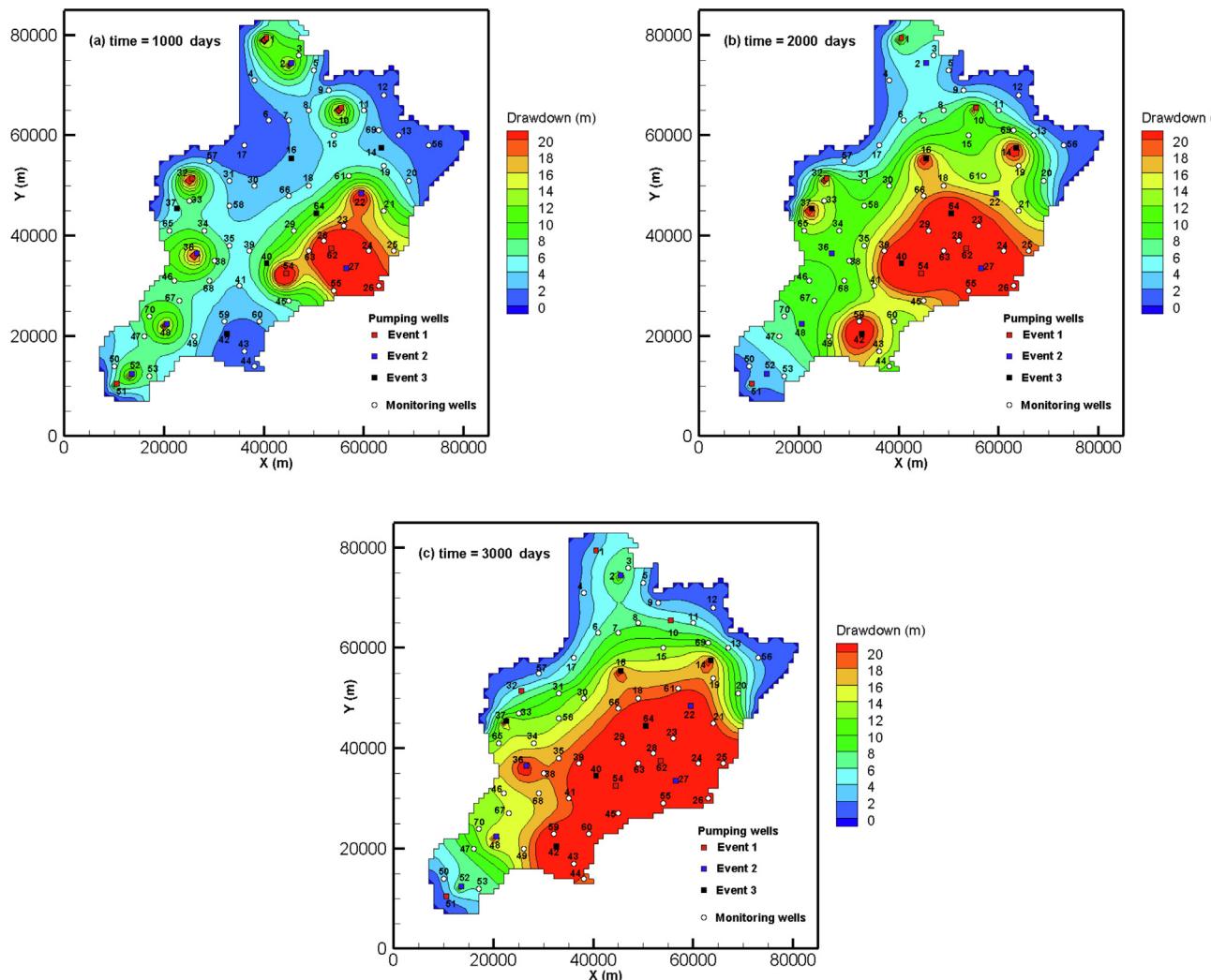


Fig. 5. Evolution of the drawdown fields at selected time 1000 (a), 2000 (b) and 3000 days (c) during the simulation period.

Quaternary, with depth to the bottom of the aquifer varying from 350 to 500 m. The hydraulic properties of the aquifer are highly variable in space. Generally, the transmissivity changes between 10 and $480 \text{ m}^2/\text{d}$, and the storage coefficient ranges from 0.001 to 0.008. The transmissivity in the low permeable parts, which is composed of silt, varies from 10 to $50 \text{ m}^2/\text{d}$, and the storage coefficient is around 0.001 (Wang, 2011). Overall, the confined groundwater flows predominantly from southwest to northeast (Fig. 1). A huge groundwater cone of depression due to long-term pumping has existed in the center of the study area. Over the past years, the difference in hydraulic head between unconfined and confined aquifers has reached more than 60 m in some areas. As a result, leakage from the upper aquifer has become the dominant recharge source of the confined aquifer with an approximate percentage of 90% (Wei, 2018). Leakage from the overlying unconfined aquifer with high TDS into the confined aquifer becomes a potential source of contamination.

The groundwater exploitation reduction program implemented by the government, via agricultural water-saving projects and water transfer from the middle route of the South-to-North Water Diversion Project, has reduced the total amount of groundwater pumping since 2014. This program has profound impacts on the groundwater and caused a series of changes in the aquifer system (Zhao et al., 2017). Well hydrographs at monitoring wells, as shown in Fig. 2, provide direct evidence of the recovery of groundwater level in some regions. The rebound of the water levels has taken place from one location to

another during the years from 2014 to 2017.

For the phreatic aquifer, relative to the mean annual groundwater level in 2014, the mean annual groundwater level at the monitoring wells M2 and G2 increased 11.83 and 5.73 m, respectively, in 2017 (Fig. 2a-d). Similarly, the groundwater level has also increased in the confined aquifer. The average annual hydraulic heads at the observation wells Q1 and Q2 ascended 15.04 and 10.03 m in 2017, compared with the corresponding values in 2014 (Fig. 2e and f). Meanwhile, the mean annual groundwater levels at the wells F1 and F2 increased 10.28 and 6.27 m in 2017, relative to those in 2015 (Fig. 2g and h).

3.2. Synthetic experiment setup

The synthetic confined aquifer was designed to mimic the confined aquifer in the HLG plain. The synthetic aquifer follows the shape of the aquifer in the HLG plain and has an area of 2695 km^2 . This aquifer is discretized into 2896 nodes that form 2720 rectangular elements. The dimensions of each element are 1 km (length) \times 1 km (width). Groundwater flows dominantly from southwest to northeast boundary (Fig. 3). The boundary segments perpendicular to the contours of the groundwater level, are no flux boundaries. The remaining boundary segments (A-B, C-D, E-F) are assigned specified head boundary condition with constant head values of 100 m (A-B and C-D segments) and 60 m (E-F segment), respectively.

Nonspatial random fields of true T and true S (Fig. 4) were

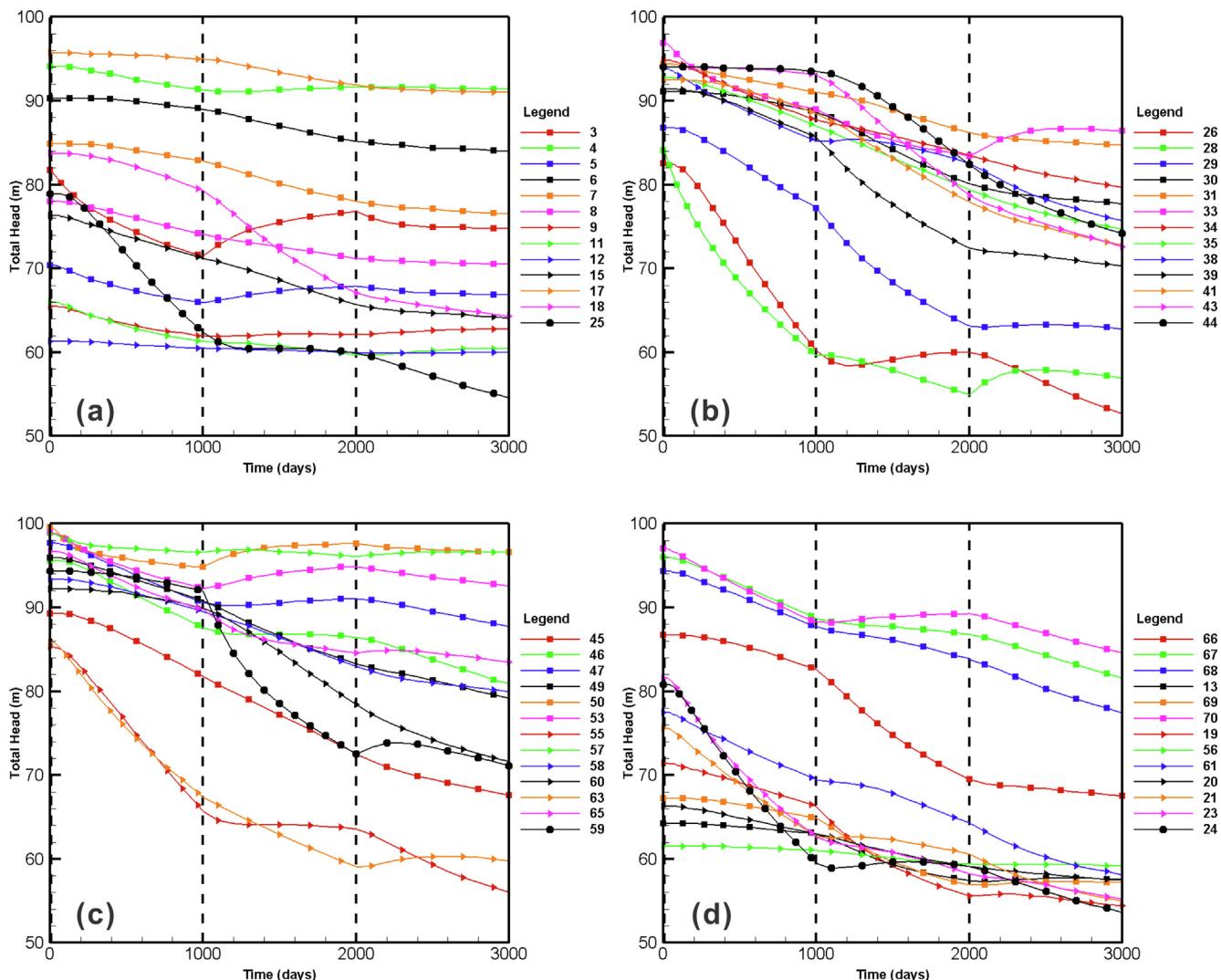


Fig. 6. Hydrographs at 52 monitoring wells in response to pumping events during the simulation period.

generated with four zones, which is analogous to the geologic characteristics of the HLG basin (Fig. 3). For each of the zones, independent random fields of T and S are generated individually using a random field generator. According to actual geological data, different ranges of transmissivity were assigned to the zone 1 (120–240 m^2/d), zone 2 ($< 120 \text{ m}^2/\text{d}$), zone 3 (240–360 m^2/d) and zone 4 (360–480 m^2/d). 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. Transmissivity is assumed to be isotropic at the scale of the model elements ($T_x = T_y$), but both the T and the S fields are assumed to be statistically anisotropic at the scale of the domain. The correlation scales are 20 km in the east-west direction and 5 km in the north-south direction. The random T and S fields are mutually exclusive.

After generating the random T field with these spatial statistics, steady-state responses of the aquifer with no pumping were simulated (Fig. 3), and then the results of the steady-state simulation were used as initial conditions for calculating transient head changes in the aquifer induced by the groundwater exploitation reduction.

The transient simulation stressed the aquifer simultaneously with multiple pumping wells in a manner analogous to municipal pumping systems in the HLG plain. The synthetic well-field is composed of 70 wells: 18 pumping and 52 observation wells. The well layout is the same as those in the real basin. The well numbers, their locations, and

steady-state initial head distribution are shown in Fig. 3.

We simulated three pumping/reduction/shutdown events, which involved different wells, pumping/reduction/shutdown periods, and pumping rates, as listed in Table 1. The pumping rates were designed such that each pumping event stresses most of the aquifer, resulting in the overlap between the stressed areas of different pumping events. The pumping well locations for each event are shown in Fig. 3; triangles are wells pumped in pumping event 1, squares in event 2 and stars in event 3. The simulation lasts 3000 days.

Specifically, for pumping event 1, each of the six wells (Table 1) was pumped at $8000 \text{ m}^3/\text{d}$ for 1000 days, then reduced to $4000 \text{ m}^3/\text{d}$ for 1000 days, and they were shut down for the rest 1000 days. For pumping event 2, six different wells (Table 1) were pumped at $8000 \text{ m}^3/\text{d}$ for 1000 days, then temporarily shut down for 1000 days, and they were pumped at a reduced rate of $4000 \text{ m}^3/\text{d}$ for the remaining 1000 days. For pumping event 3, another six different wells (Table 1) were initially closed for 1000 days, and afterward, they were pumped at $8000 \text{ m}^3/\text{d}$ for 1000 days, and reduced to $4000 \text{ m}^3/\text{d}$ for the rest 1000 days. During each pumping test, head data were collected at 52 observation wells (Fig. 3). These events were designed in such a way to imitate the current situation of groundwater exploitation reduction in the NCP.

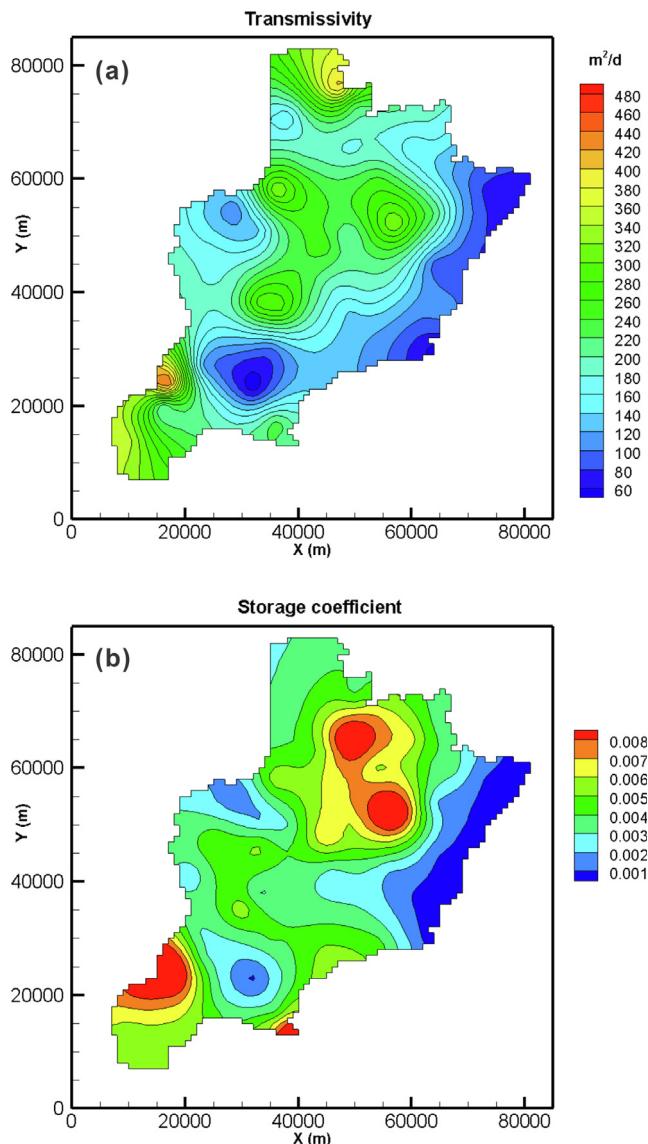


Fig. 7. The kriged fields of transmissivity (a) and storage coefficient (b) in the synthetic domain.

Table 2
Statistical summary for head inversion using different prior information.

Prior means		Uniform mean		Kriged mean		Zonal mean	
Cases		Case 1a	Case 1b	Case 2a	Case 2b	Case 3a	Case 3b
InT	R^2	0.877	0.817	0.876	0.816	0.921	0.864
	L_1	0.124	0.156	0.125	0.158	0.102	0.139
	L_2	0.028	0.042	0.029	0.043	0.018	0.032
	Slope	0.871	0.822	0.884	0.842	0.926	0.901
	Intercept	0.664	0.914	0.601	0.810	0.383	0.508
InS	R^2	0.752	0.745	0.768	0.737	0.932	0.911
	L_1	0.267	0.271	0.253	0.275	0.140	0.162
	L_2	0.127	0.132	0.120	0.136	0.034	0.046
	Slope	0.741	0.739	0.856	0.820	0.931	0.916
	Intercept	-1.422	-1.417	-0.812	-0.995	-0.379	-0.441

4. Results and discussion

4.1. Forward simulation and aquifer response to groundwater-pumping reduction

The simulated groundwater drawdown fields at 1000, 2000, and 3000 days due to the three events are shown in Fig. 5. The drawdown fields show that the changes in the operation of pumping wells, due to the groundwater exploitation reduction program, can stress the most areas of the entire aquifer. The first time frame (0–1000 days) represents the effect of normal pumping operations prior to the implementation of the groundwater exploitation reduction program at the year of 2014. Clearly, the spatial distribution of the drawdown field at 1000 days shows that simultaneous operation of multiple pumping wells can stimulate appreciable drawdown only in the finite regions of the entire aquifer (Fig. 5a). The program of groundwater exploitation reduction is a gradual process, which is reflected in the remaining time intervals (1001–2000 days and 2001–3000 days). The second time frame (1001–2000 days) and the last time frame (2001–3000 days) are the initial and late stage of regional groundwater-pumping reduction, respectively. With the groundwater exploitation reduction, the alteration of pumping wells modifies the basin-scale flow field, and the drawdown fields at 2000 and 3000 days spread throughout the region (Fig. 5b and c), indicating that the entire basin-scale aquifer is stressed due to the operational variations of pumping wells at various locations.

The changes of groundwater levels in fifty-two observation wells are shown in Fig. 6 over the three time intervals (0–1000 days, 1001–2000 days, and 2001–3000 days). According to these well hydrographs, wells respond differently to changes in stimulus caused by varying pumping events. Overall, three patterns of hydraulic head fluctuations (namely, sharp drawdown, slow drawdown and stable or rising limb) are observed during the entire period (Fig. 6). The sharp decline in water level generally occurs in the first and second time intervals when each pumping well is pumped at a normal rate of 8000 m^3/d . When the rate is reduced to 4000 m^3/d , the water level decreases slightly. During the shutdown of pumping wells in different events, the water level exhibits steady or notable rises. The head changes in most of the monitoring wells are sensitive to the change of pumping operations, suggesting a variety of patterns for all the groundwater-pumping reduction, intermittent shutdown, and resumption events.

Such a dissimilarity in the regional flow field suggests that these aquifer responses likely carry information about the hydraulic connections among pumping wells and observation wells (Wen et al., 2020). Therefore, it is fairly obvious that changes in pumping rate can be detected in well signals, and the changes in heads will be the superposition of a variety of different changes to the stresses throughout the aquifer, which is tantamount to a large-scale HT survey.

4.2. HT inversion using different prior information

4.2.1. The role of initial guess in parameter estimations

Given that the estimation of T and S in a highly-parameterized model is sensitive to the initial values of parameters (Illman et al., 2008; Zha et al., 2019), three cases with different initial guesses (uniform, kriged fields and zonation) were used in inverse model. Case 1 represents the situation in which there is little knowledge of spatial patterns of hydraulic properties at a field site. As such, the initial field for the inverse model is assumed to be homogeneous. Given the fact that the zone 1 is dominant feature in the region (Fig. 3), the mean values of T and S for zone 1 were used as the initial guess properties. Case 2 aims to demonstrate the effects of considering direct measurements of T and S from pumping tests and borehole logs on characterizing the region. That is, the hypothetically known T and S values at locations of 52 observation wells were interpolated and extrapolated by kriging method to derive the initial T and S fields for the HT analysis.

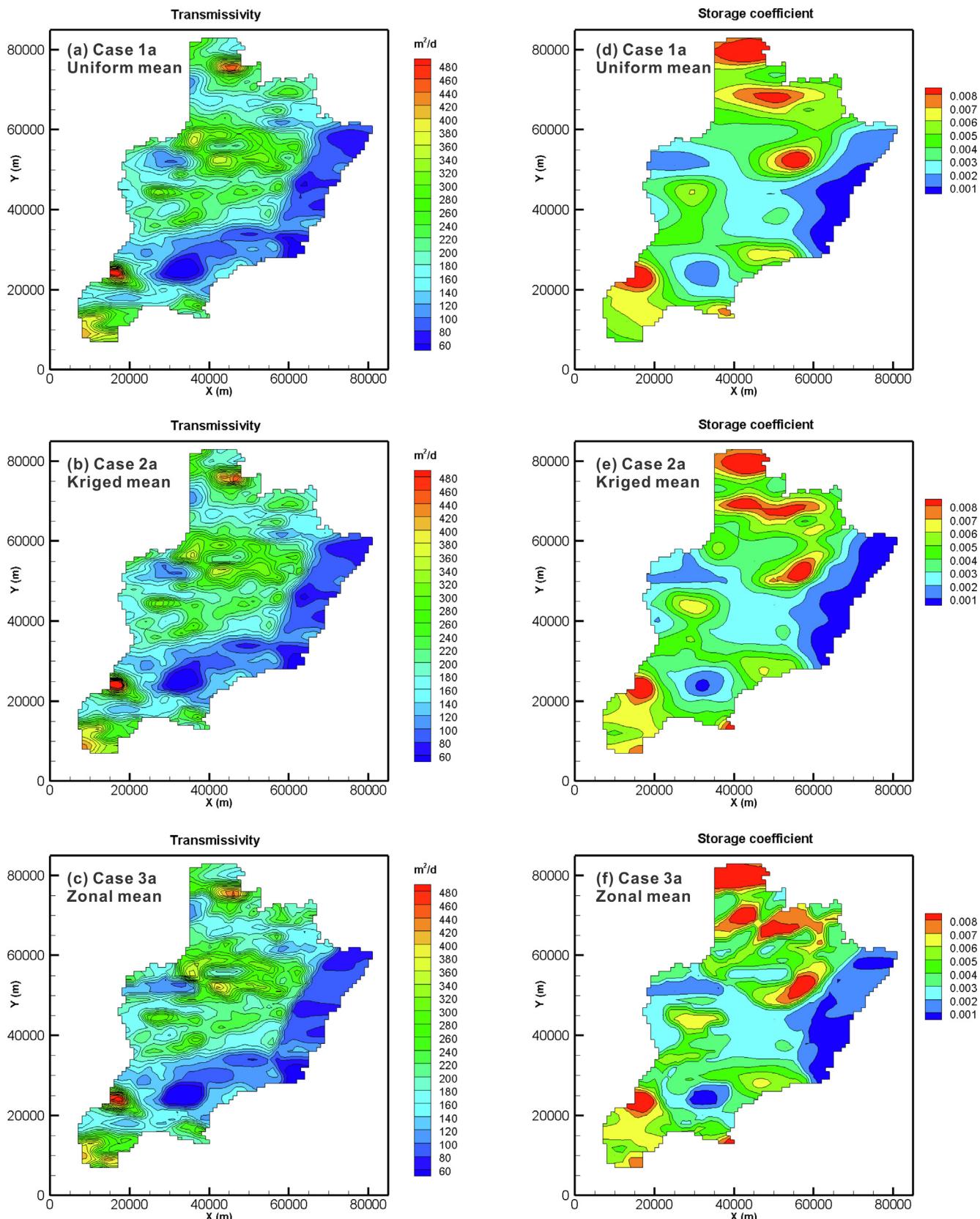


Fig. 8. Transmissivity estimates (a–c) for cases 1a, 2a and 3a, as well as the corresponding storage coefficient estimates (d–f).

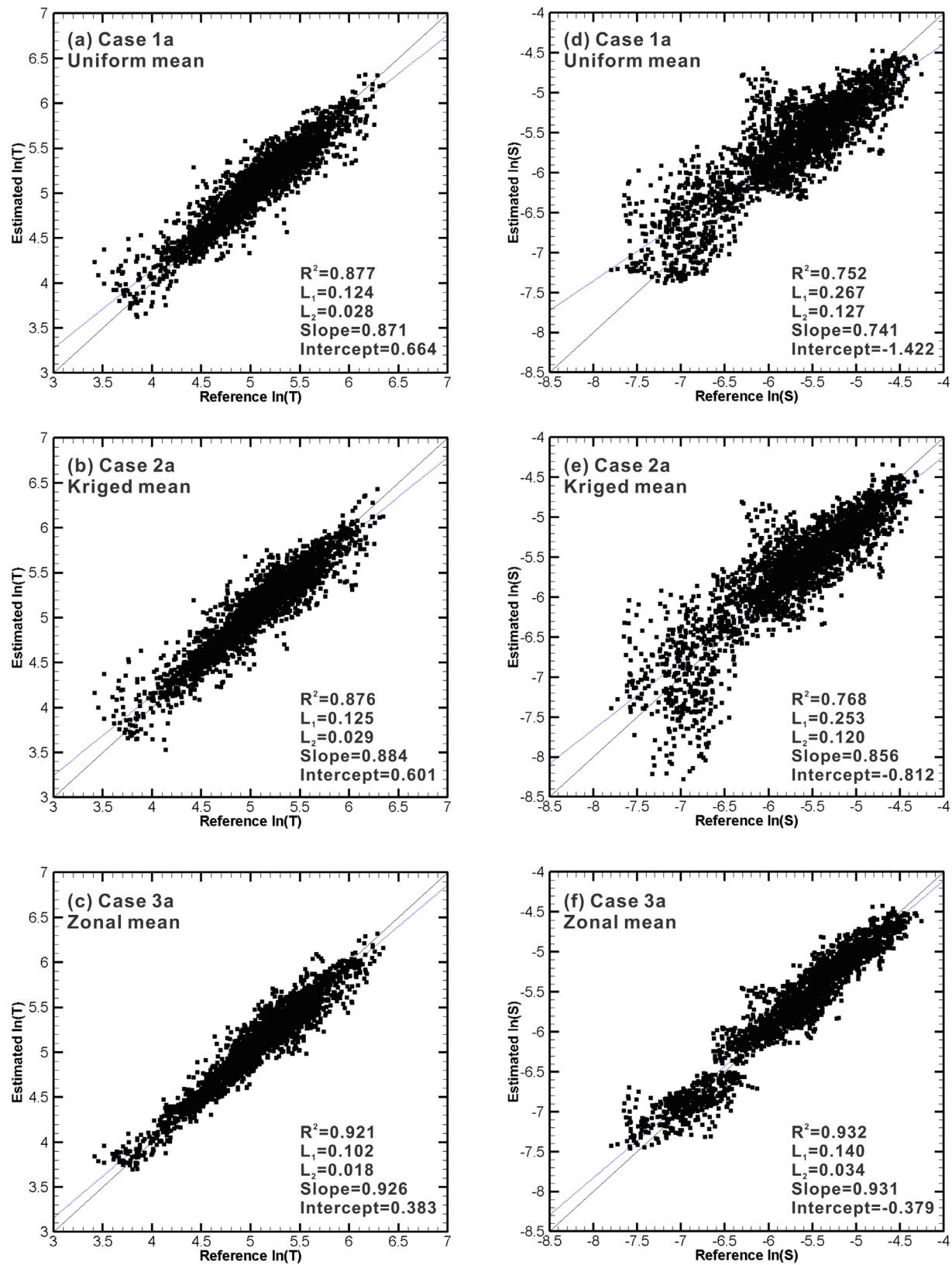


Fig. 9. Scatterplots of reference versus estimated values of the transmissivity (a–c) and storage coefficients (d–f) for Cases 1a, 2a and 3a.

Fig. 7 shows the kriged T and S fields, which have a much more smoothed appearance than the true fields in Fig. 4. The kriged fields only delineate the approximate distribution of high and low permeable zones, but fail to capture the details or accurate geometry of aquifer heterogeneity (Figs. 4 and 7). Case 3 assumes that the geometry and

locations of the four geological zones in the reference fields are known exactly (Fig. 3) via available geological maps and cross sections. For initial guesses of the zonal hydraulic properties, each zone is prescribed with a mean T or S value identical to that of the corresponding zone in the true T or S fields.

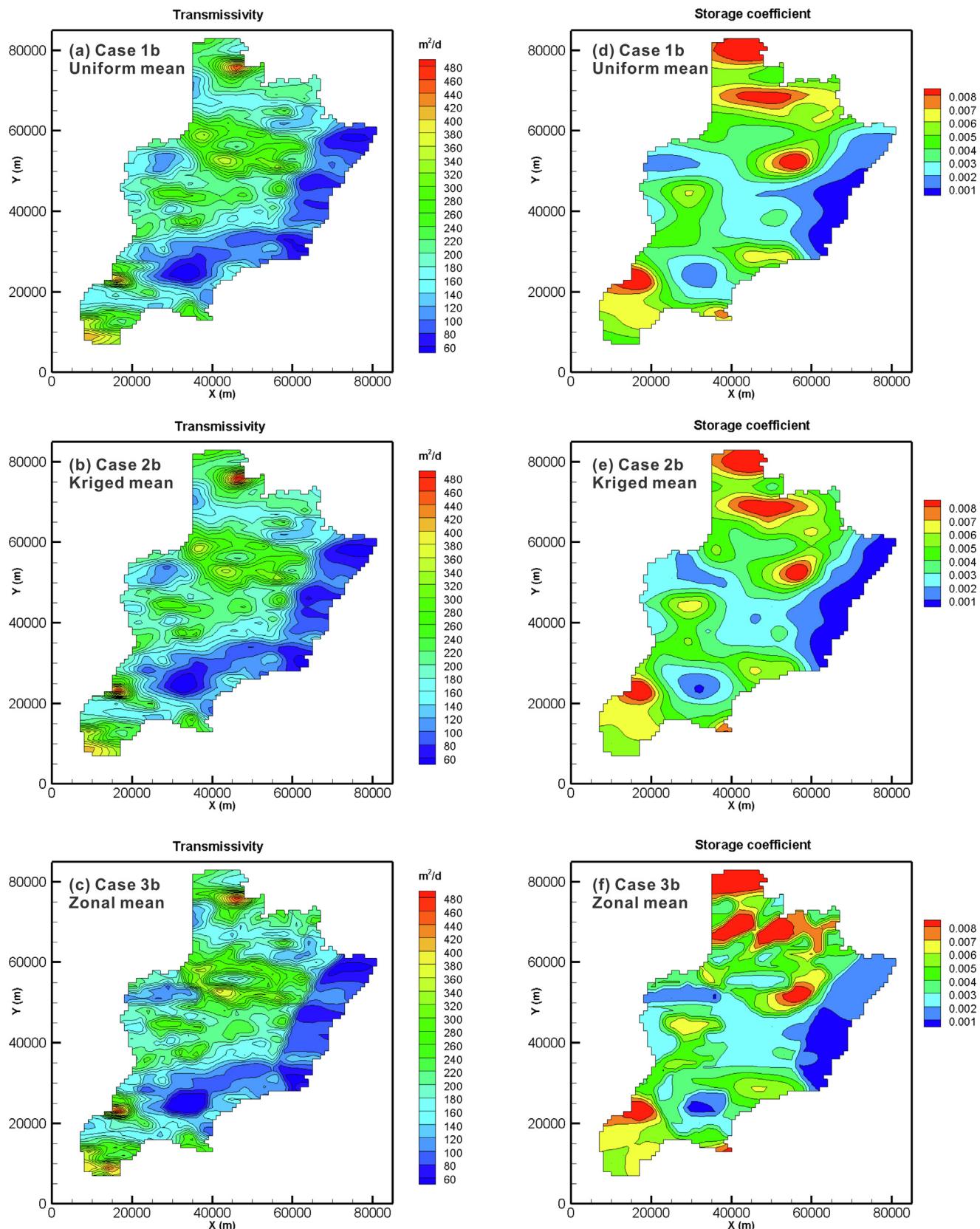


Fig. 10. Transmissivity estimates (a–c) for cases 1b, 2b and 3b, as well as the corresponding storage coefficient estimates (d–f).

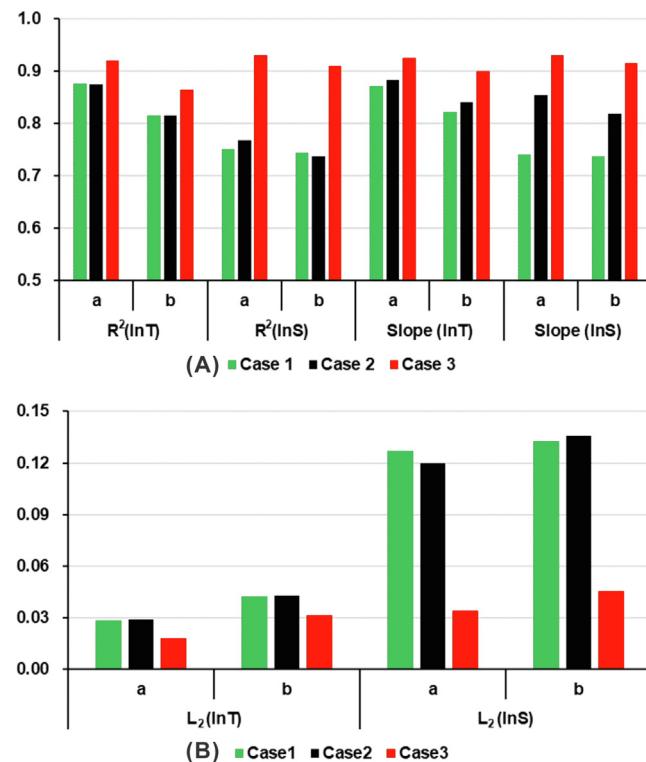


Fig. 11. Comparison of (A) R^2 and slope, (B) L_2 for parameter estimation using different prior information. Here, “a” represents Cases 1a, 2a and 3a using noise-free data, and “b” represents Cases 1b, 2b and 3b using noisy data.

As a baseline, the results of HT inversion using noise-free observation data generated from the forward simulation (Cases 1a, 2a, and 3a) were examined. The estimated fields of parameters, scatterplots and performance metrics of the estimates for the three cases are presented in Table 2, Figs. 8 and 9. For Case 1a, the main features of the heterogeneous aquifer are identified (Fig. 8a), with a correlation coefficient (R^2) of 0.877 (Fig. 9a), indicating that much of the heterogeneous information on T has been extracted. However, the estimated S pattern (Fig. 8d) is much smoother than the true values (Fig. 4b). Some localized high S zones are not identified, and the shape of the low S areas is not accurately delineated (Fig. 8d). The scatterplot for $\ln S$ in Case 1a also exhibits a more dispersed distribution than that for $\ln T$ (Fig. 9a and d). Relative to those values of metrics for $\ln T$, the performance of S estimation is worse with a lower R^2 (0.752), larger L_2 (0.127), smaller slope (0.741).

In Case 2a, the HT inversion started with kriged T and S fields. The estimated T map is similar to that of Case 1a (Fig. 8a and b), with similar values of performance metrics (Fig. 9a and b). However, the results of S estimates are better than those in Case 1a. The outlines of high S and low S zones are closer to the reference field (Fig. 8e). The index (R^2) increases to 0.768 and the slope value also ascends from 0.741 to 0.856, indicative of slightly better performance for S estimation in Case 2a than that in Case 1a.

In comparison with results of Cases 1a and 2a, the prior model with given zonal geometry (Case 3a) yields the best results. With respect to $\ln T$, the R^2 and slope increase to 0.921 and 0.926, while L_2 and intercept reduce to 0.018 and 0.383, respectively (Table 2). For $\ln S$, the R^2 and slope reach 0.932 and 0.931, while L_2 and intercept reduce to 0.034 and -0.379, respectively (Table 2). Case 3a yields a more accurate characterization of hydraulic heterogeneity relative to the reference fields, as depicted in Fig. 8c and f. Additionally, the scattered points for Case 3a (Fig. 9c and f) are closely around the 1:1 line, providing direct evidence for the excellent performance of the inverse model starting with zonal T and S means.

In summary, the HT inversion using a known stratigraphic zonation model can produce more detailed information of aquifer heterogeneity. However, the inversion of the zonation model based on inaccurate geological information may yield worse inversion results (Luo et al., 2017; Zhao et al., 2016). Therefore, using a hydraulic zonation model with accurate geological information as the prior information of HT analysis may be a preferred choice for the HT inversion, given that it can improve the delineation of high and low T or S zones.

4.2.2. Effects of noisy observations

The previous HT inversion was based on noise-free observations. However, observed well hydrographs in the field likely contain noisy signals unexplained by the synthetic model apart from effects of heterogeneity (such as measurement errors and other sources or sinks including precipitation, evaporation, lateral flow, leakage, etc.). In this study, unbiased Gaussian noises with a standard deviation of 0.1 m were superimposed onto the simulated hydrographs in order to mimic more realistic scenarios. Apart from the corruption of the observation data, the other setup of the inverse models remains the same as that in the noise-free analysis.

The estimated T and S tomograms from the Cases 1b, 2b, and 3b using noisy signals for inverse modeling are shown in Fig. 10. In terms of T tomograms, the spatial patterns of the main high and low permeable areas are mapped (Fig. 10a–c), suggesting that the estimated T tomograms using noisy head data are still comparable to the true T fields (Fig. 4a). Moreover, it is noteworthy that the localized shapes of T tomograms in the Case 3b are more accurately characterized in comparison to the inversion results in the Cases 1b and 2b. With respect to S tomograms, the estimated S patterns (Fig. 10d and e) from the Cases 1b and 2b are smoother than the true fields (Fig. 4b). The geometric shapes of the high and low S areas for the Cases 1b and 2b are not accurately captured by comparing with the true fields (Fig. 4b), indicating loss of some heterogeneous information. The S tomogram is noticeably improved in the Case 3b (Fig. 10f). Overall, in noisy data analysis, the HT inversion using an accurate zonation model can still reproduce T and S tomograms closer to the true fields.

To better assess the effects of noisy observations on the characterization of aquifer heterogeneity, the bar charts of the main performance metrics (R^2 , L_2 and slope) for all cases are plotted in Fig. 11. These bar charts illustrate that the estimates using noisy well hydrographs (Cases 1b, 2b and 3b) are inferior to those using noise-free hydrographs (Cases 1a, 2a and 3a). Although using the noisy signals in HT inversion results in the loss of some heterogeneous information, the main subareas of hydraulic properties are still captured (Figs. 10 and 11). Using zonal means as prior guessed values for inversion (Case 3b) also yields the best estimates of T and S among the Cases 1b, 2b, and 3b in noisy data cases (Fig. 11). Thus, for basin-scale aquifer characterization, we recommend the use of zonal means with accurate geological information as initial guesses for the inversion of parameters (Luo et al., 2017; Zhao et al., 2016).

4.3. Predictions of solute transport with the reconstructed heterogeneity

The spatial variation in transmissivity (T) plays a key role in the spreading of solute plumes (Ni et al., 2009). Therefore, it is necessary to assess the performance of the reconstructed hydraulic parameter fields with different precision for predicting solute transport. Here, four scenarios with different parameter fields are considered, namely homogeneous, zonal mean, kriged fields and HT inversion results. The homogeneous fields of T and S are assumed to be the mean values of T and S for zone 1, given the dominance of zone 1 in the region (Fig. 3). For the zonal mean fields, the average values of T and S are assigned to each zone of the domain, which is also widely used in the groundwater models (Liu et al., 2018; Wen et al., 2007). The resolution of the kriged fields is higher than that of the zonal mean fields, because direct measurements of the aquifer parameters are considered. Additionally,

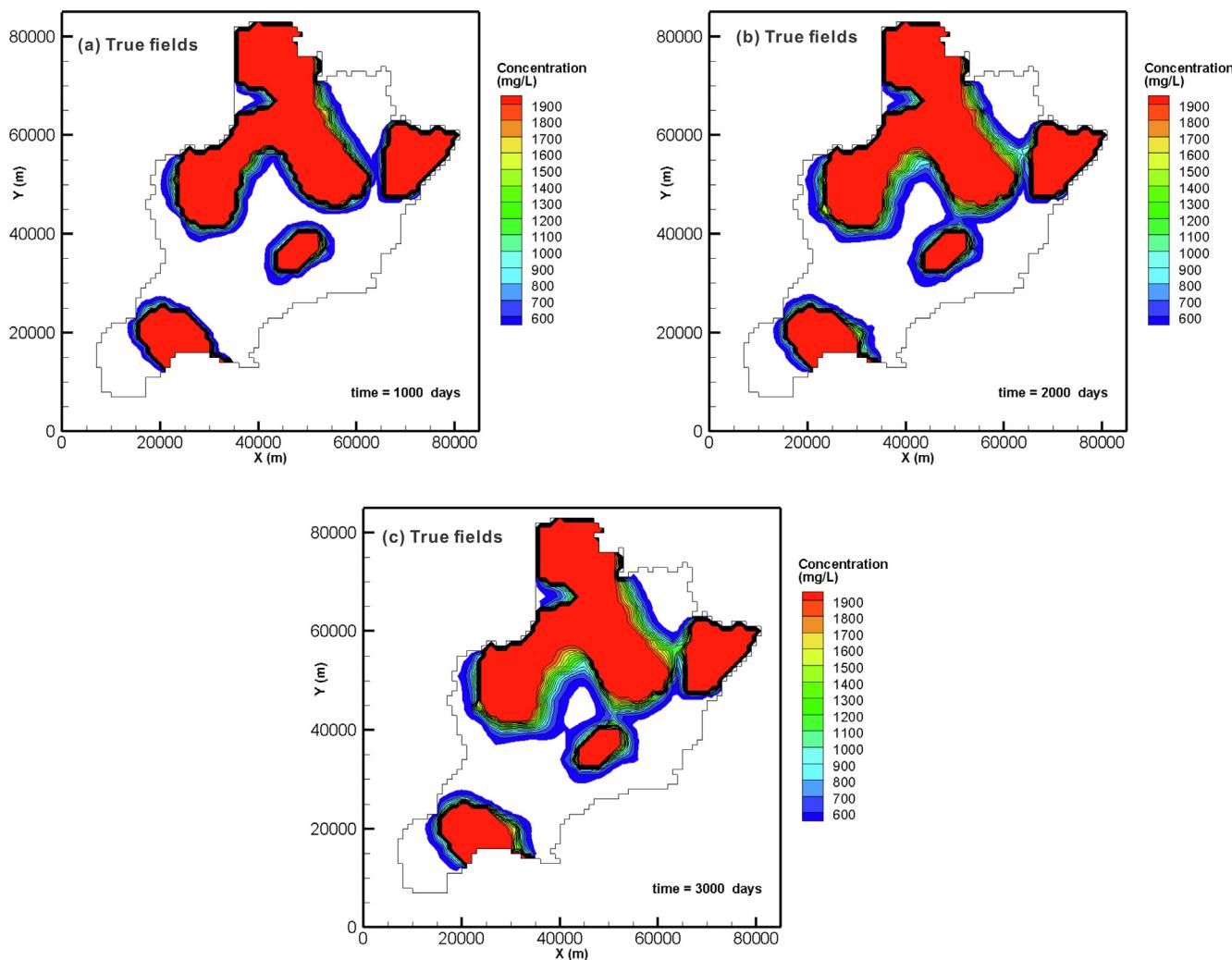


Fig. 12. Spatial snapshots of solute plume evolution using true fields in solute transport model at selected time 1000 (a), 2000 (b) and 3000 days (c).

the T and S estimates of HT inversion in Case 3a are used to predict the spreading behavior of plume, which is anticipated to provide the highest resolution of the parameter fields.

As mentioned previously, downward leakage from upper aquifer with high TDS, creates threats to the confined aquifer. Therefore, the intrusion of the overlying saline water into the confined aquifer was considered to be a meaningful and realistic scenario for solute transport prediction.

The plume source zone was assumed to be the areas of the phreatic water with $TDS > 2 \text{ g/L}$, namely the yellow, orange and red zones in Fig. 1. The leakage rate was assumed to be 0.001 m/d , with a TDS concentration of 2000 mg/L . For solute transport simulation, all nodes were assigned an initial TDS concentration of 500 mg/L , corresponding to the mean background concentration. The boundary conditions and pumping events, as well as the simulated time (3000 days), remained the same as those in the Case 3a. For comparing the solute transport simulation results based on different parameter fields, the 52 monitoring wells (Fig. 3) were used for monitoring the concentrations of TDS.

It should be noted that the transmissivity heterogeneity plays a much more important role in the transport behaviors of solute than the dispersivity (Ni et al., 2009), given the insensitivity to longitudinal dispersion coefficient for large-scale transport simulations (Sudicky et al., 2010). For transport parameters, the longitudinal and the transverse dispersivities were assumed to be 1000 and 100 m , respectively. An average porosity of 0.4 was selected for the entire domain.

The porosity value was assumed constant over the entire domain, given that it may be less likely that depth-averaged porosity varies by a factor of 2–4 within any individual grid cell at this basin-scale model, based on a small $\ln T$ variance of 0.1 .

Spatial snapshots at selected time (1000, 2000 and 3000 days) of solute transport model show that the plume size varies over time (Fig. 12). Relative to the initial plume size with 1038 km^2 , the area of solute plume increases by 266 km^2 at 1000 days (Fig. 12a) and 422 km^2 at 2000 days (Fig. 12b), respectively. Lastly, the contaminant plume size reaches 1540 km^2 at 3000 days (Fig. 12c). Given the importance of groundwater flow velocity in the solute transport prediction, the actual velocity fields (V_x and V_y) were compared with simulated velocity fields using different parameter fields (Fig. 13). Clearly, the parameter fields from HT inversion produce the best predicted velocity components analogous to those actual ones (Fig. 13a and b). The predicted velocity components using kriged fields (Fig. 13c and d) are inferior to those from HT estimation, while they are slightly better than those using zonal mean fields (Fig. 13e and f). The simulated velocity components using homogeneous fields present the lowest accuracy (Fig. 13g and h), in comparison with other scenarios.

Additionally, the breakthrough curves (BTCs) obtained at four monitoring sites (wells 19, 23, 67, and 70) with various distances from source zones, were selected to make a comparison of simulated results of different parameter resolutions (Fig. 14). With respect to the true BTC in the well 19 with approximate 1 km and 2 km away from the periphery of the northeast plume source and the west plume source of

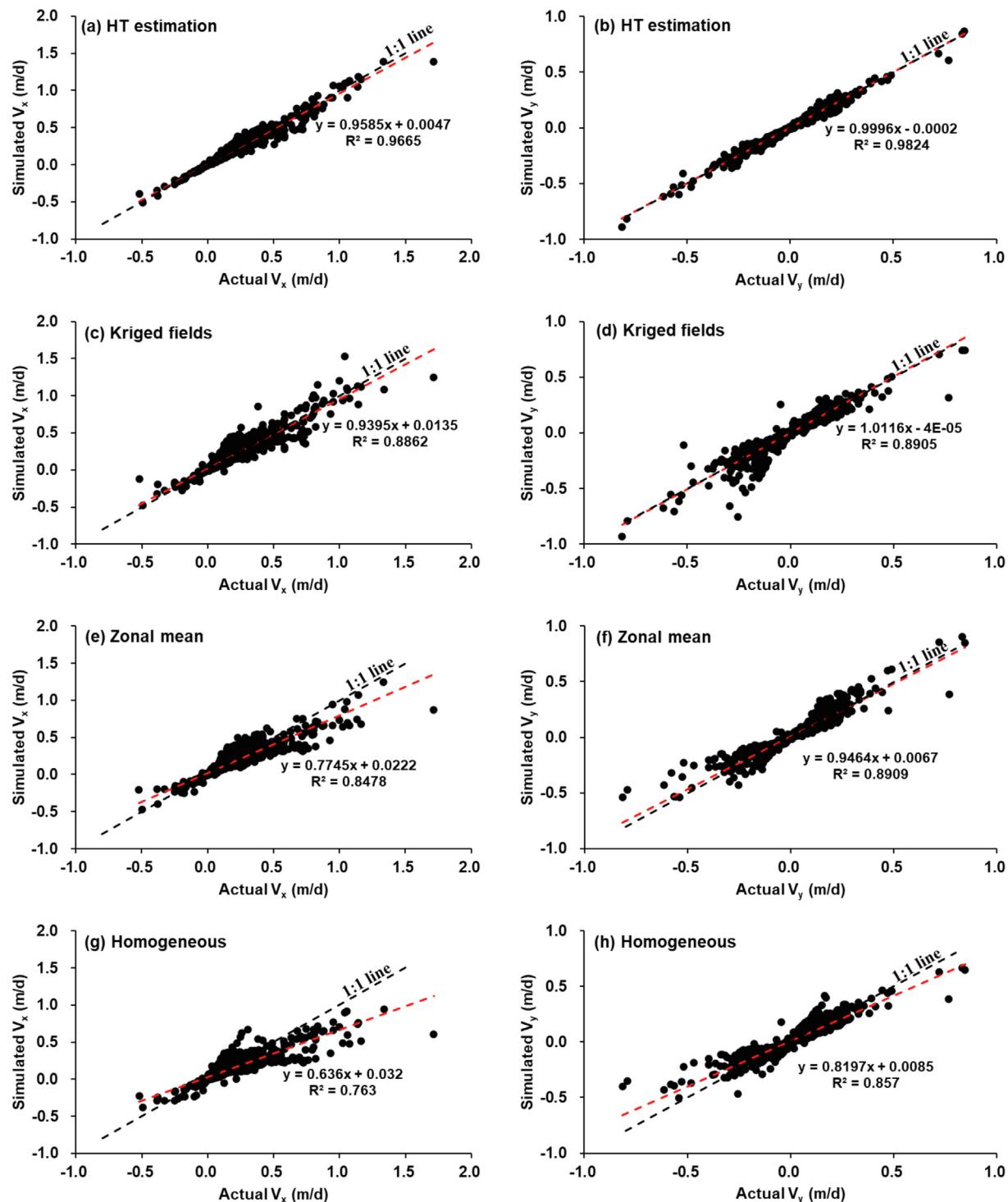


Fig. 13. Scatterplots of actual velocity components (V_x and V_y) versus simulated velocity components from HT estimation (a-b), kriged fields (c-d), zonal mean (e-f) and homogeneous fields (g-h).

well 19, nonuniformities can be identified from the curve (Fig. 14a). The BTC curve initially increases with a lower slope from the mean concentration, suggesting that the small solute flux is likely a result of dispersion from the northeast plume source at the initial stage. After about 1300 days, the TDS values increase at a faster rate, indicating additional contribution from the plume source west of this well. The superimposed effects of both two plume sources may be mainly responsible for the accelerated TDS increment. Besides, the initial arrival time is about 1200 days for the well 23 at a distance of 4 km away from the nearest point in the source zone. Thereafter, the TDS values

immediately start with a steep increase of concentration from the mean background concentration (500 mg/L), followed by gradually reduced slopes during the period from 2000 to 3000 days, and finally reach a relatively steady-state with a maximum concentration of 520 mg/L (Fig. 14b). The true BTCs from wells 67 and 70 also follow a typical shape resulting from continuous input of solutes, and vary in the arrival time and the maximum concentration (Fig. 14c and d).

In this study, the main concern is whether the reconstructed heterogeneity is accurate enough for predicting solute transport behaviors. The simulated results confirm that the parameter fields estimated from

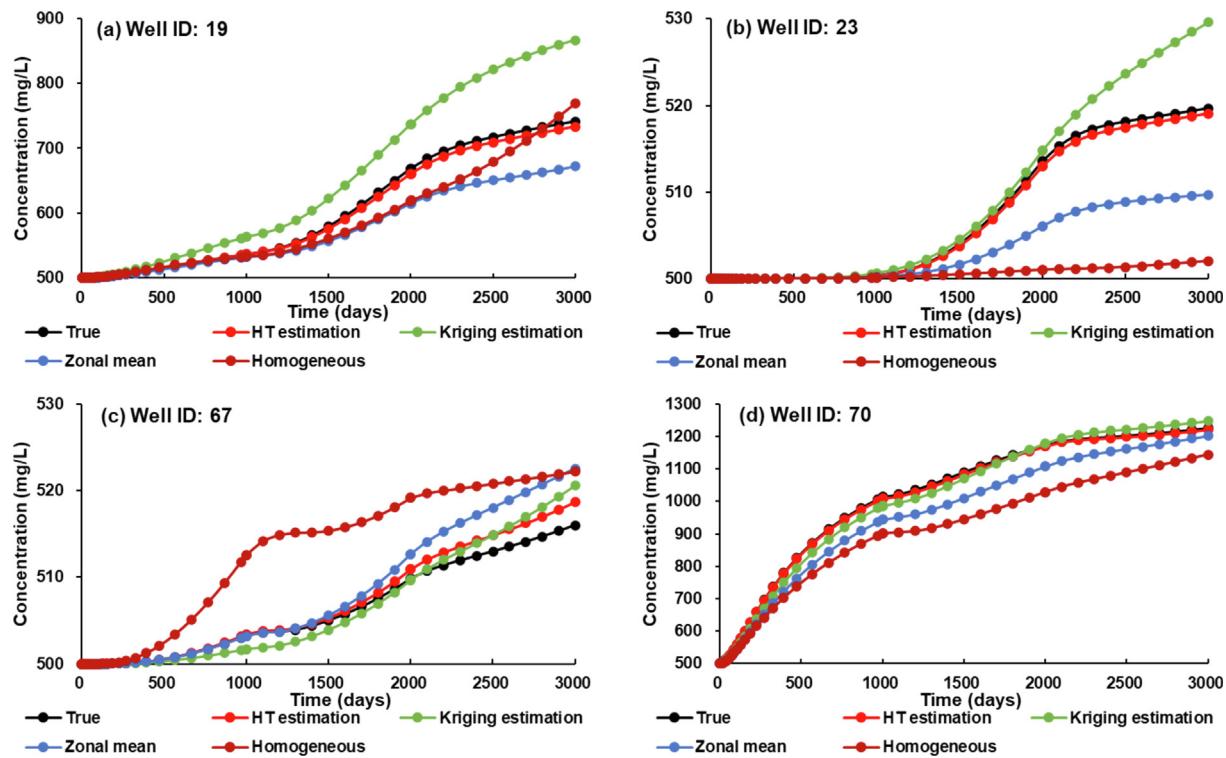


Fig. 14. Comparison of true BTCs and simulated BTCs at the selected monitoring wells.

HT inversion can yield the best prediction of BTCs at these locations. For most of the observation wells, the BTC peaks and trends can be accurately captured by using the estimated T and S values from HT inversion.

The BTCs based on kriged fields show imprecise predictions of BTCs and overestimations of peaks. Note that the kriged T and S fields were generated from 52 direct measurements that were assumed to be known. However, in most practical issues, information on hydraulic parameters is generally very limited given the huge cost for conducting a large number of pumping tests in the field. Therefore, it is an expensive way to acquire so many in-situ parameters in the real world, and the prediction of BTCs based on the kriged fields is expected to be even lower accuracy than that shown in this study.

The simulated BTCs based on zonal mean fields exhibit similar shapes of BTCs analogous to true ones but show inaccurate peak values. Relative to the true peak values, the peak value for BTC in the well 19 is underestimated by 70 mg/L, while the peak for BTC in the well 23 is underestimated by 10 mg/L, respectively. The underestimations of peak values decrease with increasing monitoring distances from the plume source zone. This may be attributed to velocity variation caused by the difference in aquifer parameters between zonal mean and true fields, since groundwater velocity controls the advective transport of solutes (Kuhlman et al., 2008).

Lastly, the homogeneous model yields incorrect shapes of BTCs at selected locations, and these curves fail to capture the true trend of solute concentration over time, which demonstrates the contribution of heterogeneous fields to improving the accuracy of the solute transport curves.

5. Conclusions

A basin-scale aquifer characterization approach based on the hydraulic tomography concept is proposed. It takes advantage of readily available aquifer signals generated from operational variations of existing pumping and monitoring wells, without additional cost to estimate the hydraulic heterogeneity over the whole basin. Specifically, the

basin-scale flow field modified by the alternation, intermittent shutdown and resumption of different pumping wells is tantamount to a basin-scale HT survey.

Because of the national policy in China, such as the groundwater exploitation reduction program, exploiting the groundwater data to resolve the basin-scale aquifer characterization problem becomes logical and feasible. In the North China Plain, abundant well hydrograph and pumping operations data over many years are readily available. These datasets may not provide a great deal of basin-scale information individually, but when they are combined as a whole, they can provide more information about aquifer heterogeneity than the sum of the individual parts. That is, aquifer responses to the change of pumping behaviors and locations of pumping, are likely to carry non-redundant information on the basin-scale heterogeneity.

This study finds that the basin-scale HT parameter estimation is sensitive to initial guesses for inverse models. In particular, compared with the HT estimates using uniform and kriged T and S fields, using accurate mean values of T and S for geological zones as prior information can yield the best estimates of T and S in noise-free and noisy data analysis. However, using the noisy signals in HT inversion results in the loss of some heterogeneous information, but the inversion captures the main distributed features of hydraulic properties.

The estimated T and S fields from the proposed approach yield the best prediction of the groundwater flow velocities and BTCs. The BTCs based on kriged and zonal mean fields show inaccurate peak values. On the other hand, the BTCs derived from homogeneous fields fail to capture the true trend of solute concentration over time.

Findings of this study suggest that the long-term head data, at as many locations as possible, in response to varying locations and behaviors of pumping wells should be integrated as a whole for HT analysis. We also recommend the use of the reasonably distributed mean properties as prior information if geologic knowledge about the heterogeneity is available.

Lastly, the results of this study support the call by Yeh and Lee (2008) "Time to change the way we collect and analyze data for aquifer characterization." Although the application of this concept to field

problems remains to be tested, the results of this study demonstrate the feasibility of using aquifer responses from operational variations of existing pumping wells to improve the characterization of basin-scale subsurface heterogeneity. The data fusion of head, flux, and chemical tracers (Tso et al., 2016; Zha et al., 2016) could be a promising perspective for further enhancement of the basin-scale aquifer characterization.

CRediT authorship contribution statement

Fei Liu: Conceptualization, Formal analysis, Writing - original draft, Writing - review & editing. **Tian-Chyi Jim Yeh:** Methodology, Writing - original draft, Writing - review & editing, Supervision. **Yu-Li Wang:** Writing - original draft. **Yonghong Hao:** Writing - review & editing. **Jet-Chau Wen:** Writing - review & editing. **Wenke Wang:** Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research was supported by the National Natural Science Foundation of China (Grant No. 41901039); the China Scholarship Council (Grant No. 201808130026); the Natural Science Foundation of Hebei Province (Grant No. D2019402045); the Department of Education of Hebei Province (Grant No. QN2018076). The authors are grateful to our colleagues and graduates for their assistance in data collection and field investigation. 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.

References

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

Bredhehoef, J.D., 2003. From models to performance assessment: the conceptualization problem. *Groundwater* 41 (5), 571–577. <https://doi.org/10.1111/j.1745-6584.2003.tb02395.x>.

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 Resour. Res.* 48 (5). <https://doi.org/10.1029/2011wr011704>.

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

Cardiff, M., Bakhos, T., Kitanidis, P.K., Barrash, W., 2013a. Aquifer heterogeneity characterization with oscillatory pumping: sensitivity analysis and imaging potential. *Water Resour. Res.* 49 (9), 5395–5410. <https://doi.org/10.1002/wrcr.20356>.

Doro, K.O., Cirkpa, O.A., Leven, C., 2015. Tracer tomography: design concepts and field experiments using heat as a tracer. *Ground Water* 53 (Suppl 1), 139–148. <https://doi.org/10.1111/gwat.12299>.

Fakhreddine, S., Lee, J., Kitanidis, P.K., Fendorf, S., Rolle, M., 2016. Imaging geochemical heterogeneities using inverse reactive transport modeling: an example relevant for characterizing arsenic mobilization and distribution. *Adv. Water Resour.* 88, 186–197. <https://doi.org/10.1016/j.advwatres.2015.12.005>.

Fischer, P., et al., 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>.

Fitterman, D.V., 1990. Developments and applications of modern airborne electromagnetic surveys : proceedings of the U.S. Geological Survey Workshop on Developments and Applications of Modern Airborne Electromagnetic Surveys, October 7–9, 1987, 1925, doi: 10.3133/b1925.

Guillemoteau, J., Sailiac, P., Béhaegel, M., 2011. Regularization strategy for the layered inversion of airborne transient electromagnetic data: application to in-loop data acquired over the basin of Franceville (Gabon). *Geophys. Prospect.* 59 (6), 1132–1143. <https://doi.org/10.1111/j.1365-2478.2011.00990.x>.

Hochstetler, D.L., et al., 2016. Hydraulic tomography: continuity and discontinuity of high-K and low-K zones. *Ground Water* 54 (2), 171–185. <https://doi.org/10.1111/gwat.12344>.

Illman, W.A., 2014. Hydraulic tomography offers improved imaging of heterogeneity in fractured rocks. *Ground Water* 52 (5), 659–684. <https://doi.org/10.1111/gwat.12119>.

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. *J. Hydrol.* 341 (3–4), 222–234. <https://doi.org/10.1016/j.jhydrol.2007.05.011>.

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., 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 (22), 8609–8614.

Illman, W.A., Berg, S.J., Yeh, T.C., 2012. Comparison of approaches for predicting solute transport: sandbox experiments. *Ground Water* 50 (3), 421–431. <https://doi.org/10.1111/j.1745-6584.2011.00859.x>.

Jiménez, S., et al., 2015. Prediction of solute transport in a heterogeneous aquifer utilizing hydraulic conductivity and specific storage tomograms. *Water Resour. Res.* 51 (7), 5504–5520.

Kitanidis, P.K., Vomvoris, E.G., 1983. A geostatistical approach to the inverse problem in groundwater modeling (steady state) and one-dimensional simulations. *Water Resour. Res.* 19 (3), 677–690.

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>.

Liu, et al., 2002. Effectiveness of hydraulic tomography: sandbox experiments. *Water Resour. Res.* 38 (4), 5–1–5–9.

Liu, F., et al., 2018. Predicting the impact of heavy groundwater pumping on groundwater and ecological environment in the Subei Lake basin, Ordos energy base, Northwestern China. *Hydrol. Res.* 49 (4), 1156–1171. <https://doi.org/10.2166/nh.2017.281>.

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 (5). <https://doi.org/10.1029/2006wr005144>.

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. *J. Hydrol.* 554, 758–779. <https://doi.org/10.1016/j.jhydrol.2017.09.045>.

Ni, C.-F., Yeh, T.-C.J., 2008. Stochastic inversion of pneumatic cross-hole tests and barometric pressure fluctuations in heterogeneous unsaturated formations. *Adv. Water Resour.* 31 (12), 1708–1718. <https://doi.org/10.1016/j.advwatres.2008.08.007>.

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 (10), 3720–3727.

Paine, J.G., Collins, E.W., 2003. Applying airborne electromagnetic induction in groundwater salinization and resource studies, West Texas. In: Symposium on the Application of Geophysics to Engineering and Environmental Problems, pp. 722–738. doi: 10.4133/1.2923218.

Palacky, G.J., 1993. Use of airborne electromagnetic methods for resource mapping. *Adv. Space Res.* 13 (11), 5–14. [https://doi.org/10.1016/0273-1177\(93\)90196-I](https://doi.org/10.1016/0273-1177(93)90196-I).

Pujol, J., 2007. The solution of nonlinear inverse problems and the Levenberg–Marquardt method. *Geophysics* 72 (4), W1–W16. <https://doi.org/10.1190/1.2732552>.

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 (7). <https://doi.org/10.1029/2006wr005287>.

Sudicky, E.A., Illman, W.A., Goltz, I.K., Adams, J.J., McLaren, R.G., 2010. Heterogeneity in hydraulic conductivity and its role on the macroscale transport of a solute plume: from measurements to a practical application of stochastic flow and transport theory. *Water Resour. Res.* 46 (1). <https://doi.org/10.1029/2008wr007558>.

Sun, R., et al., 2013. A temporal sampling strategy for hydraulic tomography analysis. *Water Resour. Res.* 49 (7), 3881–3896. <https://doi.org/10.1002/wrcr.20337>.

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 (1), 3–20. <https://doi.org/10.1002/2015wr017191>.

Wang, Y.-L., et al., 2017. Characterizing subsurface hydraulic heterogeneity of alluvial fan using riverstage fluctuations. *J. Hydrol.* 547, 650–663. <https://doi.org/10.1016/j.jhydrol.2017.02.032>.

Wang, B., 2011. Study on comprehensive quality assessment and coupling-coordination relationship of water-land resources in Heilonggang District, Chin. Acad. Geol. Sci. (in Chinese).

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., et al., 2020. Redundant and nonredundant information for model calibration or hydraulic tomography. *Groundwater* 58 (1), 79–92. <https://doi.org/10.1111/gwat.12879>.

Wen, X.H., Wu, Y.Q., Lee, L.J.E., Su, J.P., Wu, J., 2007. Groundwater flow modeling in the Zhangye Basin, Northwestern China. *Environ. Geol.* 53, 77–84. <https://doi.org/10.1007/s00254-006-0620-7>.

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 (2). <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., et al., 2008. A view toward the future of subsurface characterization: CAT

scanning groundwater basins. *Water Resour. Res.* 44 (3). <https://doi.org/10.1029/2007wr006375>.

Yeh, T.-C.J., et al., 2009. River stage tomography: a new approach for characterizing groundwater basins. *Water Resour. Res.* 45 (5). <https://doi.org/10.1029/2008wr007233>.

Yeh, T.-C.J., et al., 2014. Why hydraulic tomography works? *Ground Water* 52 (2), 168–172. <https://doi.org/10.1111/gwat.12129>.

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., Liu, S., 2000. Hydraulic tomography: development of a new aquifer test method. *Water Resour. Res.* 36 (8), 2095–2105. <https://doi.org/10.1029/2000wr900114>.

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 (1), 85–92.

Yin, C., Zhang, B., Liu, Y., Cai, J., 2016. A goal-oriented adaptive finite-element method for 3D scattered airborne electromagnetic method modeling. *Geophysics* 81 (5), E337–E346. <https://doi.org/10.1190/geo2015-0580.1>.

Zha, Y., et al., 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>.

Zha, Y., et al., 2019. Exploitation of pump-and-treat remediation systems for characterization of hydraulic heterogeneity. *J. Hydrol.* 573, 324–340. <https://doi.org/10.1016/j.jhydrol.2019.03.089>.

Zhao, Y., et al., 2017. Energy reduction effect of the south-to-north water diversion project in China. *Sci. Rep.-Uk* 7 (1), 15956. <https://doi.org/10.1038/s41598-017-16157-z>.

Zhao, Z., Illman, W.A., Berg, S.J., 2016. On the importance of geological data for hydraulic tomography analysis: laboratory sandbox study. *J. Hydrol.* 542, 156–171.

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>.

Zhou, W., 2008. Dynamic Simulation of Groundwater in Handan Plain Area, Hebei University of Engineering (in Chinese).

Zhu, J., Yeh, T.-C.J., 2005. Characterization of aquifer heterogeneity using transient hydraulic tomography. *Water Resour. Res.* 41 (7). <https://doi.org/10.1029/2004wr003790>.